Data-driven learning how oncogenic gene expression locally alters heterocellular networks

David J. Klinke II[1-3] ✉, Audry Fernandez[2,3], Wentao Deng[2,3], Habibolla Latifizadeh[4], and Anika C. Pirkey[1]

[1]Department of Chemical and Biomedical Engineering, West Virginia University, Morgantown, WV
[2]Department of Microbiology, Immunology and Cell Biology, West Virginia University, Morgantown, WV
[3]WVU Cancer Institute, West Virginia University, Morgantown, WV
[4]Department of Mathematics, West Virginia University, Morgantown, WV

Discovering and developing pharmaceutical drugs increasingly relies on mechanistic mathematical modeling and simulation. In immuno-oncology, models that capture causal relations among genetic drivers of oncogenesis, functional plasticity, and host immunity provide an important complement to wet experiments, given the cellular complexity and dynamics within tumors. Unfortunately, formulating such mechanistic cell-level models currently relies on hand curation by experts, which can bias how data is interpreted or the priority of drug targets. In modeling molecular-level networks, rules and algorithms have been developed to limit a priori biases in formulating mechanistic models. To realize an equivalent approach for cell-level networks, we combined digital cytometry with Bayesian network inference to generate causal models that link an increase in gene expression associated with oncogenesis with alterations in stromal and immune cell subsets directly from bulk transcriptomic datasets. To illustrate, we predicted how an increase in expression of Cell Communication Network factor 4 (CCN4/WISP1) altered the tumor microenvironment using data from patients diagnosed with breast cancer and melanoma. Digital cytometry and network inference predictions were then tested using two immunocompetent mouse models for melanoma, which provided consistent experimental results.

Heterocellular networks | digital cytometry | deconvolution | anti-tumor immunity | Bayesian network inference | functional plasticity

Correspondence: ***david.klinke@mail.wvu.edu***

Tissues are dynamic structures where different cell types organize to maintain function in a changing environment. For instance, the mammary epithelium reorganizes during distinct stages of the ovarian cycle in preparation for lactation (1). At the same time, immune cells clear dead cells and defend against pathogens present in the tissue microenvironment. Ultimately, the number and functional orientation of different cell types within a tissue interact to create a network, that is a heterocellular network. This heterocellular network is essential for creating and maintaining tissue homeostasis. While we know that tissue homeostasis is disrupted during oncogenesis, our understanding of how genetic alterations quantitatively and dynamically influence the heterocellular network within malignant tissues in humans is not well developed despite large efforts, like The Cancer Genome Atlas (TCGA), to characterize the genomic and transcriptomic landscape in human malignancy (2, 3). In parallel with these large scale data gathering efforts, two informatic developments, namely digital cytometry and Bayesian network inference, may be helpful in interrogating these datasets and are summarized in the next paragraphs.

In cytometry, single-cell sequencing technology elicits a lot of excitement as it enables unbiased discovery of novel cell subsets in particular disease states (4, 5). Unfortunately, persistent challenges related to confounding of batch effects with biological replicates limit the statistical power of these datasets to link oncogenic transcriptional changes with re-organization of the cellular network (6, 7). Due to the high number of biological replicates, transcriptomic datasets, such as the Cancer Genome Atlas, provide a rich resource in characterizing the heterogeneity of oncogenic transformation. Yet, these data were obtained from homogenized tissue samples and reflect the expression of genes averaged across a heterogeneous cell population. Computationally, "Digital Cytometry" can deconvolute the prevalence of individual cell types present within a mixed cell population (8). The approach stems from the idea that the influx of a particular cell subset into a tissue corresponds to an increase in a gene signature uniquely associated with this particular cell subset (9–12). Gene signatures of immune cells have been developed in a number of studies, which increasingly leverage scRNAseq data and machine-learning methods (13–16). Besides representing different cellular subsets, gene signatures can also represent intracellular processes associated with oncogenesis, like the epithelial-mesenchymal transition (17–21). Though, the predictive value of many of these tissue "features" in inferring how heterocellular networks are altered in diseased tissues remain unclear, as establishing correlations among features tends to be the end point of studies (e.g., (19, 22, 23)).

Besides facilitating data acquisition, improved computational power has also enabled probabilistic inference methods that identify relationships within biological datasets that could not be observed using simpler statistical techniques (24, 25). These relationships can be depicted as a graph, where each node represents a random variable, or "feature", and an edge represents a direct relationship between two variables. When the direction of influence can be inferred for an edge, the parent-child relationship can be, under certain conditions, interpreted as causal and defined as an arc. A parent-child relationship implies that the values of these two random variables are not independent. By repetitively testing for in-

dependence conditioned on other subsets of variables, algorithms can learn the topology of a Bayesian network, which is expressed as a directed acyclic graph (DAG), directly from data (26). As algorithms for reconstructing Bayesian networks emerged, they were used to model signaling pathways within cells (27), to identify known DNA repair networks in *E. coli* using microarray data (28) and to identify simple phosphorylation cascades in T cells using flow cytometry data (29, 30). While many more studies have been published since, a common conclusion is that the statistical confidence associated with an inferred network improves as the number of samples included in a dataset is greater than the number of random variables. However, transcriptomics data, like that obtained as part of the TCGA, typically have a large number of random variables (n_{genes}) and a small number of biological replicates ($n_{patients}$), which makes inferring gene-level networks computationally difficult and also implies greater uncertainty in the inference (31).

As summarized in Figure 1, we propose an approach that combines digital cytometry with Bayesian network inference to identify how heterocellular networks associated with functional plasticity and anti-tumor immunity change during oncogenesis in humans. Conceptually, digital cytometry improves the statistical power by projecting the transcriptomic space onto a smaller number of "features" that estimate the prevalence of stromal and immune cell types and the average differentiation state of malignant cells present within the tumor microenvironment, such that $n_{features} << n_{patients}$. The causal structure among these features can then be predicted using Bayesian network inference. While data unstructured in time, such as the TCGA datasets, are not ideal for inferring causality, we test the inferred networks using in vivo experiments using syngeneic murine tumor models.

To illustrate the approach, we focused on Cell Communication Network factor 4 (CCN4/WISP1), as it is upregulated in invasive breast cancer (32), expressed by malignant cells, and correlates with a lower overall survival in patients diagnosed with primary melanoma (33). Functionally, CCN4 promotes metastasis in melanoma by promoting a process similar to the epithelial-mesenchymal transition (33, 34). In developing state metrics that quantify functional plasticity in breast cancer and melanoma using an unsupervised approach, CCN4 was the only gene associated with both a mesenchymal state metric in breast cancer and a dedifferentiated state metric in melanoma that results in a secreted protein (21). The collective set of features, or simply nodes of a network, were quantified in three transcriptomic datasets obtained from bulk tissue samples from patients with breast cancer and melanoma and used to generate a casual network describing how expression of a gene by malignant cells, such as CCN4, more broadly alters the heterocellular network within a tissue using Bayesian network inference.

Results

Generating causal graphs that link oncogenic changes in gene expression with changes in the heterocellular network. Bayesian network inference involves inferring the structure of the network, which captures the specific causal interactions or arcs among the nodes of a network and represents them as a directed acyclic graph (DAG), and then estimating the parameters of the conditional probability distribution from the datasets. We used a four-step process to learn the causal structure associated with the cell-level networks. The four steps corresponded to specifying a "blacklist" based on prior information, generating an ensemble of potential arcs using 10 different structural learning algorithms, filtering potential arcs based on a trade-off between regression accuracy and model complexity to create a "whitelist", and learning the network structure using both the "blacklist" and "whitelist". We will discuss this process in more detail in the following paragraphs.

In learning the causal structure of a network, the network structure can be shaped by prohibiting the inclusion of specific arcs into a proposed network, that is by assigning an arc to a "blacklist". Here, the "blacklist" represents a way to incorporate prior knowledge about causal relationships associated with oncogenesis and the roles that specific immune cells play in controlling tumor cell growth. In particular, we considered only arcs into the "CD8 T cells" node (i.e., a leaf node), only arcs that originate from the "Cancer" node (i.e., a root node), mostly arcs that originate from the "CCN4" node (with exception for the "Cancer" node), and only arcs into the "CD4 T cells" and "Neutrophils" nodes. Cancer as a root node follows from contemporary understanding of oncogenesis, where mutation of either oncogenes or tumor-suppressor genes is the cause of cancer development (35). Including CCN4 as a root node follows from its frequent mutation in and scRNAseq-assayed expression by malignant cells (33, 36). Recent immunotherapies that either adoptively transfer CD8 T cells (37) or improve the function of CD8 T cells by relieving immune checkpoints (38) motivate specifying CD8 T cells as a leaf node. Specifying "CD4 T cells" and "Neutrophils" as leaf nodes follows from the high number of zero values for those features in the dataset, which were 350 and 439 samples in the BRCA dataset, respectively. These limits were implemented by assigning the corresponding arcs to the "blacklist". In a Bayesian context, assigning an arc to the "blacklist" specifies the prior probability of including this arc in network to zero. Practically, including these different categories of arcs into the "blacklist" helped infer arc direction consistently among the different structural learning algorithms (see Figure S1). As the number of arcs included in the blacklist was increased, the number of edges with unclear direction was decreased. Specifically, 17 edges had an unclear direction without specifying a blacklist (Fig. S1A), while only 6 edges had an unclear direction in the final blacklist (Fig. S1E).

As algorithms for structural learning have different underlying assumptions, we used an ensemble approach to average across the different algorithms to identify a list of potential arcs that were used to seed the final learned structure of the DAG, that is a consensus seed network. This ensemble approach mirrors the "community network" approach

described by Marbach and coworkers based on experiences with DREAM challenges for inferring intracellular networks (39, 40). Specifically, we used ten different structural learning algorithms (26), including constraint-based (Incremental association Markov Blanket - IAMB (41), Incremental association with false discovery rate control - IAMB.FDR (42), practical constraint - PC.STABLE (43), grow-shrink Markov Blanket - GS (44)), score-based (hill climbing - HC, Tabu search - Tabu (45)), and hybrid learning (max-min hill-climbing - MMHC (46), restricted maximization - RSMAX2 (47)) algorithms. Two algorithms for local discovery of undirected graphs (max-min parents and children - MMPC (41), Hiton parents and children - SI.HITON.PC (48)) were also included to provide additional evidence supporting the existence of an edge between two nodes. Bootstrap resampling, that is resampling the dataset with replacement to generate a synthetic dataset of similar size as the original and infer the network structure using the synthetic dataset, was used in learning the network structure with each algorithm, which resulted in generating 10,000 network structures. For each algorithm, an averaged network structure was then calculated from this collection of network structures, where the threshold for including an arc into the average network was automatically determined by each algorithm and was nominally 0.5 using the approach described by Scutari and Nagarajan (49). We applied the same approach to both the breast cancer (BRCA - Figures 2 and 3 and Table S1) and the two melanoma datasets (common melanocytic nevi and primary melanoma: GEO - Figures S2 and 5 and Table S2, and primary melanoma from the TCGA: SKCM - Figures S2 and 5 and Table S3). Of note, edge/arc numbers are conserved across Tables S1-S3 and Figures 2 and S2.

The strength of evidence supporting the existence of an arc, that is arc strength, can also be used to filter arcs for inclusion in the consensus seed network. For instance, all arcs with a strength below a certain threshold can be included in the "whitelist" (see Fig. 2B). As this threshold was increased, more arcs were included in the "whitelist" and the number of arcs included in the final DAG was also increased. We also note that, while exploring the impact of a strength threshold, arcs were left out of the "whitelist" if their inferred direction varied among the algorithms (yellow bars in Figures 2 and S1). While increasing the number of arcs in the DAG better approximates the joint probability distribution, a complicated network limits interpretability. We used a Bayesian Information Criterion (BIC) to quantify this trade-off between regression accuracy and model complexity, where the optimal balance is at a minimum value. When the BIC was used to analyze the entire BRCA network, an empty graph provided the maximum value and the minimum value corresponded to an empty "whitelist". As the weight of each node in contributing to the overall BIC value varied, we found that applying the BIC to individual nodes instead of the entire DAG provided better insight into including specific arcs in the "whitelist". For instance, parental arcs associated with CCN4, Mesenchymal, and CD8 T cells were readily identified without specifying a "whitelist". Adding additional arcs with higher strength values (i.e., arc strength values greater than 1E-20) associated with these nodes into the "whitelist" increased the BIC values. However, BIC values for other nodes, such as active NK cells and p(M2), were only improved when arcs with higher strength values were included. The arcs included in the "whitelist" used in the final analysis corresponded to the parental arcs associated with the minimum BIC values determined at the node-level for each dataset and that arcs didn't form cycles (see Fig. 2B and Fig. S2). For instance, the minimum BIC value for the Endothelial node of 119.4 occurred in the BRCA analysis using a strength threshold for whitelist inclusion of 1E-10. This minimum BIC value was associated with five parental arcs: arc numbers 9 ("proliferation" $\rightarrow$ "Endothelial cells"), 19 ("Cancer" $\rightarrow$ "Endothelial cells"), 20 ("naïve B cells" $\rightarrow$ "Endothelial cells"), 33 ("Mesenchymal" $\rightarrow$ "Endothelial cells") and 39 ("p(M1)" $\rightarrow$ "Endothelial cells"). However, only arc numbers 19, 33, and 39 were included in the whitelist. Arcs 9 and 20 formed cycles with arcs identified at the minimum BIC values for other nodes and were excluded from the whitelist. The final network for each dataset was generated using a hybrid learning algorithm (mmhc) using a "blacklist" specified based on prior causal knowledge and a "whitelist" corresponding to the consensus seed network. Similar to the first step, bootstrap resampling (n_{boot} = 10,000) and network averaging were used to generate the three DAGs shown in Figures 3 and 5. Each DAG was used to generate parameters for a linear Gaussian model estimated by maximum likelihood and conditioned on the network structure that approximates the joint probability distribution associated with the dataset. Values for the linear coefficients and the average node values were used to annotate the DAGs. The sign of the linear coefficient was also used to annotate whether a particular arc promotes or inhibits the target node.

The resulting DAGs imply that oncogenesis in breast cancer was associated with a shift from epithelial to mesenchymal cell state accompanied by an increase in cell proliferation and a suppression of endothelial cells, which were inferred with high confidence. In turn, endothelial cells promote the infiltration of CD4 T cells. The local structure associated with "Cancer"'s influence on the "Mesenchymal" state via "CCN4" suggests an incoherent type-3 feed-forward motif to regulate the mesenchymal state. Inference of a feed-forward motif is interesting as feed-forward loops are highly prevalent and well understood as control mechanisms in intracellular networks (50) but are less well understood in the context of intercellular networks. In addition, CCN4 inhibits active NK cells. The high confidence arc between active NK and resting NK cells follows from these features being mutually exclusive in the dataset and very few samples having zero values for both features. The mesenchymal state increased cancer-associated fibroblasts ("CAFs") with high confidence. Interestingly, oncogenesis was also associated with increasing the prevalence of a type 1 macrophage ("p(M1)"), which in turn promoted the recruitment of CD8 T cells. The prevalence of CD8 T cells are also connected to "Cancer" via a larger incoherent feed-forward motif involving "p(M1)", "CCN4", the

"Mesenchymal" state, and "CAFs" with high confidence.

As there was more data supporting the BRCA DAG, the resulting Bayesian network model was compared against the underlying experimental data and used to explore the impact of varying CCN4 expression in the context of normal and cancer tissue (Figure 4). To simulate normal and cancer tissue, we queried the conditional probability distribution by generating samples from the Bayesian network and filtered the values based on "Cancer" < 0.05 and "Cancer" > 0.95, which are colored in blue and red, respectively. The corresponding experimental data points and trend lines are overlaid upon the posterior distributions. The posterior distributions mirror the experimental data points, where there is an increase in CCN4 expression between normal ("Cancer" $<$ 0.05) and cancer ("Cancer" > 0.95) tissue. The posterior distributions mirror the variability observed in the experimental data when comprised of non-zero values, such as CD8 T cells and CAFs. In contrast, the prevalence of zero values increased the range of the posterior distribution, such as for CD4 T cells. In comparing normal to cancer tissue, CD8 T cells was the only feature, on average, increased in cancer tissue, while CD4 T cells and CAFs were decreased and active NK cells, B cells, and Macrophages exhibited similar trends. Slopes of the trend lines highlight the influence of CCN4 gene expression on the prevalence of different immune cell populations. Increased CCN4 had the most pronounced inhibition on active NK cells and also suppressed CD8 T cells. CCN4 also had a pronounced positive impact on the prevalence of CAFs, macrophages, and CD4 T cells. CCN4 seemed to have no impact on B cells.

The breast cancer dataset contained 582 samples, of which 8.8% were from normal mammary tissue. In contrast, the two melanoma datasets contained 78 GEO samples, which includes 34.6% benign nevi, and 94 SKCM samples of primary melanoma only. While a lower number of samples limits the inferential power of a dataset, we decided to analyze them separately as they had different distributions in transcript abundance as a function of transcript length. As the Bayesian network inference algorithm leverages differences in the magnitude of a feature within a population, approaches to harmonize these two datasets may introduce a systemic bias that is convoluted with oncogenic transformation, as the GEO dataset has many samples obtained from benign nevi while the SKCM dataset does not. In analyzing the melanoma datasets separately, we developed separate consensus seed networks using the node-level BIC values (see Fig. S2). Using these consensus seed networks to specify the corresponding "whitelist", the structure and parameters associated with Bayesian network was inferred independently from each melanoma dataset (see Figure 5).

Given the high prevalence of samples from benign nevi in the GEO dataset, high confidence arcs in the GEO network focus on changes associated with oncogenesis. Similar to the breast cancer analysis, oncogenesis was associated with a shift from an epithelial to a mesenchymal-like cell state. The mesenchymal cell state is promoted by both oncogenesis and CCN4 via a feed-forward motif, depicted here as the more common type 1 coherent motif. In addition, CCN4 indirectly impacted CAFs by promoting a mesenchymal-like cell state. Similar to the breast cancer analysis, oncogenesis promoted an increase in CD8 T cells, but indirectly by recruiting active NK cells. In analyzing the SKCM dataset, less emphasis is placed on the changes associated with oncogenesis but how expression of CCN4 influenced the network. Similarly to the GEO analysis, the SKCM analysis suggested that CCN4 directly impacted the mesenchymal state that then influenced CAFs. In addition, CCN4 directly promoted resting NK cells, which is similar to the BRCA analysis considering the reciprocal relationship between active and resting NK cells. In both melanoma datasets, CD8 T cells were directly promoted by macrophages and by NK cells either directly in the GEO dataset or indirectly through macrophages in the SKCM dataset. In addition, neutrophils, proliferation, and B cells were independent of all nodes. In all three analysis, there was high confidence associated with the arcs among the nodes quantifying macrophage polarization, which is likely an artifact of the formula used to calculate $p(M\Phi i)$'s. Queries of the conditional probability distribution based on the SKCM DAG for active NK cells, Macrophages, B cells, and CAFs were similar to the BRCA analysis (Fig. S3). Similar to the BRCA analysis, a high number of zero values for the CD4 T cell features in the SKCM dataset suggests caution in interpreting differences in CD4 T cell predictions. To validate these predictions, we focused on two strategies: using flow cytometric analysis of tumor infiltrating lymphocytes in syngeneic mouse tumor models to validate the digital cytometry predictions and targeted in vitro experiments to validate the presence or absence of arcs predicted by the Bayesian network inference.

Validating the impact of CCN4 on heterocellular networks using syngeneic mouse models. Syngeneic immunocompetent mouse models of cancer provide an important complement to retrospective studies of human data as they can aid in causally linking genetic alterations with cellular changes the tumor microenvironment. Here we used two syngeneic transplantable models for melanoma to test the predictions generated by the collective approach: the spontaneous B16F0 model and the YUMM1.7 model that displays Braf$^{V600E/WT}$ Pten$^{-/-}$ Cdkn2$^{-/-}$ genotype. As these cell lines basally produce CCN4, we generated CCN4 knockout (KO) variants using a CRISPR/Cas9 approach and confirmed CCN4 KO by testing conditioned media for CCN4 by ELISA. Tumors were generated by injecting the cell variants subcutaneously in 6-8 week old female C57BL/6 mice and monitoring for tumor growth. Once wild-type (WT) tumors reached between 1000 and 1500 mm^3 in size, tumors were surgically removed from all mice that were not considered outliers and processed into single cell suspensions (n = 7 for YUMM1.7 variants and n = 4 for B16F0 variants). The single cell suspensions were aliquoted among three antibody panels to characterize the tumor infiltrating lymphocytes by flow cytometry (see Figures S4-S6 for gating strategies). While the B16F0 and YUMM1.7 KO variants were generated using a double nickase CRISPR/Cas9 approach,

similar results were obtained using a homology directed repair strategy (34, 51). Additional controls for puromycin selection of CRIPSR/Cas9 edited cells using B16F0 cells transfected with a pBabe-puromycin retrovirus also behaved functionally similar in vitro and in vivo as wild-type B16F0 cells (34).

The percentage of CD45$^+$ cells among total live cells exhibited a semi-log dependence on tumor size (Fig. 6A - B16F0: R^2 = 0.607, F-test p-value = 7.27E-6; YUMM1.7: R^2 = 0.830, F-test p-value = 1.48E-7), where CCN4 KO resulted in smaller tumors in both cell models with greater CD45$^+$ cell infiltration. As illustrated in Figure 6A, YUMM1.7 variants had a much higher dependence on tumor size than B16F0 variants. Conventionally, flow cytometry data are normalized to tumor size to estimate the prevalence of a particular cell type per tumor volume. Yet, the dependence on tumor size could be a confounding factor in addition to CCN4 expression that could skew the results. Moreover, the Bayesian network analysis predicts the impact of CCN4 alone on the prevalence of specific immune cell subsets. Thus, we focused instead on the prevalence of a particular cell type within the live CD45$^+$ TIL compartment to validate the digital cytometry predictions.

In comparing the WT B16F0 and YUMM1.7 models, the relative prevalence of NK (Live CD45$^+$ CD3e$^-$ NK1.1$^+$ B220$^-$ events), CD4$^+$ T (Live CD45$^+$ CD3e$^+$ CD4$^+$ CD8a$^-$ events), and CD8$^+$ T (Live CD45$^+$ CD3e$^+$ CD4$^-$ CD8a$^+$ events) cells were similar while B cells (Live CD45$^+$ CD3e$^-$ NK1.1$^-$ B220$^+$ events) were almost 10-times more prevalent in the B16F0 tumors compared to YUMM1.7 tumors (Fig. 6B). The prevalence of these different cell types changed within the CD45$^+$ TIL compartment upon CCN4 KO (Fig. 6C and D). Figure 6C highlights the trends among the mouse models and compares against the digital cytometry predictions obtained from the BRCA and SKCM datasets. Predictions for the change in cell type prevalence by CCN4 expression were obtained by comparing the prevalence of the indicated cell type estimated by digital cytometry with CCN4 expression. Given the uncertainty associated with how particular tissue samples represent the full dynamic range of biologically relevant CCN4 expression, the BRCA and SKCM samples were separated into 5 quantiles based on CCN4 expression (see Figs. S7 and S8). We used the feature values from the highest CCN4 quantile to represent wild-type tumors and from the lowest CCN4 quantile to represent CCN4 knock-out tumors. We then calculated the difference in cell type abundance upon CCN4 knock-out. Specifically, CD4 and CD8 T cells and B cells had analogous digital cytometry features as assayed in the flow panel, while NK cells were mapped to the "active NK cells" feature. The relative change in abundance was largely consistent among the four systems, with the YUMM1.7 model being the most different. The BRCA and SKCM datasets predicted that NK cells were most reduced by CCN4, which was observed in both the B16F0 (p-value = 0.047, n = 4) and YUMM1.7 (p-value = 0.051, n = 7) models. The BRCA dataset predicted that CCN4 reduced CD8$^+$ T cells, which was observed in the YUMM1.7 model (YUMM1.7 p-value = 0.002) with a similar trend observed with the B16F0 model (p-value = 0.770). The CD4$^+$ T cells seemed to vary in response to CCN4 expression as the BRCA and B16F0 results showed an increase while the SKCM and YUMM1.7 results showed a decrease. As stated previously, the BRCA and SKCM predictions for CD4 T cells should be interpreted with caution given the high frequency of zero values for the features. B cell response was mixed with both the BRCA and SKCM results suggesting no change. In mice, B cells were observed to decrease in the B16F0 model and increase in the YUMM1.7 model upon CCN4 KO, with the low number of B cells infiltrating YUMM1.7 tumors rendered the results more variable. Given the small sample size of the experimental mouse cohorts, only the extremes were statistically significant, with NK cells significantly increased (p-value = 0.047) and B cells significantly decreased (p-value = 0.002) in B16F0 CCN4 KO tumors and CD8 T cells significantly increased (p-value = 0.002) in YUMM1.7 CCN4 KO tumors (Fig. 6D).

Concordance in CCN4-induced changes in the myeloid compartment are less clear. In addition to changes in T and NK cells within the live CD45$^+$ compartment, we also assayed myeloid subsets in tumors generated by WT and CCN4 KO variants of the B16F0 and YUMM1.7 cell lines. Using the gating strategy summarized in Figure S6, we focused on CD11c$^+$ and CD11c$^-$ macrophages (live CD45$^+$ CD11b$^+$ GR1$^-$ F4/80$^+$ events), neutrophils (live CD45$^+$ CD11b^{int} CD11c$^-$ GR1$^+$ F4/80$^-$ events), dendritic cells (live CD45$^+$ CD11b$^{lo/int}$ CD11c$^+$ GR1$^-$ F4/80$^-$ events), and two different myeloid-derived suppressor cell (MDSC) subsets: CD11c$^-$ and CD11c$^+$ MDSC (live CD45$^+$ CD11b$^+$ GR1$^+$ F4/80$^{+/-}$ MHCII$^+$ events). In comparing tumors derived from WT cell lines, CD11c$^+$ macrophages were the most predominant infiltrating myeloid cell subset and most subsets were consistent between the two mouse models (Fig. 7A). Upon CCN4 KO in the mouse models, the CD11c$^+$ macrophage subset increased while the MDSC subsets decreased (Fig. 7B-E) within the CD45$^+$ compartment, while the response varied for the two least abundant subsets: neutrophils and CD11c$^-$ macrophages. The reduction in CD11c$^+$ MDSC in CCN4 KO variants were most pronounced and statistically significant (p=0.004 in YUMM1.7 and p = 0.153 in B16F0). While Ly6G and Ly6C staining may have been a better staining strategy for distinguishing among monocytic (Mo-) and polymorphonuclear (PMN-) MDSC subsets, we observed a reduction in PMN-MDSCs in YUMM1.7 tumors upon CCN4 KO using Ly6G/Ly6C antibodies (51). Consistent with the idea that PMN-MDSCs arise from impaired differentiation of granulocytes, neutrophils were increased within the CD45$^+$ compartment in CCN4 KO tumors derived from YUMM1.7 cells (p = 0.002) but not statistically different in the B16F0 model (p = 0.097). Other myeloid subsets trended similarly but with differences that were not statistically significant. In addition, we noted that a dendritic cell subset (live CD45$^+$ CD11b$^{lo/int}$ CD11c$^+$ GR1$^-$ F4/80$^-$ cells) increased upon CCN4 KO (p=0.045 in YUMM1.7 and p=0.011 in B16F0).

Comparing the trends in the myeloid compartment observed among the mouse models and the Bayesian network predictions obtained from the BRCA and SKCM datasets is less clear, given the uncertainty as to how the digital cytometry features map onto the quantified myeloid subsets in these mouse models. In particular, the digital cytometry features, that is the subset of particular genes and their relative expression that constitute a particular cell's gene expression signature, were defined using differentiated cell subsets. In contrast, MDSCs are prevalent within the tumor microenvironment and result from impaired differentiation of myeloid precursors into mature myeloid cells, like macrophages, dendritic cells, and neutrophils. Uncertainty in mapping how the gene expression signature in an immature myeloid cell subset overlaps with a differentiated myeloid cell creates uncertainty in predicting cell abundance by digital cytometry. Despite those concerns, key myeloid features in the Bayesian networks were macrophages oriented towards a M1 phenotype. Correspondingly, CD11c$^+$ macrophages, a subset that has been associated with pro-inflammatory M1 tumor-associated macrophages (52), were the most predominant myeloid subset in WT B16F0 and YUMM1.7 tumors and didn't change upon CCN4 KO. In the BRCA dataset, the prevalence of macrophages was influenced by CCN4 expression; yet, the functional orientation away from the M2 and towards the M1 phenotype depended solely on oncogenic transformation. Similarly, the prevalence of macrophages was influenced by both CCN4 expression and oncogenic transformation in both melanoma datasets. In contrast to the BRCA results, functional orientation of macrophages were independent of both oncogenic transformation and CCN4 expression. Neutrophils were predicted to be independent of CCN4 in the melanoma datasets, which is not surprising considering that the the majority of tumors had zero values for the Neutrophil feature. Similarly, neutrophils were about 10 times less abundant than CD11c+ macrophages in the mouse models. Given the significant changes observed in MDSCs and the corresponding differentiated cell subsets upon CCN4 KO in the mouse models, challenging digital cytometry predictions in this way highlights features that can be improved, such as discriminating among terminally differentiated and immature subsets, like Mo-MDSC and PMN-MDSC.

CCN4 has no direct effect on T cell proliferation but alters CD8$^+$ T cell function. The local proliferation of CD8$^+$ T cells correlates with clinical response to immune checkpoint blockade (53, 54). In addition, the DAGs inferred from both the breast cancer and melanoma datasets suggest that a decrease in CD8$^+$ T cells is driven indirectly through CCN4 via modulating cancer-associated fibroblasts or the activity of NK cells. While the structural learning algorithms rejected a direct edge between CCN4 and CD8$^+$ cells, we tested whether CCN4 directly inhibits T cell proliferation (see Fig. 8A and B) using a statistical analysis of Cell Trace distributions in CD4$^+$ and CD8$^+$ T cells stimulated in vitro (see Table S4). Specifically, splenocytes were stimulated in vitro with αCD3/αCD28-loaded beads in the presence of media conditioned by WT or CCN4 KO B16F0 cells or supplemented with 10 ng/ml recombinant mouse CCN4. In both the CD4$^+$ and CD8$^+$ T cell populations, the presence of tumor-conditioned media significantly inhibited the fraction of cells that divided at least once (Dil - CD4 p-value = 0.022, CD8 p-value = 0.018) and the probability that a cell will divide at least once (PF - CD4 p-value = 0.024, CD8 p-value = 0.013) while CCN4 exposure was not a statistically significant factor. For responding cells, the average number of divisions they undergo (PI) was not different among experimental conditions for CD4$^+$ T cells (p-value = 0.22) but reduced in CD8$^+$ T cells exposed to tumor-conditioned media (p-value = 0.0077). Overall, the presence of tumor-conditioned media and not CCN4 influenced T cell proliferation, which was consistent with the DAGs.

Another characteristic of CD8$^+$ T cells present within the tumor microenvironment is that they are dysfunctional (55). As the digital cytometry approach used here doesn't estimate the functional state of CD8$^+$ T cells only their prevalence within a tissue sample, we decided to test whether CCN4 had a direct impact on CD8$^+$ T cell function, as quantified by target-specific ex vivo cytokine release as measured by ELISpot. First we generated YUMM1.7-reactive CD8$^+$ T cells by immunizing C57BL/6mice against YUMM1.7 cells and isolated CD8a$^+$ T cells from splenocytes three days after re-priming with live YUMM1.7 cells. We also created a variant of CCN4 KO YUMM1.7 cells with CCN4 expression induced by doxycycline and vector controls that were used as target cells (see Fig. S9). IFNγ and TNFα ELISpots were used to quantify the CD8$^+$ T cell functional response to the different tumor targets in the presence or absence of tumor-produced CCN4. As expected for a re-call CD8a$^+$ effector T cell response, the most prominent IFNγ and lowest TNFα responses were against WT and CCN4 KO YUMM1.7 cells, with a slightly higher IFNγ response to WT YUMM1.7 targets (see Fig. 8C, p-value = 0.098). Interestingly, re-expression of CCN4 by CCN4 KO YUMM1.7 cells following doxycycline induction significantly reduced both IFNγ and TNFα production (p-values < 0.001), which suggests that CCN4 plays a direct role in inhibiting CD8a$^+$ T cell function. Of note, CCN4 was predicted to directly inhibit the activity of NK cells, which share cytokine release, like IFNγ, and cytotoxicity mechanisms with CD8$^+$ T cells. Overall, the changes observed between WT and CCN4 KO variants of the B16F0 and YUMM1.7 mouse models were consistent with the causal networks inferred from the breast cancer and melanoma datasets.

Discussion

Validating the role that a particular molecule plays in driving the disease state using targeted experiments is central for improving understanding of biological mechanisms or selecting among competing drug targets. Given the limited observability of the biological response in experimental models and patients, mechanistic modeling and simulation is playing an increasing role in helping answer many central questions in discovering, developing, and receiving federal approval of pharmaceutical drugs and also basic biology (56). In immuno-

oncology, there is increasing interest in modeling the heterocellular network of relevance for a specific immunotherapy. The first step in creating mathematical models of cell-level networks is to create the topology of the network, which is expressed in terms of which nodes to include and how they influence each other. The structure of these cell-level models is created using a fully supervised approach, which means by hand using expert knowledge (57). For instance, systems of ordinary differential equations have been developed to capture multiple spatial compartments containing interacting malignant, antigen presenting, and T cells and to predict a general immune response (58), a response to immune checkpoint blockade using CTLA-4, PD-1, and PD-L1 antibodies (59) or adoptive cell transfer (60).

While leveraging the knowledge of experts is a great starting point, hand-curated models can also implicitly impose bias on how data is interpreted. In the context of molecular-level networks, rules and algorithms have been developed to elaborate causal networks based on a limited set of rules (61–64). The rules constrain the types of interactions, or arcs, that are realistic between the nodes while the algorithms generate all possible arcs that are consistent with the rules and collection of nodes. The resulting rule-based networks are then used to interpret data by filter the arcs for the most consistent and, in the process, may reveal previously unappreciated pathways. For instance, a rule-based model was used to interpret single-molecule detection of multisite phosphorylation on intact EGFR to reveal a new role for the abundance of adaptor proteins to redirect signaling (65). Given the challenges with representing the various activation states of a 12-subunit Ca^{2+}/calmodulin-dependent protein kinase II (CaMKII) holoenzyme that is essential for memory function, a rule-based model identified a molecular mechanism stabilizing protein activity that was obscured in prior reduced models (66). Inspired by engineering better CAR T cells, Rohrs et al. developed a rule-based model to interpret site-specific phosphorylation dynamics associated with Chimeric Antigen Receptors (67).

To our knowledge, no equivalent approach exists in the context of modeling cell-level networks.[1] We posit that coupling digital cytometry with Bayesian network inference is analogous to rule-based modeling in the context of modeling cell-level networks. Here, the rules comprise a limited set of constraints, or heuristics, related to the direction of information flow. Specifically, the rules limit how changes in gene expression within the malignant cell introduced during oncogenesis propagate to stromal and immune cells present within the tumor microenvironment and are implemented as a "black list". In formulating a black list, specifying cancer as a root node and $CD8^+$ T cells as a leaf should be generalizable in the context of most solid tumors. Due to immune privilege, glioblastoma multiform may be an exception for $CD8^+$ T cells (71). Specifying the oncogene as a child of cancer and a potential root node to everything else should also be generalizable. Single cell RNAseq data can help support this assumption that malignant cells express the oncogene. Specifying nodes that have a high percentage of zero values as leaves follows from network inference arguments as trends in the non-zero values, which ideally are related to other nodes, may be swamped by differences between the zero and non-zero values and may give rise to spurious arcs. The algorithms that underpin Bayesian network inference search over all possible network topologies for arcs that are consistent with the data. The resulting networks can be used in multiple ways. As an unsupervised approach, the network topology could complement existing workflows for creating mechanistic mathematical models fit for use in testing molecular targets (57, 72). In addition, DAGs represent explicit hypotheses generated from pre-existing human data that motivate new experiments to validate the predictions, as illustrated by the B16F0 and YUMM1.7 results.

While the focus here is in the context of breast cancer and melanoma due the pre-existing breadth of data, the approach could be generally applied to other biological contexts and motivate new experimental studies. For instance, one of the limitations of inferring the network topology in the form of directed acyclic graphs is that some direct and indirect causal relationships can be confounded, such as reciprocal feedback modes of communication between cells (73). Discerning the difference between a direct and indirect causal relationship has practical importance, such as for selecting therapeutic targets (74). Methods, like Granger causality and dynamic Bayesian networks (31, 75, 76), do exist that could reveal direct and indirect causal relationships, but time-series data is required. Unfortunately, human tissue samples, like those in the TGCA, are very rarely sampled with time. Analysis of pre-existing human datasets can be complemented by a more focused experimental study of a pre-clinical model. Specifically, single-cell RNAseq to identify the cell types present and their associated gene signatures can be combined with bulk transcriptomic seqencing to capture the prevalence of all of the cell types within the tissue sample and provide a large number of biological replicates spanning the disease space - normal homeostasis; initiation; early, middle and late progression; and productive resolution or adverse outcomes. Similar network topologies would suggest similar biological mechanisms and help select relevant pre-clinical models for drug development. In short, we feel that combining digital cytometry with Bayesian network inference has the potential to become an indispensable unsupervised approach for discovering relevant heterocellular networks associated with disease.

[1] One might consider agent-based or cellular automata models to apply as the cellular interactions are specified by rules. In rule-based modeling of molecular networks, the rules and algorithms elaborate a network space that encompasses all possible topologies of the network and data is used to prune the network to the most relevant. Similarly, the arcs included in the "blacklist" and "whitelists" can be considered as a Bayesian prior, where the strength of inclusion in the final DAG and the coefficient associated with a particular arc in the conditional probability function depend on the data. In contrast, agent-based or cellular automata models require specifying all interactions between cells as rules a priori and are validated qualitatively by comparing emergent behavior against experimental observations (68–70).

Methods

Digital Cytometry. Transcriptomics profiling of bulk tissue samples using Illumina RNA sequencing for the breast cancer (BRCA) and cutaneous melanoma (SKCM) arms of the Cancer Genome Atlas were downloaded from TCGA data commons, where values for gene expression were expressed in counts using the "TCGAbiolinks" (V2.8.2) package in R (V3.6.1) and converted to TPM. RNA-seq data assayed in samples acquired from benign melanocytic nevi and untreated primary melanoma tissue and associated sample annotation were downloaded from GEO entry GSE98394 and converted to TPM. TCGA data and the benign nevi and melanoma data were filtered to remove sample outliers and normalized based on housekeeping gene expression (77). Digital cytometry features associated with the functional plasticity of malignant cells within an epithelial to mesenchymal-like state space were calculated based on state metrics developed separately for bulk breast cancer and melanoma tissue samples (21). Cell proliferation features were calculated based on the median expression of genes associated with cell proliferation that were identified previously using human cell line data (34). Features corresponding to the prevalence of endothelial cells, cancer-associated fibroblasts, macrophages, and CD4$^+$ T cells were calculated using CIBERSORTx (https://cibersortx.stanford.edu) using the gene signatures derived from single cell RNAseq data (36) while the prevalence of B cells naïve, CD8$^+$ T cells, Macrophage M0 ($M\Phi 0$), Macrophage M1 ($M\Phi 1$), Macrophage M2 ($M\Phi 2$), activated NK cells, resting NK cells, and neutrophils were calculated using the LM22 immune cell gene signatures in CIBERSORTx run in absolute mode.

Given the potential lack of independence among the macrophage features, the LM22 macrophage features were combined to estimate the probability of the average functional orientation using the formula described previously(71):

$$p(M\Phi i) = \frac{M\Phi i}{M\Phi 0 + M\Phi 1 + M\Phi 2}, \qquad \textbf{(1)}$$

where $i = \{0, 1, 2\}$ and denotes the specific macrophage subtype. Additional cellular features were excluded from the analysis as they tended to have a large number of zero values across the datasets, which resulted in being disconnected from the rest of the network in preliminary structural inference. Sample attributes were transformed to numerical values, which were assumed to be extremes of a continuous variable. For instance, if the sample was obtained from normal tissue, the value for "Cancer" was set equal to 0; if the sample was obtained from cancer tissue, the value for "Cancer" was set equal to 1. The sample attributes, CCN4 gene expression, and estimated cellular features extracted from the bulk RNAseq data calculated for each sample are included in the GitHub repository.

Bayesian Network Inference. Prior to network inference, feature values were log transformed, normalized to values between 0 and 1, and discretized (BRCA: 15 intervals; GEO and SKCM: 6 intervals).The features were then assigned to nodes. The relationships among the nodes, or arcs, were represented by directed acyclic graphs inferred from the datasets using a process involving four steps, as detailed in the results section. Given the inferred structure, a Bayesian network in the form of a linear Gaussian model was fit to the datasets using maximum likelihood estimation of the model parameters. A Bayesian Information Criterion (BIC) was used to evaluate the trade-off between a L1 loss function, which numerically quantifies the ability of a specified directed acyclic graph to represent the conditional independence of the joint probability distribution, and a penalty associated with network complexity. BIC values were calculated by:

$$BIC = n_{arcs} \cdot log_{10}(n_{obs}) + \sum_{i=1}^{n_{obs}} |\hat{Y}_{ij} - Y_{ij}|, \qquad \textbf{(2)}$$

where $\hat{Y}_{ij}$ and Y_{ij} are the observed and predicted values for feature j, n_{arcs} is the number of parental arcs associated with the entire DAG or node, and n_{obs} is the number of samples. Conditional probability queries of the Bayesian networks were performed by logic sampling with 10^5 samples. Bayesian network inference was performed using the 'bnlearn' package (V4.5) in R (V3.6.1).

Reagents and Cell Culture. Cytokines and antibodies were obtained from commercial sources and used according to the suppliers' recommendations unless otherwise indicated. The mouse melanoma line B16F0 (purchased in 2008, RRID: CVCL_0604) was obtained from American Tissue Culture Collection (ATCC, Manassas, VA). The mouse melanoma line YUMM1.7 (received in September 2017, RRID: CVCL_JK16) was a gift from Drs. William E. Damsky and Marcus W. Bosenberg (Yale University) (78). B16F0 and YUMM1.7 cells were cultured at 37°C in 5% CO_2 in high-glucose DMEM (Cellgro/Corning) supplemented with L-glutamine (Lonza), penicillin-streptomycin (Gibco), and 10% heat-inactivated fetal bovine serum (Hyclone). All cell lines were revived from frozen stock, used within 10-15 passages that did not exceed a period of 6 months, and routinely tested for mycoplasma contamination by PCR. CCN4 knock-out variants of B16F0 and YUMM1.7 cells were generated using a double-nickase CRISPR/Cas9 editing strategy described previously (33). Briefly, two pairs of mouse CCN4 double nickase plasmids that target the mouse CCN4 gene at different locations were purchased from Santa Cruz Biotechnology, Inc. (Dallas, TX) and transfected into B16F0 and YUMM1.7 cells following the manufacturer's instructions. Following antibiotic selection, surviving single clones were isolated and expanded on 6-well plates. The concentration of CCN4 in the cell culture media from those wells was assayed using the Human WISP-1/CCN4 DuoSet ELISA Kit (R&D Systems, Minneapolis, MN) to confirm CCN4 knockout. CCN4-knockout cells were further expanded and aliquoted to create a low passage frozen stock.

In vivo Tumor Assays and in vitro T cell proliferation assays. All animal experiments were approved by West Virginia University (WVU) Institutional Animal Care and Use Committee and performed on-site. C57BL/6Ncrl mice (6-8 week-old female) were from Charles River Laboratories. Mice were randomly assigned to treatment groups and co-housed following tumor initiation. Subcutaneous tumors were initiated by injecting mice subcutaneously with 3×10^5 of the indicated YUMM1.7 cells and 2.2×10^5 of the indicated B16F0 cells in 100 μL and, once palpable, tumor sizes were recorded every other day via caliper. Tumor volume was calculated using the formula: 0.5236 x width^2 x length, where the width is the smaller dimension of the tumor. Once WT tumors reached between 1000 and 1500 mm^3 in size, the tumors were surgically removed from mice in both arms of the study (WT and CCN4 KO) after euthanasia and processed into single cell suspensions. This normally occurred at Day 14 with the B16F0 model and at Day 27 with the YUMM1.7 model. Seven tumors were processed separately for each YUMM1.7 variant while four tumors were processed for each B16F0 variant. Single-cell suspensions were obtained by enzymatically digesting the excised tumors using the Tumor Dissociation Kit and gentleMACS C system (Miltenyi Biotec, Auburn, CA). In addition to following the manufacturer's instructions, the gentleMACS program 37C_m_TDK_1 was used for B16F0 tumors and 37C_m_TDK_2 was used for YUMM1.7 tumors. Following lysing of the red blood cells, the remaining single-cell suspensions were washed and stained with Live/Dead Fixable Pacific Blue Dead Cell Stain Kit (ThermoFisher). Following blocking with Mouse BD Fc Block (BD Biosciences), the surface of the cells were stained with one of three different antibody mixes that focused on T cells (CD45, CD3, CD4, CD8, and PD1), NK and B cells (CD45, CD3, B220, NK11, DX5, and PD1), and myeloid cells (CD45, CD11b, CD11c, Gr-1, F4/80, and MHCII) and quantified by flow cytometry. The specific antibodies used are listed in Table S2.

To assess the impact of CCN4 on T cell proliferation in vitro, splenocytes were obtained from naïve C57BL/6 mice and stained with CellTrace Pacific Blue Cell Proliferation Kit (ThermoFisher). Stained splenocytes (2.5×10^5) were stimulated for 3 days in 96 well plate with MACSiBeads loaded with anti-mouse CD3 and anti-mouse CD28 antibodies (AP beads, Miltenyi Biotec), at a 1:1 proportion. Fresh serum-free DMEM media conditioned for 24 hours by either confluent wild-type (WT TCM) or confluent CCN4 KO (CCN4 KO TCM) melanoma B16F0 cells were collected, centrifuged to remove cells and cell debris, and added at 50% final volume during T cell stimulation with AP beads. In addition, splenocytes were either left unstimulated or stimulated with AP beads alone, or stimulated in the presence of recombinant mouse CCN4 (rCCN4, R&D) at a final concentration of 10 ng/mL. Each experimental arm was performed in biological triplicate. After 72h, cells were washed and stained with Live/Dead Fixable Green Dead Cell Stain Kit (ThermoFisher). Surface staining with anti-mouse CD8/APC (Miltenyi Biotec), anti-mouse CD4/APC-Cy7 (BD Biosciences), anti-mouse CD62L/PE (eBioscience, ThermoFisher) and anti-mouse CD44/PerCP-Cy5.5 (eBioscience, ThermoFisher) was performed after incubating the cells with Mouse BD Fc Block (BD Biosciences). The proliferation of both CD4$^+$ and CD8$^+$ T cells were quantified by flow cytometry. Results representative of one of two independent experiments.

In vitro suppression of CD8$^+$ T cell function. Inducible mouse CCN4 expression lentiviral vector (IDmCCN4) was constructed with Gateway cloning using Tet-on destination lentiviral vector pCW57.1 (Addgene Plasmid #41393, a gift from David Root) and pShuttle Gateway PLUS ORF Clone for mouse CCN4 (GC-Mm21303, GeneCopoeia). Lentiviruses were packaged as described (33) to transduce YUMM1.7 cell with *Ccn4* CRISPR knockout (Ym1.7-KO1) (33). After puromycin selection, two pools of cells with inducible mCCN4 (Ym1.7-KO1-IDmCCN4) or vector control (Ym1.7-KO1-IDvector) were obtained. ELISA tests with doxycycline (Dox, final 0.5μg/ml) induction revealed the mCCN4 expression was under stringent control and the secreted protein was in the similar level as compared with wild-type YUMM1.7 cells (see Fig. S9).

To generate YUMM1.7-reactive CD8+ T cells, healthy C57BL/6Ncrl mice were inoculated subcutaneously with irradiated YUMM1.7 cells (10^5/mouse), followed by live YUMM1.7 cells (3×10^5/mouse) 3 weeks later. The mice without tumor growth in the next five weeks were maintained. Three days before the assay, the mice were injected again with live YUMM1.7 cells (10^5/mouse). On the day of assay, these mice were euthanized and the YUMM1.7-reactive cells were isolated from mouse splenocytes using mouse CD8a+ T Cell Isolation Kit (130-104-075, Miltenyi Biotec), resuspended in a concentration of 10^6/ml. 50μl (5×10^4) of the YUMM1.7-reactive CD8+ T cells were aliquoted into each well on a 96-well plate for ELISpot assay using Mouse IFNγ/TNFα Double-Color ELISpot kit (Cellular Technology Limited, CTL) following manufacturer's instructions. Briefly, target tumor cells were stimulated with IFNγ (200U/ml, or, 20ng/ml) for 24 hours, harvested and resuspended in a concentration of 2×10^6/ml. 50μl (10^5) of indicated tumor cells in triplicates were aliquoted into each well, with or without doxycycline (Dox, 0.5μg/ml). The reactions were incubated at 37°C for 24 hours and colored spots were developed (red for IFNγ and blue for TNFα). The spots were counted and imaged using an Olympus MVX10 Microscope and the result was plotted and analyzed by GraphPad Prism (version 5).

Flow Cytometry. Single cell suspensions described above were stained with specific antibodies or isotype controls using conventional protocols. Fluorescence-activated cell counting was performed using a BD LSRFortessa and FACSDiva software (BD Biosciences), where the fluorescence intensity for each parameter was reported as a pulse area with 18-bit resolution. Unstained samples were used as negative flow cytometry controls. Single-stain controls were used to establish fluorescence compensation parameters. For TIL

analysis, greater than 5×10^5 events were acquired in each antibody panel in each biological replicate. In analyzing enriched cell populations, 2×10^4 events were acquired in each biological replicate. Flow cytometric data were exported as FCS3.0 files and analyzed using R/Bioconductor (V3.6.1), as described previously (79). The typical gating strategies for T cells, NK and B cells, and myeloid cells are shown in supplementary Figures S3-S5, respectively. The statistical difference in tumor infiltrating lymphocytes between WT and CCN4 KO variants was assessed using log-transformed values and a two-tailed homoscedastic Student's t test. Cell proliferation was quantified using four metrics: fraction diluted (Dil), Precursor frequency, %dividing cells (PF), Proliferation index (PI), and proliferation variance (SD^D) (80). Statistical differences among these proliferation parameters were assessed using type III repeated measures ANOVA in the "car" (V3.0-7) package in R. A p-value < 0.05 was considered statistically significant.

Data Availability. The key datasets used in the analysis can be obtained from the following sources:

- The datasets supporting the conclusions of this article are available in Gene Expression Omnibus repository with the following GEO accession number: GSE98394.
- Transcriptomics profiling of bulk tissue samples using Illumina RNA sequencing for the breast cancer (BRCA) and cutaneous melanoma (SKCM) arms of the Cancer Genome Atlas were downloaded from TCGA data commons, where values for gene expression were expressed in counts, using the "TCGAbiolinks" (V2.8.2) package in R (V3.6.1).

Code Availability. The code used in the analysis can be obtained from the following GitHub repository:

- https://github.com/KlinkeLab/CellNetwork_2020

ACKNOWLEDGEMENTS

This work was supported by National Science Foundation (NSF CBET-1644932 to DJK) and National Cancer Institute (NCI 1R01CA193473 to DJK). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NSF or NCI. We also used equipment from the WVU Flow Cytometry & Single Cell core, which was supported by the National Institutes of Health Grants GM103488/RR032138, GM104942, GM103434, and OD016165.

AUTHOR CONTRIBUTIONS

These contributions follow the International Committee of Medical Journal Editors guidelines: http://www.icmje.org/recommendations/. Conceptualization: DJK; Study Design: DJK; Data Acquisition: DJK, AF, and WD; Data Analysis: DJK, HL, and WD; Data Interpretation: DJK; Funding acquisition: DJK; Methodology: DJK; Project administration: DJK; Software: DJK and ACP; Supervision: DJK; Writing - original draft: DJK; Writing - review & editing: all authors.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Bibliography

1. D. J. Klinke. Eavesdropping on altered cell-to-cell signaling in cancer by secretome profiling. *Mol Cell Oncol*, 3(1):e1029061, Jan 2016.
2. K. A. Hoadley, C. Yau, T. Hinoue, D. M. Wolf, A. J. Lazar, E. Drill, R. Shen, A. M. Taylor, A. D. Cherniack, V. Thorsson, R. Akbani, R. Bowlby, C. K. Wong, M. Wiznerowicz, F. Sanchez-Vega, A. G. Robertson, B. G. Schneider, M. S. Lawrence, H. Noushmehr, T. M. Malta, Cancer Genome Atlas Network, J. M. Stuart, C. C. Benz, and P. W. Laird. Cell-of-Origin Patterns Dominate the Molecular Classification of 10,000 Tumors from 33 Types of Cancer. *Cell*, 173(2):291–304, 04 2018.
3. A. Wells and H. S. Wiley. A systems perspective of heterocellular signaling. *Essays Biochem.*, 62(4):607–617, 10 2018.
4. E. Papalexi and R. Satija. Single-cell RNA sequencing to explore immune cell heterogeneity. *Nat. Rev. Immunol.*, 18(1):35–45, Jan 2018.
5. M. Singer and A. C. Anderson. Revolutionizing Cancer Immunology: The Power of Next-Generation Sequencing Technologies. *Cancer Immunol Res*, 7(2):168–173, Feb 2019.
6. D. Grun, L. Kester, and A. van Oudenaarden. Validation of noise models for single-cell transcriptomics. *Nat. Methods*, 11(6):637–640, Jun 2014.
7. T. Stuart and R. Satija. Integrative single-cell analysis. *Nat. Rev. Genet.*, 20(5):257–272, May 2019.
8. A. M. Newman, C. B. Steen, C. L. Liu, A. J. Gentles, A. A. Chaudhuri, F. Scherer, M. S. Khodadoust, M. S. Esfahani, B. A. Luca, D. Steiner, M. Diehn, and A. A. Alizadeh. Determining cell type abundance and expression from bulk tissues with digital cytometry. *Nat. Biotechnol.*, 37(7):773–782, 07 2019.
9. S. Shen-Orr, R. Tibshirani, and A.J. Butte. Gene expression deconvolution in linear space. *Nature Methods*, 9(1):9, 01 2012.
10. K. Yoshihara, M. Shahmoradgoli, E. Martínez, R. Vegesna, H. Kim, W. Torres-Garcia, V. Treviño, H. Shen, P.W. Laird, D.A. Levine, S. L. Carter, G. Getz, K. Stemke-hale, G. B. Mills, and R.G. Verhaak. Inferring tumour purity and stromal and immune cell admixture from expression data. *Nature Communications*, 4:2612, 10 2013.
11. Z. Wang, S. Cao, J. S. Morris, J. Ahn, R. Liu, S. Tyekucheva, F. Gao, B. Li, W. Lu, X. Tang, I. I. Wistuba, M. Bowden, L. Mucci, M. Loda, G. Parmigiani, C. C. Holmes, and W. Wang. Transcriptome Deconvolution of Heterogeneous Tumor Samples with Immune Infiltration. *iScience*, 9:451–460, Nov 2018.
12. K. Zaitsev, M. Bambouskova, A. Swain, and M.N. Artyomov. Complete deconvolution of cellular mixtures based on linearity of transcriptional signatures. *Nature communications*, 10(1):2209, May 17 2019.
13. S. S. Shen-Orr, R. Tibshirani, P. Khatri, D. L. Bodian, F. Staedtler, N. M. Perry, T. Hastie, M. M. Sarwal, M. M. Davis, and A. J. Butte. Cell type-specific gene expression differences in complex tissues. *Nat. Methods*, 7(4):287–289, Apr 2010.
14. E. Becht, N. A. Giraldo, L. Lacroix, B. Buttard, N. Elarouci, F. Petitprez, J. Selves, P. Laurent-Puig, C. Sautes-Fridman, W. H. Fridman, and A. de Reynies. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biol.*, 17(1):218, Oct 2016.
15. M. Schelker, S. Feau, J. Du, N. Ranu, E. Klipp, G. MacBeath, B. Schoeberl, and A. Raue. Estimation of immune cell content in tumour tissue using single-cell RNA-seq data. *Nat Commun*, 8(1):2032, Dec 2017.
16. A. Torang, P. Gupta, and D. J. Klinke. An elastic-net logistic regression approach to generate classifiers and gene signatures for types of immune cells and T helper cell subsets. *BMC Bioinformatics*, 20(1):433, Aug 2019.
17. T. Z. Tan, Q. H. Miow, Y. Miki, T. Noda, S. Mori, R. Y. Huang, and J. P. Thiery. Epithelial-mesenchymal transition spectrum quantification and its efficacy in deciphering survival and drug responses of cancer patients. *EMBO Mol Med*, 6(10):1279–1293, Oct 2014.
18. S. Koplev, K. Lin, A. B. Dohlman, and A. Ma'ayan. Integration of pan-cancer transcriptomics with RPPA proteomics reveals mechanisms of epithelial-mesenchymal transition. *PLoS Comput. Biol.*, 14(1):e1005911, 01 2018.
19. T. M. Malta, A. Sokolov, A. J. Gentles, T. Burzykowski, L. Poisson, J. N. Weinstein, B. Kami/'nska, J. Huelsken, L. Omberg, O. Gevaert, A. Colaprico, P. Czerwi/'nska, S. Mazurek, L. Mishra, H. Heyn, A. Krasnitz, A. K. Godwin, A. J. Lazar, Cancer Genome Atlas Research Network, K. A. Stuart, J. M. Hoadley, P. W. Laird, H. Noushmehr, and M. Wiznerowicz. Machine Learning Identifies Stemness Features Associated with Oncogenic Dedifferentiation. *Cell*, 173(2):338–354, 04 2018.
20. J. T. George, M. K. Jolly, S. Xu, J. A. Somarelli, and H. Levine. Survival Outcomes in Cancer Patients Predicted by a Partial EMT Gene Expression Scoring Metric. *Cancer Res.*, 77(22): 6415–6428, 11 2017.
21. D. J. Klinke and A. Torang. An unsupervised feature extraction and selection strategy for identifying epithelial-mesenchymal transition state metrics in breast cancer and melanoma. *iScience*, 23(5):101080, 2020.
22. M. Tosolini, A. Kirilovsky, B. Mlecnik, T. Fredriksen, S. Mauger, G. Bindea, A. Berger, P. Bruneval, W. Fridman, F. Pagos, and J. Galon. Clinical impact of different classes of infiltrating t cytotoxic and helper cells (th1, th2, treg, th17) in patients with colorectal cancer. *Cancer Research*, 71(4):1263–1271, February 2011.
23. V. Thorsson, D. L. Gibbs, S. D. Brown, D. Wolf, D. S. Bortone, T. H. Ou Yang, E. Porta-Pardo, G. F. Gao, C. L. Plaisier, J. A. Eddy, E. Ziv, A. C. Culhane, E. O. Paull, I. K. A. Sivakumar, A. J. Gentles, R. Malhotra, F. Farshidfar, A. Colaprico, J. S. Parker, L. E. Mose, N. S. Vo, J. Liu, Y. Liu, J. Rader, V. Dhankani, S. M. Reynolds, R. Bowlby, A. Califano, A. D. Cherniack, D. Anastassiou, D. Bedognetti, A. Rao, K. Chen, A. Krasnitz, H. Hu, T. M. Malta, H. Noushmehr, C. S. Pedamallu, S. Bullman, A. I. Ojesina, A. Lamb, W. Zhou, H. Shen, T. K. Choueiri, J. N. Weinstein, J. Guinney, J. Saltz, and et al. The Immune Landscape of Cancer. *Immunity*, 48(4):812–830, 04 2018.
24. S. M. Hill, L. M. Heiser, T. Cokelaer, M. Unger, N. K. Nesser, D. E. Carlin, Y. Zhang, A. Sokolov, E. O. Paull, C. K. Wong, K. Graim, A. Bivol, H. Wang, F. Zhu, B. Afsari, L. V. Danilova, A. V. Favorov, W. S. Lee, D. Taylor, C. W. Hu, B. L. Long, D. P. Noren, A. J. Bisberg, G. B. Mills, J. W. Gray, M. Kellen, T. Norman, S. Friend, A. A. Qutub, E. J. Fertig, Y. Guan, M. Song, J. M. Stuart, P. T. Spellman, H. Koeppl, G. Stolovitzky, J. Saez-Rodriguez, S. Mukherjee, B. Afsari, R. Al-Ouran, B. Anton, T. Arodz, O. A. Sichani, N. Bagheri, N. Berlow, A. J. Bisberg, A. Bivol, A. Bohler, J. Bonet, R. Bonneau, G. Budak, R. Bunescu, M. Caglar, B. Cai, C. Cai, D. E. Carlin, A. Carlon, L. Chen, M. F. Ciaccio, T. Cokelaer, G. Cooper, S. Coort, C. J. Creighton, S. M. Daneshmand, A. de la Fuente, B. Di Camillo, L. V. Danilova, J. Dutta-Moscato, K. Emmett, C. Evelo, M. K. Fassia, A. V. Favorov, E. J. Fertig, J. D. Finkle, F. Finotello, S. Friend, X. Gao, J. Gao, J. Garcia-Garcia, S. Ghosh, A. Giaretta, K. Graim, J. W. Gray, R. Grosseholz, Y. Guan, J. Guinney, C. Hafemeister, O. Hahn, S. Haider, T. Hase, L. M. Heiser, S. M. Hill, J. Hodgson, B. Hoff, C. H. Hsu, C. W. Hu, Y. Hu, X. Huang, M. Jalili, X. Jiang, T. Kacprowski, L. Kaderali, M. Kang, V. Kannan, M. Kellen, K. Kikuchi, D. C. Kim, H. Kitano, B. Knapp, G. Komatsoulis, H. Koeppl, A. Kramer, M. B. Kursa, M. Kutmon, W. S. Lee, Y. Li, X. Liang, Z. Liu, Y. Liu, B. L. Long, S. Lu, X. Lu, M. Manfrini, M. R. Matos, D. Meerzaman, G. B. Mills, W. Min, S. Mukherjee, C. L. Muller, R. E. Neapolitan, N. K. Nesser, D. P. Noren, T. Norman, B. Oliva, S. O. Opiyo, R. Pal, A. Palinkas, E. O. Paull, J. Planas-Iglesias, D. Poglayen, A. A. Qutub, J. Saez-Rodriguez, F. Sambo, T. Sanavia, A. Sharifi-Zarchi, J. Slawek, A. Sokolov, M. Song, P. T.

Spellman, A. Streck, G. Stolovitzky, S. Strunz, J. M. Stuart, D. Taylor, J. Tegner, K. Thobe, G. M. Toffolo, E. Trifoglio, M. Unger, Q. Wan, H. Wang, L. Welch, C. K. Wong, J. J. Wu, A. Y. Xue, R. Yamanaka, C. Yan, S. Zairis, M. Zengerling, H. Zenil, S. Zhang, Y. Zhang, F. Zhu, and Z. Zi. Inferring causal molecular networks: empirical assessment through a community-based effort. *Nat. Methods*, 13(4):310–318, Apr 2016.

25. N. Friedman. Inferring cellular networks using probabilistic graphical models. *Science*, 303 (5659):799–805, February 2004.
26. M. Scutari. Learning Bayesian Networks with the bnlearn R Package. *J Stat Software*, 35: 1–22, 2010.
27. K. Sachs, D. Gifford, T. Jaakkola, P. Sorger, and D. A. Lauffenburger. Bayesian network approach to cell signaling pathway modeling. *Science Signaling*, 2002(148):pe38, September 2002.
28. B.-E. Perrin, L. Ralaivola, A. Mazurie, S. Bottani, J. Mallet, and F. dlchBuc. Gene networks inference using dynamic bayesian networks. *Bioinformatics*, 19(suppl 2):ii138–ii148, September 2003.
29. K. Sachs, O. Perez, D. Pe'er, D. A. Lauffenburger, and G. P. Nolan. Causal protein-signaling networks derived from multiparameter single-cell data. *Science*, 308(5721):523–529, April 2005.
30. K. Sachs, S. Itani, J. Carlisle, G. P. Nolan, D. Pe'er, and D. A. Lauffenburger. Learning signaling network structures with sparsely distributed data. *J Comp Biol*, 16(2):201–212, February 2009.
31. M. Zou and S. D. Conzen. A new dynamic bayesian network (DBN) approach for identifying gene regulatory networks from time course microarray data. *Bioinformatics*, 21(1):71–79, January 2005.
32. D. J. Klinke. Induction of Wnt-inducible signaling protein-1 correlates with invasive breast cancer oncogenesis and reduced type 1 cell-mediated cytotoxic immunity: a retrospective study. *PLoS Comput. Biol.*, 10(1):e1003409, Jan 2014.
33. W. Deng, A. Fernandez, S. L. McLaughlin, and D. J. Klinke. WNT1-inducible signaling pathway protein 1 (WISP1/CCN4) stimulates melanoma invasion and metastasis by promoting the epithelial-mesenchymal transition. *J. Biol. Chem.*, 294(14):5261–5280, 04 2019.
34. W. Deng, A. Fernandez, S. L. McLaughlin, and D. J. Klinke. Cell Communication Network Factor 4 (CCN4/WISP1) Shifts Melanoma Cells from a Fragile Proliferative State to a Resilient Metastatic State. *Cell Mol Bioeng*, 13(1):45–60, Feb 2020.
35. D. Hanahan and R. A. Weinberg. Hallmarks of cancer: the next generation. *Cell*, 144(5): 646–674, Mar 2011.
36. I. Tirosh, B. Izar, S. M. Prakadan, M. H. Wadsworth, D. Treacy, J. J. Trombetta, A. Rotem, C. Rodman, C. Lian, G. Murphy, M. Fallahi-Sichani, K. Dutton-Regester, J. R. Lin, O. Cohen, P. Shah, D. Lu, A. S. Genshaft, T. K. Hughes, C. G. Ziegler, S. W. Kazer, A. Gaillard, K. E. Kolb, A. C. Villani, C. M. Johannessen, A. Y. Andreev, E. M. Van Allen, M. Bertagnolli, P. K. Sorger, R. J. Sullivan, K. T. Flaherty, D. T. Frederick, J. Jane-Valbuena, C. H. Yoon, O. Rozenblatt-Rosen, A. K. Shalek, A. Regev, and L. A. Garraway. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science*, 352(6282): 189–196, Apr 2016.
37. S. A. Rosenberg and N. P. Restifo. Adoptive cell transfer as personalized immunotherapy for human cancer. *Science*, 348(6230):62–68, Apr 2015.
38. S. C. Wei, C. R. Duffy, and J. P. Allison. Fundamental Mechanisms of Immune Checkpoint Blockade Therapy. *Cancer Discov*, 8(9):1069–1086, 09 2018.
39. D. Marbach, J. C. Costello, R. Küffner, N. M. Vega, R. J. Prill, D. M. Camacho, K. R. Allison, M. Kellis, J. J. Collins, G. Stolovitzky, A. Aderhold, K. R. Allison, R. Bonneau, D. M. Camacho, Y. Chen, J. J. Collins, F. Cordero, J. C. Costello, M. Crane, F. Dondelinger, M. Drton, R. Esposito, R. Foygel, A. de la Fuente, J. Gertheiss, P. Geurts, A. Greenfield, M. Grzegorczyk, A. C. Haury, B. Holmes, T. Hothorn, D. Husmeier, V. A. Huynh-Thu, A. Irrthum, M. Kellis, G. Karlebach, R. Küffner, S. Lèbre, V. De Leo, A. Madar, S. Mani, D. Marbach, F. Mordelet, H. Ostrer, Z. Ouyang, R. Pandya, T. Petri, A. Pinna, C. S. Poultney, R. J. Prill, S. Rezny, H. J. Ruskin, Y. Saeys, R. Shamir, A. Sîrbu, M. Song, N. Soranzo, A. Statnikov, G. Stolovitzky, N. Vega, P. Vera-Licona, J. P. Vert, A. Visconti, H. Wang, L. Wehenkel, L. Windhager, Y. Zhang, and R. Zimmer. Wisdom of crowds for robust gene network inference. *Nat. Methods*, 9(8):796–804, Jul 2012.
40. D. Marbach, R. J. Prill, T. Schaffter, C. Mattiussi, D. Floreano, and G. Stolovitzky. Revealing strengths and weaknesses of methods for gene network inference. *Proc. Natl. Acad. Sci. U.S.A.*, 107(14):6286–6291, Apr 2010.
41. I. Tsamardinos, C.F. Aliferis, and Statnikov. A. Algorithms for Large Scale Markov Blanket Discovery. *Proc Sixteenth Int'l Florida Artificial Intelligence Research Society Conf.*, pages 376–381, 2003.
42. M. Gasse, A. Aussem, and H. Elghazel. A Hybrid Algorithm for Bayesian Network Structure Learning with Application to Multi-Label Learning. *Expert Systems with Applications*, 41 (15):6755–6772, 2014.
43. D. Colombo and M.H. Maathuis. Order-Independent Constraint-Based Causal Structure Learning. *Journal of Machine Learning Research*, 14:3921–3962, 2014.
44. D. Margaritis. *Learning Bayesian Network Model Structure from Data.* PhD thesis, School of Computer Science, Carnegie-Mellon University, Pittsburgh, PA, 2003.
45. S.J. Russell and P. Norvig. *Artificial Intelligence: A Modern Approach.* Prentice Hall, 3rd edition, 2009.
46. I. Tsamardinos, L.E. Brown, and C.F. Aliferis. The Max-Min Hill-Climbing Bayesian Network Structure Learning Algorithm. *Machine Learning*, 65(1):31–78, 2006.
47. N. Friedman, I. Nachman, and D. Pe'er. Learning Bayesian Network Structure from Massive Datasets: the Sparse Candidate Algorithm. *Proceedings of the Fifteenth Conference on Uncertainty in Artificial Intelligence (UAI)*, pages 206–215, 1999.
48. F.C. Aliferis, A. Statnikov, I. Tsamardinos, M. Subramani, and X.D. Koutsoukos. Local Causal and Markov Blanket Induction for Causal Discovery and Feature Selection for Classification Part I: Algorithms and Empirical Evaluation. *Journal of Machine Learning Research*, 11:171–234, 2010.
49. M. Scutari and R. Nagarajan. Identifying significant edges in graphical models of molecular networks. *Artif Intell Med*, 57(3):207–217, Mar 2013.
50. U. Alon. Network motifs: theory and experimental approaches. *Nat Rev Genet*, 8(6):450–461, Jun 2007.
51. A. Fernandez, W. Deng, S.L. McLaughlin, A.C. Pirkey, S.L. Rellick, and D. J. Klinke. Cell communication network factor 4 (ccn4/wisp1) promotes tumor-induced immunosuppression in melanoma. *bioRxiv*, page (submitted), 2020.
52. H. Jeong, I. Hwang, S. H. Kang, H. C. Shin, and S. Y. Kwon. Tumor-Associated Macrophages as Potential Prognostic Biomarkers of Invasive Breast Cancer. *J Breast Cancer*, 22(1):38–51, Mar 2019.
53. A. C. Huang, M. A. Postow, R. J. Orlowski, R. Mick, B. Bengsch, S. Manne, W. Xu, S. Harmon, J. R. Giles, B. Wenz, M. Adamow, D. Kuk, K. S. Panageas, C. Carrera, P. Wong, F. Quagliarello, B. Wubbenhorst, K. D'Andrea, K. E. Pauken, R. S. Herati, R. P. Staupe, J. M. Schenkel, S. McGettigan, S. Kothari, S. M. George, R. H. Vonderheide, R. K. Amaravadi, G. C. Karakousis, L. M. Schuchter, X. Xu, K. L. Nathanson, J. D. Wolchok, T. C. Gangadhar, and E. J. Wherry. T-cell invigoration to tumour burden ratio associated with anti-PD-1 response. *Nature*, 545(7652):60–65, 05 2017.
54. C. Twyman-Saint Victor, A. J. Rech, A. Maity, R. Rengan, K. E. Pauken, E. Stelekati, J. L. Benci, B. Xu, H. Dada, P. M. Odorizzi, R. S. Herati, K. D. Mansfield, D. Patsch, R. K. Amaravadi, L. M. Schuchter, H. Ishwaran, R. Mick, D. A. Pryma, X. Xu, M. D. Feldman, T. C. Gangadhar, S. M. Hahn, E. J. Wherry, R. H. Vonderheide, and A. J. Minn. Radiation and dual checkpoint blockade activate non-redundant immune mechanisms in cancer. *Nature*, 520(7547):373–377, Apr 2015.
55. H. Li, A. M. van der Leun, I. Yofe, Y. Lubling, D. Gelbard-Solodkin, A. C. J. van Akkooi, M. van den Braber, E. A. Rozeman, J. B. A. G. Haanen, C. U. Blank, H. M. Horlings, E. David, Y. Baran, A. Bercovich, A. Lifshitz, T. N. Schumacher, A. Tanay, and I. Amit. Dysfunctional CD8 T Cells Form a Proliferative, Dynamically Regulated Compartment within Human Melanoma. *Cell*, 176(4):775–789, 02 2019.
56. H. Moore and R. Allen. What Can Mathematics Do for Drug Development? *Bull. Math. Biol.*, 81(9):3421–3424, 09 2019.
57. K. Gadkar, D. C. Kirouac, D. E. Mager, P. H. van der Graaf, and S. Ramanujan. A Six-Stage Workflow for Robust Application of Systems Pharmacology. *CPT Pharmacometrics Syst Pharmacol*, 5(5):235–249, 05 2016.
58. S. Palsson, T. P. Hickling, E. L. Bradshaw-Pierce, M. Zager, K. Jooss, P. J. O'Brien, M. E. Spilker, B. O. Palsson, and P. Vicini. The development of a fully-integrated immune response model (FIRM) simulator of the immune response through integration of multiple subset models. *BMC Syst Biol*, 7:95, Sep 2013.
59. O. Milberg, C. Gong, M. Jafarnejad, I. H. Bartelink, B. Wang, P. Vicini, R. Narwal, L. Roskos, and A. S. Popel. A QSP Model for Predicting Clinical Responses to Monotherapy, Combination and Sequential Therapy Following CTLA-4, PD-1, and PD-L1 Checkpoint Blockade. *Sci Rep*, 9(1):11286, Aug 2019.
60. D. J. Klinke and Q. Wang. Inferring the Impact of Regulatory Mechanisms that Underpin CD8+ T Cell Control of B16 Tumor Growth In vivo Using Mechanistic Models and Simulation. *Front Pharmacol*, 7:515, 2016.
61. L. A. Chylek, L. A. Harris, C. S. Tung, J. R. Faeder, C. F. Lopez, and W. S. Hlavacek. Rule-based modeling: a computational approach for studying biomolecular site dynamics in cell signaling systems. *Wiley Interdiscip Rev Syst Biol Med*, 6(1):13–36, 2014.
62. J. A. Sekar and J. R. Faeder. Rule-based modeling of signal transduction: a primer. *Methods Mol. Biol.*, 880:139–218, 2012.
63. P. Boutillier, M. Maasha, X. Li, H. F. Medina-Abarca, J. Krivine, J. Feret, I. Cristescu, A. G. Forbes, and W. Fontana. The Kappa platform for rule-based modeling. *Bioinformatics*, 34 (13):i583–i592, 07 2018.
64. Sergio Vernuccio and Linda J. Broadbelt. Discerning complex reaction networks using automated generators. *AIChE Journal*, 65(8):e16663, 2019.
65. E. Salazar-Cavazos, C. F. Nitta, E. D. Mitra, B. S. Wilson, K. A. Lidke, W. S. Hlavacek, and D. S. Lidke. Multisite EGFR phosphorylation is regulated by adaptor protein abundances and dimer lifetimes. *Mol. Biol. Cell*, 31(7):695–708, Mar 2020.
66. M. C. Pharris, N. M. Patel, T. G. VanDyk, T. M. Bartol, T. J. Sejnowski, M. B. Kennedy, M. I. Stefan, and T. L. Kinzer-Ursem. A multi-state model of the CaMKII dodecamer suggests a role for calmodulin in maintenance of autophosphorylation. *PLoS Comput. Biol.*, 15(12): e1006941, 12 2019.
67. J. A. Rohrs, D. Zheng, N. A. Graham, P. Wang, and S. D. Finley. Computational Model of Chimeric Antigen Receptors Explains Site-Specific Phosphorylation Kinetics. *Biophys. J.*, 115(6):1116–1129, 09 2018.
68. M. Hwang, M. Garbey, S. A. Berceli, and R. Tran-Son-Tay. Rule-Based Simulation of Multi-Cellular Biological Systems-A Review of Modeling Techniques. *Cell Mol Bioeng*, 2(3):285–294, Sep 2009.
69. Á. G. López, J. M. Seoane, and M. A. F. Sanjuán. Dynamics of the cell-mediated immune response to tumour growth. *Philos Trans A Math Phys Eng Sci*, 375(2096), Jun 2017.
70. D. G. Mallet and L. G. De Pillis. A cellular automata model of tumor-immune system interactions. *J. Theor. Biol.*, 239(3):334–350, Apr 2006.
71. J. L. Kaiser, C. L. Bland, and D. J. Klinke. Identifying causal networks linking cancer processes and anti-tumor immunity using Bayesian network inference and metagene constructs. *Biotechnol. Prog.*, 32(2):470–479, 03 2016.
72. S. Ramanujan, J. R. Chan, C. M. Friedrich, and C. J. Thalhauser. A Flexible Approach for Context-Dependent Assessment of Quantitative Systems Pharmacology Models. *CPT Pharmacometrics Syst Pharmacol*, 8(6):340–343, Jun 2019.
73. X. Zhou, R. A. Franklin, M. Adler, J. B. Jacox, W. Bailis, J. A. Shyer, R. A. Flavell, A. Mayo, U. Alon, and R. Medzhitov. Circuit Design Features of a Stable Two-Cell System. *Cell*, 172 (4):744–757, 02 2018.
74. J. Pearl. Direct and indirect effects. In *Proc Am Stat Assoc Joint Stat Meetings (Minneapolis, MN)*, pages 1572–1581. MIRA Digital Publishing, 2005.
75. J. D. Finkle, J. J. Wu, and N. Bagheri. Windowed Granger causal inference strategy improves discovery of gene regulatory networks. *Proc. Natl. Acad. Sci. U.S.A.*, 115(9):2252–2257, 02 2018.
76. P. Li, P. Gong, H. Li, E. J. Perkins, N. Wang, and C. Zhang. Gene regulatory network inference and validation using relative change ratio analysis and time-delayed dynamic Bayesian network. *EURASIP J Bioinform Syst Biol*, 2014:12, Dec 2014.
77. E. Eisenberg and E. Y. Levanon. Human housekeeping genes, revisited. *Trends Genet.*, 29 (10):569–574, Oct 2013.

78. K. Meeth, J. X. Wang, G. Micevic, W. Damsky, and M. W. Bosenberg. The YUMM lines: a series of congenic mouse melanoma cell lines with defined genetic alterations. *Pigment Cell Melanoma Res*, 29(5):590–597, 09 2016.
79. D. J. Klinke and K. M. Brundage. Scalable analysis of flow cytometry data using R/Bioconductor. *Cytometry A*, 75(8):699–706, Aug 2009.
80. M. Roederer. Interpretation of cellular proliferation data: avoid the panglossian. *Cytometry A*, 79(2):95–101, Feb 2011.

Figure Legends

Figure 1 - A computational workflow combining digital cytometry with Bayesian network inference to estimate how a genetic alteration associated with disease, such as over-expression of a gene by malignant cells, impacts the heterocellular network within a tissue. Digital cytometry deconvolutes a bulk transcriptomic profile using gene signatures that correspond to different stromal, malignant, and immune cell types. The results estimate the prevalence of the different cell types within the tissue sample, that is the digital cytometry features. By using bulk transcriptomic profiles of defined patient populations, underlying variation in the inferred cellular composition coupled with features associated with a patient sample, such as over-expression of a gene by malignant cells, can be used to estimate how the heterocellular network is impacted by a genetic alteration intrinsic to the malignant cell using Bayesian Network inference. To illustrate the approach, we focused on malignant cell expression of Cell Communication Network factor 4 (CCN4). The resulting directed acyclic graph represents the collective conditional independence among the modeled features, or nodes, of the network.

Figure 2 - Summary of the evidence obtained from the TCGA breast cancer dataset supporting the consensus edges in the seed network. (A) Edges ordered based on the number of algorithms that detected that an edge was enriched (bar graph - left axis) and the strength of enrichment (dotted lines - right axis). The strength of enrichment, that is Edge Strength, corresponds to the probability of a partial correlation between the two nodes of an arc being explained by random chance, given the rest of the network. The lines associated with the strength of enrichment represent the minimum (dashed line) and maximum (dotted line) values obtained by the different algorithms for each edge. Bar graph color indicates whether an edge was significantly enriched with a clear direction and contained within the set of arcs identified at the minimum BIC (green), significantly enriched without a clear direction but contained within the set of arcs identified at the minimum BIC (green/yellow), significantly enriched without a clear direction (yellow), or excluded from the consensus seed network list (tan). (B) Dependence of overall network connectivity (top) and node connectivity (bottom) on the number of arcs included in the whitelist. A threshold value for the edge strength (x-axis) was used to select arcs for including in the whitelist (red curve), which resulted in the inferred DAG connectivity (black circles). Values for the Bayesian Information Criterion (BIC) were calculated for the entire DAG (top - black squares) and for each node given the inferred parents (bottom table: BIC / number of parental arcs). The cells highlighted in green indicate the minimum BIC value and the number of corresponding arcs that were included in the consensus whitelist.

Figure 3 - A directed acyclic graph (DAG) representing the conditional probability distribution inferred using the digital cytometry and sample features extracted from the breast cancer arm of the TCGA. The nodes of the graph represent features, such as CCN4 gene expression (rectangle), sample attribute (hexagon), or the prevalence of a particular cell type/state (oval). The edges represent inferred causal relationships among the nodes. The black lines with arrow heads represent a positive causal relation while red lines with horizontal bars represent a negative or inhibitory causal relation, where the extent of influence of the parental node is annotated by the number beside the edge. The number included within the node symbol represents the average normalized value of the digital cytometry feature within the dataset with values of all of the parental nodes set to zero. The width of the edge is proportional to the posterior probability of inclusion into the DAG.

Figure 4 - Conditional probability query of the BRCA DAG compared against digital cytometry estimates obtained from experimental data. Experimental samples obtained from normal mammary and tumor tissue are shown as blue versus red dots, respectively. Samples of the conditional probability model for "Cancer" < 0.05 (blue contour) and "Cancer" > 0.95 (red contour) for CD8 T cells (A), CD4 T cells (B), active NK cells (C), B cells (D), Macrophages (E) and Cancer Associated Fibroblasts (F). Linear trend lines are superimposed on the conditional probability samples.

Figure 5 - Two DAGs representing the conditional probability distributions inferred using the digital cytometry and sample features extracted from the two melanoma-related datasets. (A) Analysis of a bulk RNAseq dataset obtained from patients with common pigmented nevi and primary melanoma ($n_{samples}$ = 78). (B) Analysis of primary melanoma samples extracted from the SKCM arm of the TCGA ($n_{samples}$ = 94). The DAGs are summarized using similar notation as described in Figure 3. Dotted lines indicate edges that were included in the consensus seed network but, as the samples were all from patients with cancer, had no evidence in the TCGA dataset.

Figure 6 - CCN4 knock-out in two syngeneic mouse models of melanoma induced a similar shift in cytolytic lymphocytes as observed in human breast cancer and melanoma. (A) The percentage of live CD45+ cells isolated from tumors generated by inoculating s.c. with WT (red) and CCN4 KO (blue) variants of B16F0 (o and x's) and YUMM1.7 (□ and +'s) cells, where the log-linear trends are highlighted by dotted lines. CD45+ values were obtained from three different antibody panels that quantified T cells, B/NK cells, and myeloid cells in TIL isolates from each mouse. (B) A comparison of the ratio of NK cells (black), CD8+ T cells (red), CD4+ T cells (blue), and B cells (green) to live CD45+ TILs in s.c. tumors generated using WT B16F0 and YUMM1.7 cells (mean $\pm$ s.d.). (C) The difference in the mean prevalence of the infiltrating immune cell types was compared when CCN4 is present (WT) versus absent (CCN4 KO) as predicted by digital cytometry from the BRCA (dark gray) and SKCM (light gray) datasets and as observed experimentally using the B16F0 (red) and YUMM1.7 (black) mouse models. (D) TIL comparison upon CCN4 KO in B16F0 and YUMM1.7 mouse models stratified by NK cells, CD8+ T cells, CD4+ T cells, and B cells (top to bottom) (n = 7 for YUMM1.7 and n = 4 for B16F0 variants and mean $\pm$ s.d.). p-values calculated between WT and CCN4 KO pairs using Student's t-test.

Figure 7 - Myeloid immune cell subsets differentially infiltrated tumors derived from WT B16F0 and YUMM1.7 cells and shifted in prevalence upon CCN4 knock-out. (A) A comparison of the ratio of CD11c- (black) and CD11c+ (gray) macrophages, Dendritic cells (yellow), CD11c+ MDSC (green), MDSC (blue), and Neutrophils (red) to live CD45+ TILs in s.c. tumors generated using WT B16F0 and YUMM1.7 cells. (B) The difference in prevalence of the myeloid cell types was compared when CCN4 is present (WT) versus absent (CCN4 KO) as predicted by digital cytometry of the BRCA (dark gray) and SKCM (light gray) data sets and as observed experimentally using the

B16F0 (red) and YUMM1.7 (black) mouse models. Macrophages are the only myeloid cell subset inferred from the BRCA and SKCM datasets and are assumed to be related to CD11c+ macrophages in mouse models. (C) A representative scatter plot of GR1 versus CD11c expression in gated live CD45+ CD11b+ TILs obtained from WT (top) and CCN4 KO (bottom) YUMM1.7 tumors. (D - F) TIL comparison upon CCN4 KO in B16F0 and YUMM1.7 mouse models stratified by myeloid-derived suppressor cell subsets (D: MDSC (top) and CD11c+ MDSC (bottom)) and other myeloid cell subsets (E: CD11c- (top) and CD11c+ (bottom) macrophages, F: neutrophils (top) and dendritic cells (bottom)) (n = 7 for YUMM1.7 and n = 4 for B16F0 variants and mean $\pm$ s.d.). p-values calculated between WT and CCN4 KO pairs using Student's t-test.

Figure 8 - CCN4 had no direct effect on T cell proliferation but impaired CD8$^+$ T cell function. The distribution in cell trace staining among live CD4$^+$ (A) and CD8$^+$ (B) T cells stimulated with αCD3/αCD28 (AP beads) alone or in the presence of media conditioned by WT B16F0 cells (AP beads + WT TCM), media conditioned by CCN4 KO B16F0 cells (AP beads + CCN4 KO TCM), or with 10 ng/ml of recombinant mouse CCN4 (AP beads + rCCN4). The distribution in the corresponding unstimulated cells (gray) are shown at the bottom. The colored vertical lines indicate the predicted dilution of cell trace staining in each generation based on the unstimulated controls. (C) CD8+ T cells isolated from the spleens of C57BL/6 mice that rejected YUMM1.7 tumors were cultured in an in vitro ELISpot assay using variants of the YUMM1.7 cell line as targets (WT YUMM1.7 - yellow, CCN4 KO YUMM1.7 - light green, CCN4 KO YUMM1.7 with a blank inducible expression vector - dark green and blue, CCN4 KO YUMM1.7 with a CCN4 inducible expression vector - purple and red). Variants containing the inducible expression vector were also cultured in the absence (dark green and purple) or presence of doxycycline (blue and red). CD8+ T cells expressing IFNγ and TNFα were quantified following 24 hour co-culture (bar graph). Statistical significant between pairs was assessed using a Student's t-test.

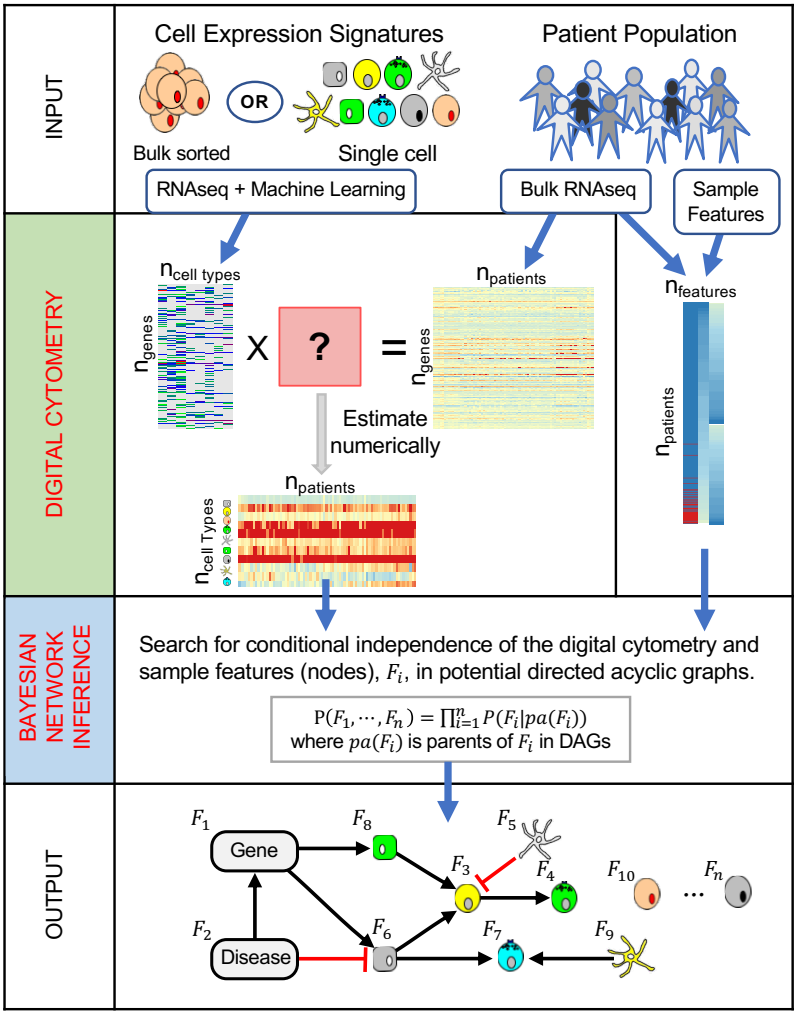
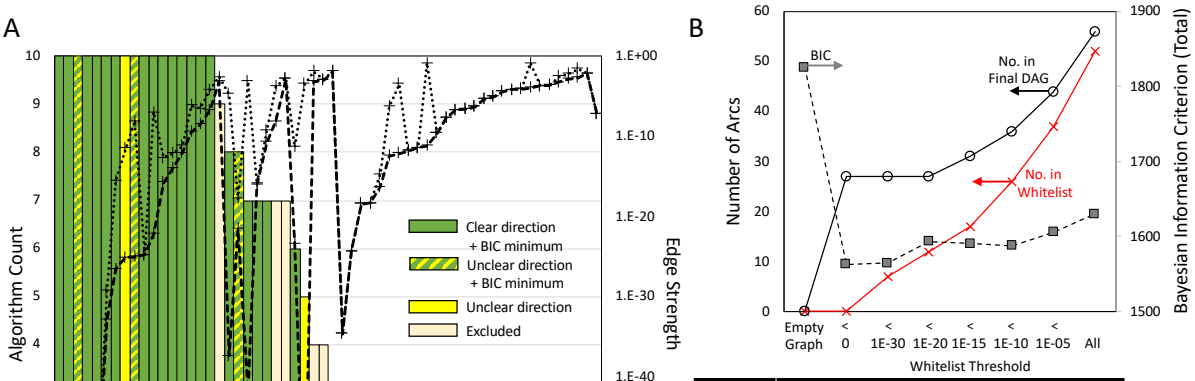


Fig. 1. A computational workflow combining digital cytometry with Bayesian network inference to estimate how a genetic alteration associated with disease, such as over-expression of a gene by malignant cells, impacts the heterocellular network within a tissue.



Node	BIC / Number of Parental Arcs							
Cancer	102.8 / 0	102.8 / 0	102.8 / 0	102.8 / 0	102.8 / 0	102.8 / 0	102.8 / 0	102.8 / 0
CCN4	85.2 / 0	71.0 / 1	71.0 / 1	71.0 / 1	71.0 / 1	71.0 / 1	71.0 / 1	71.0 / 1
B Cells naive	67.8 / 0	67.6 / 2	65.0 / 1	65.0 / 1	65.0 / 1	67.0 / 1	70.4 / 2	72.8 / 3
CAF	61.3 / 0	49.9 / 2	47.4 / 2	47.4 / 2	48.2 / 4	50.2 / 5	50.2 / 5	54.0 / 6
CD4 T cells	195.1 / 0	172.5 / 3	172.5 / 3	172.5 / 3	172.5 / 3	172.5 / 3	172.3 / 4	173.7 / 6
CD8 T cells	72.8 / 0	65.8 / 4	65.8 / 4	65.8 / 4	65.8 / 4	65.8 / 4	67.7 / 5	72.7 / 7
Endothelial cells	142.7 / 0	128.5 / 2	133.6 / 1	128.5 / 2	126.3 / 2	119.4 / 5	119.4 / 5	122.3 / 4
Epithelial	85.9 / 0	80.0 / 1	81.8 / 2	80.0 / 1	80.0 / 1	80.0 / 1	80.0 / 1	80.0 / 1
Macrophages	76.3 / 0	76.3 / 0	76.3 / 0	76.3 / 0	76.3 / 0	76.3 / 0	76.4 / 1	76.4 / 1
Mesenchymal	99.2 / 0	56.8 / 2	57.5 / 1	57.5 / 1	56.8 / 2	58.0 / 3	58.0 / 3	60.0 / 4
Neutrophils	102.4 / 0	105.3 / 3	105.3 / 3	105.3 / 3	105.3 / 3	105.3 / 3	106.5 / 4	107.8 / 5
NK cells active	174.7 / 0	174.7 / 0	174.7 / 0	174.7 / 0	174.7 / 0	174.7 / 0	174.7 / 0	174.0 / 1
NK cells resting	143.8 / 0	119.6 / 1	119.6 / 1	119.6 / 1	119.6 / 1	119.6 / 1	121.0 / 2	122.4 / 3
pM0	123.8 / 0	20.8 / 2	20.8 / 2	73.0 / 2	73.0 / 2	73.0 / 2	115.9 / 3	115.9 / 3
pM1	109.6 / 0	109.6 / 0	108.8 / 1	92.3 / 2	88.6 / 3	88.6 / 3	88.6 / 3	88.6 / 3
pM2	108.1 / 0	93.1 / 1	93.1 / 1	93.1 / 1	93.1 / 1	93.1 / 1	60.5 / 2	60.5 / 2
Proliferation	74.2 / 0	68.4 / 3	68.7 / 4	68.4 / 3	71.5 / 3	70.5 / 3	70.5 / 3	75.3 / 6

Fig. 2. Summary of the evidence obtained from the TCGA breast cancer dataset supporting the consensus edges in the seed network.

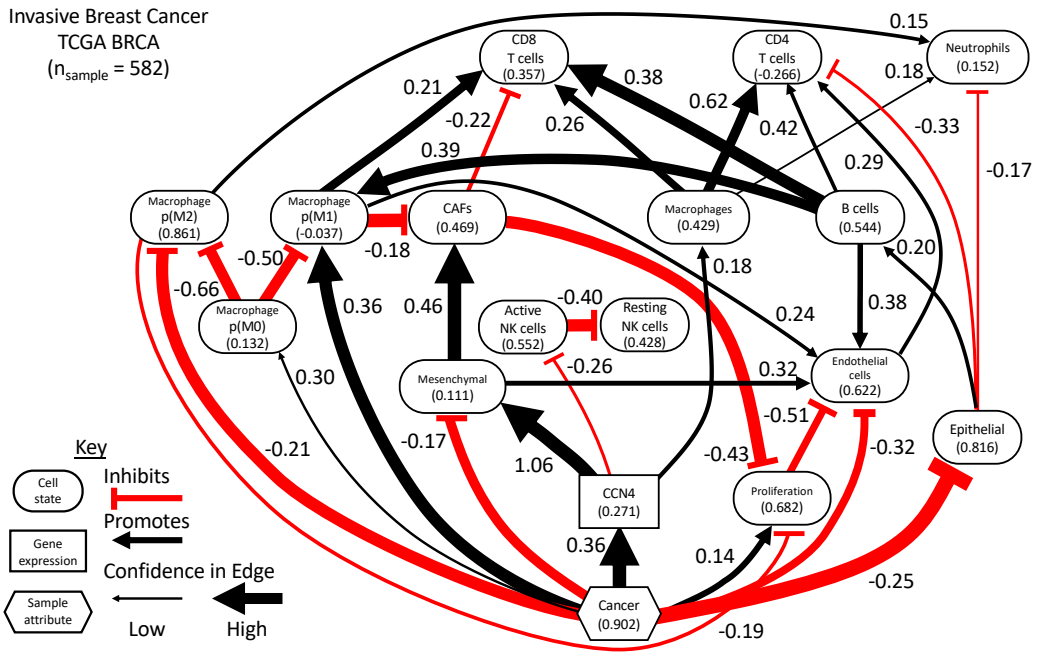


Fig. 3. A directed acyclic graph (DAG) representing the conditional probability distribution inferred using the digital cytometry and sample features extracted from the breast cancer arm of the TCGA.

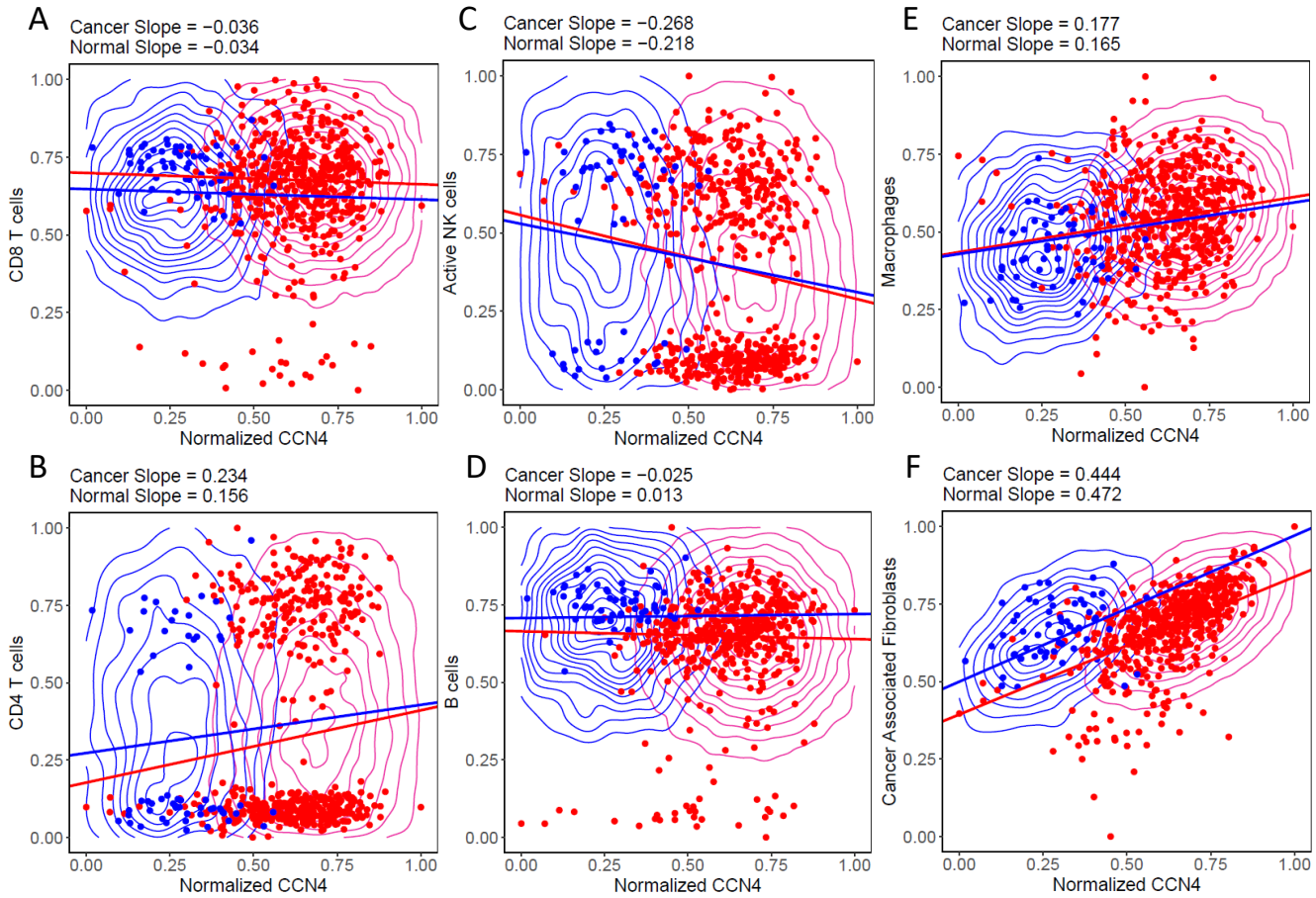


Fig. 4. Conditional probability query of the BRCA DAG compared against digital cytometry estimates obtained from experimental data.

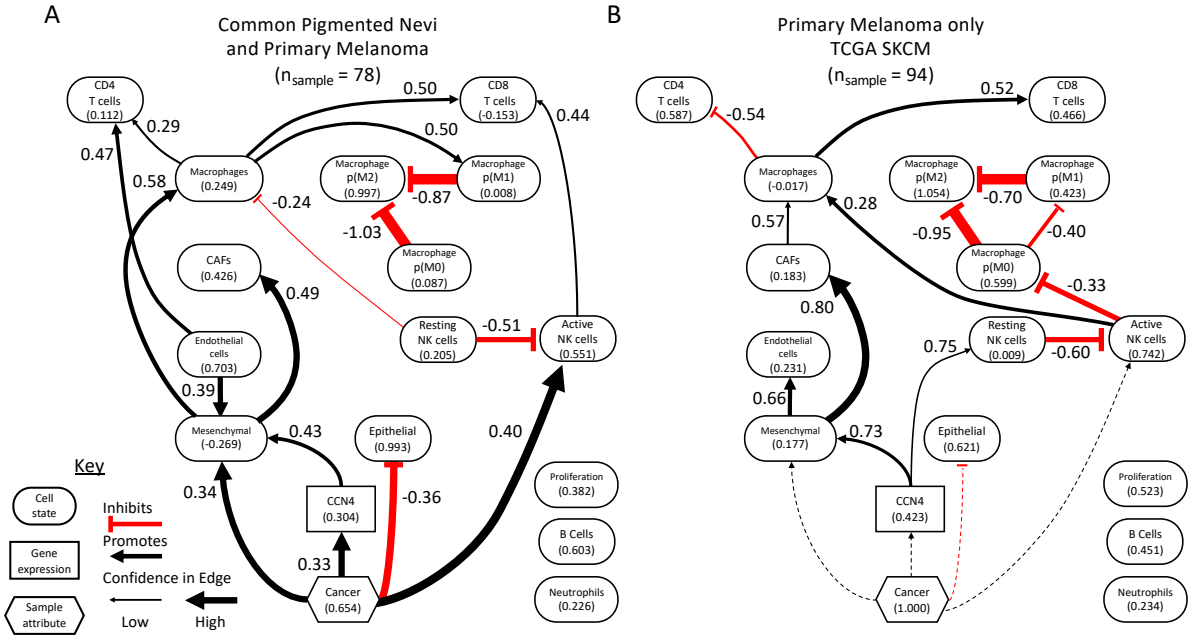


Fig. 5. Two DAGs representing the conditional probability distributions inferred using the digital cytometry and sample features extracted from the two melanoma-related datasets.

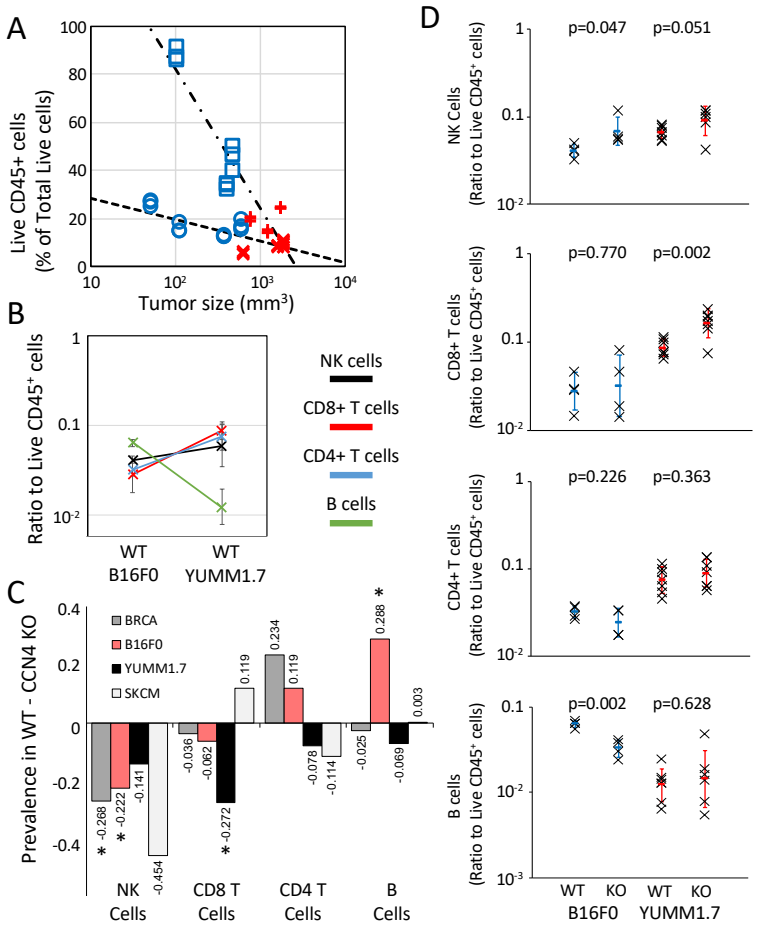


Fig. 6. CCN4 knock-out in two syngeneic mouse models of melanoma induced a similar shift in cytolytic lymphocytes as observed in human breast cancer and melanoma.

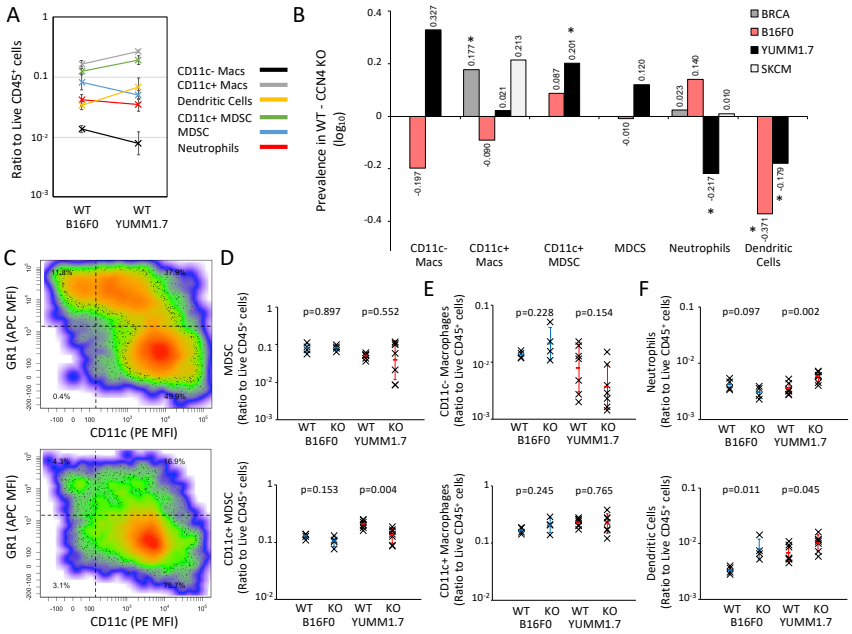


Fig. 7. Myeloid immune cell subsets differentially infiltrated tumors derived from WT B16F0 and YUMM1.7 cells and shifted in prevalence upon CCN4 knock-out.

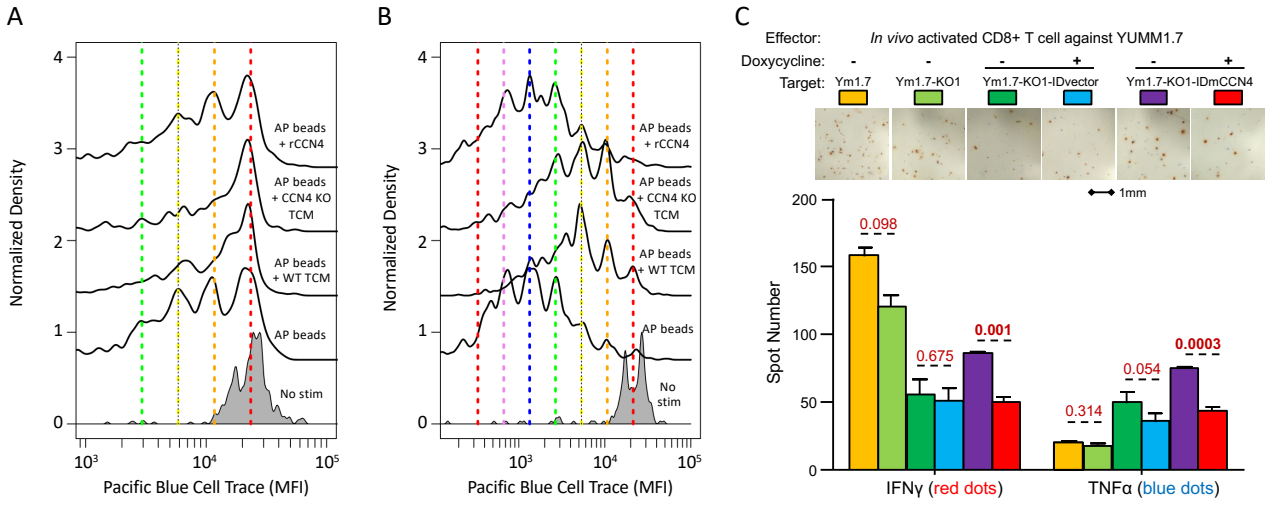


Fig. 8. CCN4 had no direct effect on T cell proliferation but impaired CD8$^+$ T cell function.

Table S1. List of edges and corresponding arcs, whether an arc was predicted to promote (+) or inhibit (-) the target node, and the strength inferred using the different structure learning algorithms in analyzing the features present in TCGA breast cancer dataset. Rows highlighted in green indicate arcs that were included in the consensus seed network, yellow indicate that the directionality was unclear, and red indicate arcs included in the “blacklist". The edge numbers correspond to the x-axis in Figures 2 and S1.

Edge No	from	to	mmpc CorSign	hiton CorSign	pc_stable strength	pcstable CorSign	gs strength	gs CorSign	iamb strength	iamb CorSign	iambfdr strength	iambfdr CorSign	tabu strength	tabu CorSign	hc strength	hc CorSign	mmhc strength	mmhc CorSign	rsmax2 strength	rsmax2 CorSign	Hit Count	Min strength	Max strength
1	CCN4	Mesenchymal	+	+	1.79E-139	+	2.02E-138	+	2.02E-138	+	2.02E-138	+	1.95E-148	+	1.95E-148	+	1.79E-139	+	1.79E-139	+	10	1.95E-148	2.02E-138
1	Mesenchymal	CCN4	+	+																			
2	pM0	pM1	-	-									3.77E-48	-	3.77E-48	-					10	0.00E+00	3.77E-48
2	pM1	pM0	-	-	0.00E+00	-	0.00E+00	-	0.00E+00	-	0.00E+00	-					0.00E+00	-	0.00E+00	-			
3	pM0	pM2	-	-									0.00E+00	-	0.00E+00	-					10	0.00E+00	0.00E+00
3	pM2	pM0	-	-	0.00E+00	-	0.00E+00	-	0.00E+00	-	0.00E+00	-					0.00E+00	-	0.00E+00	-			
4	CAF_lg	Mesenchymal	+	+																	10	1.86E-108	9.62E-73
4	Mesenchymal	CAF_lg	+	+	8.08E-86	+	9.62E-73	+	9.62E-73	+	9.62E-73	+	2.94E-97	+	1.86E-108	+	9.56E-73	+	2.35E-86	+			
5	Cancer	CCN4	+	+	2.06E-55	+	2.06E-55	+	2.06E-55	+	2.06E-55	+	2.06E-55	+	2.06E-55	+	2.06E-55	+	2.06E-55	+	10	2.06E-55	2.06E-55
5	CCN4	Cancer	+	+																			
6	NK.cells.active_lg	NK.cells.rest_lg	-	-	4.80E-30	-	4.80E-30	-	4.80E-30	-	4.80E-30	-	1.15E-33	-	1.15E-33	-	4.80E-30	-	4.80E-30	-	10	1.15E-33	4.80E-30
6	NK.cells.rest_lg	NK.cells.active_lg	-	-																			
7	B.cells.naive_lg	T.cells.CD8_lg	+	+	2.92E-27	+	9.25E-27	+	9.25E-27	+	2.92E-27	+	3.10E-16	+	3.10E-16	+	2.92E-27	+	2.92E-27	+	10	2.92E-27	3.10E-16
7	T.cells.CD8_lg	B.cells.naive_lg	+	+																			
8	CAF_lg	proliferation	-	-											9.87E-15	-	7.61E-26	-	7.61E-26	-	10	7.61E-26	3.37E-12
8	proliferation	CAF_lg	-	-	1.80E-18	-	6.37E-21	-	6.37E-21	-	6.37E-21	-	3.37E-12	-									
9	Endothelial.cells_lg	proliferation	-	-									5.75E-10	-	7.36E-09	-					10	9.05E-26	7.36E-09
9	proliferation	Endothelial.cells_lg	-	-	6.23E-18	-	9.05E-26	-	2.88E-15	-	9.05E-26	-					2.88E-15	-	2.88E-15	-			
10	Cancer	Epithelial	-	-	1.66E-25	-	1.66E-25	-	1.66E-25	-	1.66E-25	-	1.66E-25	-	1.66E-25	-	1.66E-25	-	1.66E-25	-	10	1.66E-25	1.66E-25
10	Epithelial	Cancer	-	-																			
11	Cancer	proliferation	+	+	1.93E-08	+	6.68E-23	+	6.68E-23	+	6.68E-23	+	4.70E-12	+	1.09E-07	+	3.72E-10	+	3.72E-10	+	10	6.68E-23	1.09E-07
11	proliferation	Cancer	+	+																			
12	CD4Tcell_sc_lg	Macrophages_sc_lg	+	+																	10	1.91E-16	2.01E-13
12	Macrophages_sc_lg	CD4Tcell_sc_lg	+	+	2.01E-13	+	1.91E-16	+	1.91E-16	+	5.09E-16	+	3.07E-14	+	3.07E-14	+	5.09E-16	+	5.09E-16	+			
13	pM1	T.cells.CD8_lg	+	+	3.53E-14	+	8.18E-13	+	8.18E-13	+	3.53E-14	+	8.98E-15	+	8.98E-15	+	3.53E-14	+	3.53E-14	+	10	8.98E-15	8.18E-13
13	T.cells.CD8_lg	pM1	+	+																			
14	Macrophages_sc_lg	T.cells.CD8_lg	+	+	2.38E-12	+	9.15E-13	+	9.15E-13	+	2.38E-12	+	6.56E-12	+	6.56E-12	+	2.38E-12	+	2.38E-12	+	10	9.15E-13	6.56E-12
14	T.cells.CD8_lg	Macrophages_sc_lg	+	+																			
15	CD4Tcell_sc_lg	Endothelial.cells_lg	+	+																	10	3.64E-10	7.84E-07
15	Endothelial.cells_lg	CD4Tcell_sc_lg	+	+	7.84E-07	+	5.65E-08	+	5.65E-08	+	2.93E-07	+	3.64E-10	+	3.64E-10	+	2.93E-07	+	2.93E-07	+			
16	CAF_lg	T.cells.CD8_lg	-	-	3.06E-07	-	2.84E-07	-	2.84E-07	-	3.06E-07	-	3.42E-09	-	3.42E-09	-	3.06E-07	-	3.06E-07	-	10	3.42E-09	3.06E-07
16	T.cells.CD8_lg	CAF_lg	-	-																			
17	B.cells.naive_lg	CD4Tcell_sc_lg	+	+	7.31E-05	+	2.95E-06	+	2.95E-06	+	1.95E-06	+	2.17E-07	+	2.17E-07	+	1.95E-06	+	1.95E-06	+	10	2.17E-07	7.31E-05
17	CD4Tcell_sc_lg	B.cells.naive_lg	+	+																			
18	Epithelial	Neutrophils_lg	-	-	6.79E-04	-	2.82E-03	-	2.82E-03	-			1.62E-03	-	1.62E-03	-	6.79E-04	-	6.79E-04	-	9	6.79E-04	2.82E-03
18	Neutrophils_lg	Epithelial	-	-																			
19	Cancer	Endothelial.cells_lg	-	-	2.10E-05	-			7.40E-09	-			3.67E-38	-	3.67E-38	-	7.40E-09	-	7.40E-09	-	8	3.67E-38	2.10E-05
19	Endothelial.cells_lg	Cancer	-	-																			
20	B.cells.naive_lg	Endothelial.cells_lg	+	+	4.01E-20	+															8	2.71E-22	1.52E-18
20	Endothelial.cells_lg	B.cells.naive_lg	+	+			2.71E-22	+	2.71E-22	+	2.71E-22	+					1.52E-18	+	2.71E-22	+			
21	Cancer	pM2	-	-	1.91E-42	-							7.00E-04	-	7.00E-04	-	1.91E-42	-	1.91E-42	-	7	1.91E-42	7.00E-04
21	pM2	Cancer	-	-																			
22	Cancer	Mesenchymal	+	+	1.16E-16	+							8.89E-17	+	8.89E-17	+	1.16E-16	+	1.16E-16	+	7	8.89E-17	1.16E-16
22	Mesenchymal	Cancer	+	+																			
23	pM2	proliferation	-	-	3.30E-10	-							1.02E-10	-	6.65E-10	-	2.09E-11	-	2.09E-11	-	7	2.09E-11	6.65E-10
23	proliferation	pM2	-	-																			
24	Neutrophils_lg	pM2	+	+																	7	7.51E-09	1.57E-04
24	pM2	Neutrophils_lg	+	+	1.57E-04	+							7.51E-09	+	7.51E-09	+	1.57E-04	+	1.57E-04	+			
25	Macrophages_sc_lg	Neutrophils_lg	+	+	1.86E-03	+	1.59E-03	+	1.59E-03	+							1.86E-03	+	1.86E-03	+	7	1.59E-03	1.86E-03
25	Neutrophils_lg	Macrophages_sc_lg	+	+																			
26	CAF_lg	pM1	-	-																	6	3.74E-24	4.37E-12
26	pM1	CAF_lg	-	-	1.06E-21	-							4.37E-12	-	2.49E-13	-			3.74E-24	-			
27	pM1	pM2											0.00E+00	-	0.00E+00	-					5	0.00E+00	3.43E-04
27	pM2	pM1					3.43E-04	-	3.43E-04	-	3.43E-04	-											
28	NK.cells.rest_lg	T.cells.CD8_lg					1.24E-02	-	1.24E-02	-			7.88E-04	-	7.88E-04	-					4	7.88E-04	1.24E-02
28	T.cells.CD8_lg	NK.cells.rest_lg																					
29	CD4Tcell_sc_lg	NK.cells.active_lg																			4	1.08E-03	1.26E-03
29	NK.cells.active_lg	CD4Tcell_sc_lg					1.26E-03	-	1.26E-03	-			1.08E-03	-	1.08E-03	-							
30	CD4Tcell_sc_lg	pM1	+	+																	3	1.23E-02	1.23E-02
30	pM1	CD4Tcell_sc_lg	+	+	1.23E-02	+																	
31	Cancer	pM1											2.35E-35	+	2.35E-35	+					2	2.35E-35	2.35E-35
31	pM1	Cancer																					
32	Cancer	pM0											5.03E-25	+	5.03E-25	+					2	5.03E-25	5.03E-25
32	pM0	Cancer																					
33	Endothelial.cells_lg	Mesenchymal																			2	4.02E-19	4.02E-19
33	Mesenchymal	Endothelial.cells_lg											4.02E-19	+	4.02E-19	+							
34	B.cells.naive_lg	pM1											4.18E-19	+	4.18E-19	+					2	4.18E-19	4.18E-19
34	pM1	B.cells.naive_lg																					
35	CAF_lg	Macrophages_sc_lg																			2	5.33E-17	1.86E-15
35	Macrophages_sc_lg	CAF_lg											5.33E-17	-	1.86E-15	-							
36	B.cells.naive_lg	Cancer																			2	3.36E-13	4.79E-07
36	Cancer	B.cells.naive_lg											3.36E-13	-	4.79E-07	-							
37	CAF_lg	pM2																			2	8.24E-13	3.71E-04
37	pM2	CAF_lg											3.71E-04	+	8.24E-13	+							
38	Mesenchymal	pM0																			2	1.51E-12	1.51E-12
38	pM0	Mesenchymal											1.51E-12	+	1.51E-12	+							
39	Endothelial.cells_lg	pM1																			2	4.23E-12	4.23E-12
39	pM1	Endothelial.cells_lg											4.23E-12	+	4.23E-12	+							
40	CCN4	proliferation											7.29E-12	-	1.51E-01	-					2	7.29E-12	1.51E-01
40	proliferation	CCN4																					
41	Cancer	T.cells.CD8_lg											2.30E-10	-	2.30E-10	-					2	2.30E-10	2.30E-10
41	T.cells.CD8_lg	Cancer																					
42	Macrophages_sc_lg	pM0											2.14E-08	-	2.14E-08	-					2	2.14E-08	2.14E-08
42	pM0	Macrophages_sc_lg																					
43	NK.cells.rest_lg	pM0											2.12E-07	+	2.12E-07	+					2	2.12E-07	2.12E-07
43	pM0	NK.cells.rest_lg																					
44	Mesenchymal	Neutrophils_lg											2.39E-07	+	2.39E-07	+					2	2.39E-07	2.39E-07
44	Neutrophils_lg	Mesenchymal																					
45	CCN4	Macrophages_sc_lg											4.84E-07	+	4.84E-07	+					2	4.84E-07	4.84E-07
45	Macrophages_sc_lg	CCN4																					
46	Cancer	NK.cells.rest_lg											4.25E-06	-	4.25E-06	-					2	4.25E-06	4.25E-06
46	NK.cells.rest_lg	Cancer																					
47	CD4Tcell_sc_lg	Epithelial																			2	9.98E-06	9.98E-06
47	Epithelial	CD4Tcell_sc_lg											9.98E-06	-	9.98E-06	-							
48	B.cells.naive_lg	NK.cells.rest_lg											3.72E-05	-	3.72E-05	-					2	3.72E-05	3.72E-05
48	NK.cells.rest_lg	B.cells.naive_lg																					
49	Epithelial	Mesenchymal											7.35E-05	+	7.35E-05	+					2	7.35E-05	7.35E-05
49	Mesenchymal	Epithelial																					
50	Endothelial.cells_lg	Neutrophils_lg											7.61E-05	-	7.61E-05	-					2	7.61E-05	7.61E-05
50	Neutrophils_lg	Endothelial.cells_lg																					
51	Macrophages_sc_lg	proliferation											1.53E-01	-	1.18E-04	-					2	1.18E-04	1.53E-01
51	proliferation	Macrophages_sc_lg																					
52	Mesenchymal	T.cells.CD8_lg											2.04E-04	+	2.04E-04	+					2	2.04E-04	2.04E-04
52	T.cells.CD8_lg	Mesenchymal																					
53	CCN4	NK.cells.active_lg											2.64E-04	-	2.64E-04	-					2	2.64E-04	2.64E-04
53	NK.cells.active_lg	CCN4																					
54	CAF_lg	Epithelial																			2	5.82E-04	2.94E-03
54	Epithelial	CAF_lg											2.94E-03	+	5.82E-04	+							
55	B.cells.naive_lg	CAF_lg											1.48E-03	-	6.27E-03	-					2	1.48E-03	6.27E-03
55	CAF_lg	B.cells.naive_lg																					
56	NK.cells.rest_lg	proliferation											2.80E-02	+	2.15E-03	+					2	2.15E-03	2.80E-02
56	proliferation	NK.cells.rest_lg																					
57	B.cells.naive_lg	Epithelial	+																		2	8.19E-03	8.19E-03
57	Epithelial	B.cells.naive_lg	+														8.19E-03	+					
58	B.cells.naive_lg	CCN4																			1	5.87E-08	5.87E-08
58	CCN4	B.cells.naive_lg											5.87E-08	+									
59	Cancer	Macrophages_sc_lg																			0	1.00E+00	1.00E+00
59	Macrophages_sc_lg	Cancer																					
60	CAF_lg	Endothelial.cells_lg																			0	1.00E+00	1.00E+00
60	Endothelial.cells_lg	CAF_lg																					
61	CCN4	pM0																			0	1.00E+00	1.00E+00
61	pM0	CCN4																					
62	Endothelial.cells_lg	T.cells.CD8_lg																			0	1.00E+00	1.00E+00
62	T.cells.CD8_lg	Endothelial.cells_lg																					
63	Macrophages_sc_lg	Mesenchymal																			0	1.00E+00	1.00E+00
63	Mesenchymal	Macrophages_sc_lg																					
64	B.cells.naive_lg	Macrophages_sc_lg																			0	1.00E+00	1.00E+00
64	Macrophages_sc_lg	B.cells.naive_lg																					
65	Mesenchymal	pM2																			0	1.00E+00	1.00E+00
65	pM2	Mesenchymal																					
66	NK.cells.active_lg	pM0																			0	1.00E+00	1.00E+00
66	pM0	NK.cells.active_lg																					
67	B.cells.naive_lg	Mesenchymal																			0	1.00E+00	1.00E+00
67	Mesenchymal	B.cells.naive_lg																					
68	CCN4	Endothelial.cells_lg																			0	1.00E+00	1.00E+00
68	Endothelial.cells_lg	CCN4																					
69	CAF_lg	Cancer																			0	1.00E+00	1.00E+00
69	Cancer	CAF_lg																					
70	Neutrophils_lg	pM0																			0	1.00E+00	1.00E+00
70	pM0	Neutrophils_lg																					
71	CCN4	Neutrophils_lg																			0	1.00E+00	1.00E+00
71	Neutrophils_lg	CCN4																					
72	Neutrophils_lg	pM1																			0	1.00E+00	1.00E+00
72	pM1	Neutrophils_lg																					
73	Macrophages_sc_lg	NK.cells.rest_lg																			0	1.00E+00	1.00E+00
73	NK.cells.rest_lg	Macrophages_sc_lg																					
74	Epithelial	pM2																			0	1.00E+00	1.00E+00
74	pM2	Epithelial																					
75	Endothelial.cells_lg	pM0																			0	1.00E+00	1.00E+00
75	pM0	Endothelial.cells_lg																					
76	CCN4	NK.cells.rest_lg																			0	1.00E+00	1.00E+00
76	NK.cells.rest_lg	CCN4																					
77	NK.cells.active_lg	T.cells.CD8_lg																			0	1.00E+00	1.00E+00
77	T.cells.CD8_lg	NK.cells.active_lg																					
78	CAF_lg	CD4Tcell_sc_lg																			0	1.00E+00	1.00E+00
78	CD4Tcell_sc_lg	CAF_lg																					
79	CAF_lg	CCN4																			0	1.00E+00	1.00E+00
79	CCN4	CAF_lg																					
80	Endothelial.cells_lg	Epithelial																			0	1.00E+00	1.00E+00
80	Epithelial	Endothelial.cells_lg																					
81	Macrophages_sc_lg	NK.cells.active_lg																			0	1.00E+00	1.00E+00
81	NK.cells.active_lg	Macrophages_sc_lg																					
82	Macrophages_sc_lg	pM1																			0	1.00E+00	1.00E+00
82	pM1	Macrophages_sc_lg																					
83	Cancer	NK.cells.active_lg																			0	1.00E+00	1.00E+00
83	NK.cells.active_lg	Cancer																					
84	Epithelial	pM0																			0	1.00E+00	1.00E+00
84	pM0	Epithelial																					

Table S2. List of edges and corresponding arcs, whether an arc was predicted to promote (+) or inhibit (-) the target node, and the strength inferred using the different structure learning algorithms in analyzing the features present in dataset containing common melanocytic nevi and primary melanoma tissue samples (GEO). Rows highlighted in green indicate arcs that were included in the consensus seed network, yellow indicate that the directionality was unclear, and red indicate arcs included in the "blacklist". The edge numbers correspond to the x-axis in Figure S2, panel A.

Edge No	from	to	mmpc CorSign	hiton CorSign	pc_stable strength	pcstable CorSign	gs strength	gs CorSign	iamb strength	iamb CorSign	iambfdr strength	iambfdr CorSign	tabu strength	tabu CorSign	hc strength	hc CorSign	mmhc strength	mmhc CorSign	rsmax2 strength	rsmax2 CorSign	Hit Count	Min strength	Max strength
1	CCN4	Mesenchymal	+	+	2.29E-02	+	6.48E-05	+	6.48E-05	+			8.38E-07	+	8.38E-07	+	6.48E-05	+	6.48E-05	+	9	8.38E-07	2.29E-02
1	Mesenchymal	CCN4	+	+																			
2	pM0	pM1											2.28E-02	-	2.28E-02	-					2	2.28E-02	2.28E-02
2	pM1	pM0																					
3	pM0	pM2	-	-	3.31E-38	-	3.31E-38	-	3.31E-38	-	3.31E-38	-	3.31E-38	-	3.31E-38	-	3.31E-38	-	3.31E-38	-	10	3.31E-38	3.31E-38
3	pM2	pM0	-	-																			
4	CAF_lg	Mesenchymal	+	+	2.20E-12	+					6.70E-15	+									10	6.70E-15	5.98E-05
4	Mesenchymal	CAF_lg	+	+			2.75E-14	+	2.75E-14	+			7.46E-12	+	5.98E-05	+	2.75E-14	+	2.75E-14	+			
5	Cancer	CCN4											5.06E-11	+	5.06E-11	+					2	5.06E-11	5.06E-11
5	CCN4	Cancer																					
6	NK.cells.active_lg	NK.cells.rest_lg	-	-													1.02E-06	-			10	1.01E-09	1.68E-05
6	NK.cells.rest_lg	NK.cells.active_lg	-	-	1.68E-05	-	1.01E-09	-	1.01E-09	-	1.68E-05	-	1.01E-09	-	1.01E-09	-			1.68E-05	-			
7	B.cells.ive_lg	T.cells.CD8_lg																			0	1.00E+00	1.00E+00
7	T.cells.CD8_lg	B.cells.ive_lg																					
8	CAF_lg	proliferation																			0	1.00E+00	1.00E+00
8	proliferation	CAF_lg																					
9	Endothelial.cells_lg	proliferation											5.19E-05	-	5.19E-05	-					2	5.19E-05	5.19E-05
9	proliferation	Endothelial.cells_lg																					
10	Cancer	Epithelial	-	-	4.24E-15	-	4.24E-15	-	4.24E-15	-	4.24E-15	-	4.24E-15	-	2.79E-16	-	4.24E-15	-	4.24E-15	-	10	2.79E-16	4.24E-15
10	Epithelial	Cancer	-	-																			
11	Cancer	proliferation					1.47E-09	+	1.47E-09	+			6.39E-11	+	6.39E-11	+					4	6.39E-11	1.47E-09
11	proliferation	Cancer																					
12	CD4Tcell_sc_lg	Macrophages_sc_lg	+	+																	9	4.58E-04	4.58E-04
12	Macrophages_sc_lg	CD4Tcell_sc_lg	+	+	4.58E-04	+	4.58E-04	+	4.58E-04	+			4.58E-04	+	4.58E-04	+	4.58E-04	+	4.58E-04	+			
13	pM1	T.cells.CD8_lg	+	+	3.70E-03	+							5.96E-03	+	5.96E-03	+	3.70E-03	+	3.70E-03	+	7	3.70E-03	5.96E-03
13	T.cells.CD8_lg	pM1	+	+																			
14	Macrophages_sc_lg	T.cells.CD8_lg	+	+	1.79E-04	+	9.32E-06	+	9.32E-06	+	1.82E-11	+	7.68E-06	+	7.68E-06	+	1.79E-04	+	1.79E-04	+	10	1.82E-11	1.79E-04
14	T.cells.CD8_lg	Macrophages_sc_lg	+	+																			
15	CD4Tcell_sc_lg	Endothelial.cells_lg	+	+																	10	8.60E-08	4.15E-07
15	Endothelial.cells_lg	CD4Tcell_sc_lg	+	+	4.15E-07	+	4.15E-07	+	4.15E-07	+	8.60E-08	+	4.15E-07	+	4.15E-07	+	4.15E-07	+	4.15E-07	+			
16	CAF_lg	T.cells.CD8_lg											8.87E-03	+	8.87E-03	+					2	8.87E-03	8.87E-03
16	T.cells.CD8_lg	CAF_lg																					
17	B.cells.ive_lg	CD4Tcell_sc_lg																			0	1.00E+00	1.00E+00
17	CD4Tcell_sc_lg	B.cells.ive_lg																					
18	Epithelial	Neutrophils_lg																			0	1.00E+00	1.00E+00
18	Neutrophils_lg	Epithelial																					
19	Cancer	Endothelial.cells_lg																			0	1.00E+00	1.00E+00
19	Endothelial.cells_lg	Cancer																					
20	B.cells.ive_lg	Endothelial.cells_lg																			2	3.41E-03	3.41E-03
20	Endothelial.cells_lg	B.cells.ive_lg											3.41E-03	+	3.41E-03	+							
21	Cancer	pM2																			0	1.00E+00	1.00E+00
21	pM2	Cancer																					
22	Cancer	Mesenchymal	+	+	2.46E-12	+	6.29E-09	+	6.29E-09	+	2.09E-18	+	8.02E-13	+	8.02E-13	+	6.29E-09	+	6.29E-09	+	10	2.09E-18	6.29E-09
22	Mesenchymal	Cancer	+	+																			
23	pM2	proliferation																			0	1.00E+00	1.00E+00
23	proliferation	pM2																					
24	Neutrophils_lg	pM2																			0	1.00E+00	1.00E+00
24	pM2	Neutrophils_lg																					
25	Macrophages_sc_lg	Neutrophils_lg																			0	1.00E+00	1.00E+00
25	Neutrophils_lg	Macrophages_sc_lg																					
26	CAF_lg	pM1																			0	1.00E+00	1.00E+00
26	pM1	CAF_lg																					
27	pM1	pM2	-	-	2.40E-41	-	2.40E-41	-	2.40E-41	-	2.40E-41	-	2.40E-41	-	2.40E-41	-	2.40E-41	-	2.40E-41	-	10	2.40E-41	2.40E-41
27	pM2	pM1	-	-																			
28	NK.cells.rest_lg	T.cells.CD8_lg																			0	1.00E+00	1.00E+00
28	T.cells.CD8_lg	NK.cells.rest_lg																					
29	CD4Tcell_sc_lg	NK.cells.active_lg																			0	1.00E+00	1.00E+00
29	NK.cells.active_lg	CD4Tcell_sc_lg																					
30	CD4Tcell_sc_lg	pM1																			0	1.00E+00	1.00E+00
30	pM1	CD4Tcell_sc_lg																					
31	Cancer	pM1																			0	1.00E+00	1.00E+00
31	pM1	Cancer																					
32	Cancer	pM0																			0	1.00E+00	1.00E+00
32	pM0	Cancer																					
33	Endothelial.cells_lg	Mesenchymal											2.75E-10	+	2.75E-10	+					2	2.75E-10	2.75E-10
33	Mesenchymal	Endothelial.cells_lg																					
34	B.cells.ive_lg	pM1																			0	1.00E+00	1.00E+00
34	pM1	B.cells.ive_lg																					
35	CAF_lg	Macrophages_sc_lg																			0	1.00E+00	1.00E+00
35	Macrophages_sc_lg	CAF_lg																					
36	B.cells.ive_lg	Cancer																			0	1.00E+00	1.00E+00
36	Cancer	B.cells.ive_lg																					
37	CAF_lg	pM2																			0	1.00E+00	1.00E+00
37	pM2	CAF_lg																					
38	Mesenchymal	pM0																			0	1.00E+00	1.00E+00
38	pM0	Mesenchymal																					
39	Endothelial.cells_lg	pM1																			0	1.00E+00	1.00E+00
39	pM1	Endothelial.cells_lg																					
40	CCN4	proliferation																			0	1.00E+00	1.00E+00
40	proliferation	CCN4																					
41	Cancer	T.cells.CD8_lg																			0	1.00E+00	1.00E+00
41	T.cells.CD8_lg	Cancer																					
42	Macrophages_sc_lg	pM0																			0	1.00E+00	1.00E+00
42	pM0	Macrophages_sc_lg																					
43	NK.cells.rest_lg	pM0	+	+	1.17E-03	+							2.22E-04	+	2.22E-04	+					7	6.30E-05	1.17E-03
43	pM0	NK.cells.rest_lg	+	+													6.30E-05	+	1.17E-03	+			
44	Mesenchymal	Neutrophils_lg																			0	1.00E+00	1.00E+00
44	Neutrophils_lg	Mesenchymal																					
45	CCN4	Macrophages_sc_lg																			0	1.00E+00	1.00E+00
45	Macrophages_sc_lg	CCN4																					
46	Cancer	NK.cells.rest_lg																			0	1.00E+00	1.00E+00
46	NK.cells.rest_lg	Cancer																					
47	CD4Tcell_sc_lg	Epithelial																			0	1.00E+00	1.00E+00
47	Epithelial	CD4Tcell_sc_lg																					
48	B.cells.ive_lg	NK.cells.rest_lg																			0	1.00E+00	1.00E+00
48	NK.cells.rest_lg	B.cells.ive_lg																					
49	Epithelial	Mesenchymal																			0	1.00E+00	1.00E+00
49	Mesenchymal	Epithelial																					
50	Endothelial.cells_lg	Neutrophils_lg																			0	1.00E+00	1.00E+00
50	Neutrophils_lg	Endothelial.cells_lg																					
51	Macrophages_sc_lg	proliferation																			0	1.00E+00	1.00E+00
51	proliferation	Macrophages_sc_lg																					
52	Mesenchymal	T.cells.CD8_lg																			0	1.00E+00	1.00E+00
52	T.cells.CD8_lg	Mesenchymal																					
53	CCN4	NK.cells.active_lg																			0	1.00E+00	1.00E+00
53	NK.cells.active_lg	CCN4																					
54	CAF_lg	Epithelial																			0	1.00E+00	1.00E+00
54	Epithelial	CAF_lg																					
55	B.cells.ive_lg	CAF_lg																			0	1.00E+00	1.00E+00
55	CAF_lg	B.cells.ive_lg																					
56	NK.cells.rest_lg	proliferation																			0	1.00E+00	1.00E+00
56	proliferation	NK.cells.rest_lg																					
57	B.cells.ive_lg	Epithelial																			0	1.00E+00	1.00E+00
57	Epithelial	B.cells.ive_lg																					
58	B.cells.ive_lg	CCN4																			0	1.00E+00	1.00E+00
58	CCN4	B.cells.ive_lg																					
59	Cancer	Macrophages_sc_lg																			0	1.00E+00	1.00E+00
59	Macrophages_sc_lg	Cancer																					
60	CAF_lg	Endothelial.cells_lg			3.11E-06	+															3	3.92E-05	3.92E-05
60	Endothelial.cells_lg	CAF_lg											8.04E-04	+	4.19E-04	+							
61	CCN4	pM0																			0	1.00E+00	1.00E+00
61	pM0	CCN4																					
62	Endothelial.cells_lg	T.cells.CD8_lg																			0	1.00E+00	1.00E+00
62	T.cells.CD8_lg	Endothelial.cells_lg																					
63	Macrophages_sc_lg	Mesenchymal																			2	3.92E-05	3.92E-05
63	Mesenchymal	Macrophages_sc_lg											7.24E-09	+	7.24E-09	+							
64	B.cells.ive_lg	Macrophages_sc_lg																			0	1.00E+00	1.00E+00
64	Macrophages_sc_lg	B.cells.ive_lg																					
65	Mesenchymal	pM2																			0	1.00E+00	1.00E+00
65	pM2	Mesenchymal																					
66	NK.cells.active_lg	pM0																			0	1.00E+00	1.00E+00
66	pM0	NK.cells.active_lg																					
67	B.cells.ive_lg	Mesenchymal																			0	1.00E+00	1.00E+00
67	Mesenchymal	B.cells.ive_lg																					
68	CCN4	Endothelial.cells_lg																			0	1.00E+00	1.00E+00
68	Endothelial.cells_lg	CCN4																					
69	CAF_lg	Cancer																			0	1.00E+00	1.00E+00
69	Cancer	CAF_lg																					
70	Neutrophils_lg	pM0																			0	1.00E+00	1.00E+00
70	pM0	Neutrophils_lg																					
71	CCN4	Neutrophils_lg											7.22E-04	+	7.22E-04	+					2	3.92E-05	3.92E-05
71	Neutrophils_lg	CCN4																					
72	Neutrophils_lg	pM1																			0	1.00E+00	1.00E+00
72	pM1	Neutrophils_lg																					
73	Macrophages_sc_lg	NK.cells.rest_lg																			2	3.92E-05	3.92E-05
73	NK.cells.rest_lg	Macrophages_sc_lg											2.08E-02	-	2.08E-02	-							
74	Epithelial	pM2																			0	1.00E+00	1.00E+00
74	pM2	Epithelial																					
75	Endothelial.cells_lg	pM0		-																	3	3.92E-05	3.92E-05
75	pM0	Endothelial.cells_lg		-									1.27E-04	-					1.27E-04	-			
76	CCN4	NK.cells.rest_lg																			0	1.00E+00	1.00E+00
76	NK.cells.rest_lg	CCN4																					
77	NK.cells.active_lg	T.cells.CD8_lg	+	+	3.93E-03	+	1.05E-04	+	1.05E-04	+			1.19E-03	+	1.19E-03	+	3.93E-03	+	3.93E-03	+	9	1.05E-04	3.93E-03
77	T.cells.CD8_lg	NK.cells.active_lg	+	+																			
78	CAF_lg	CD4Tcell_sc_lg																			0	1.00E+00	1.00E+00
78	CD4Tcell_sc_lg	CAF_lg																					
79	CAF_lg	CCN4																			1	2.35E-01	2.35E-01
79	CCN4	CAF_lg													2.35E-01	+							
80	Endothelial.cells_lg	Epithelial													9.69E-04	+					1	9.69E-04	9.69E-04
80	Epithelial	Endothelial.cells_lg																					
81	Macrophages_sc_lg	NK.cells.active_lg																			0	1.00E+00	1.00E+00
81	NK.cells.active_lg	Macrophages_sc_lg																					
82	Macrophages_sc_lg	pM1											5.29E-07	+	5.29E-07	+					2	5.29E-07	5.29E-07
82	pM1	Macrophages_sc_lg																					
83	Cancer	NK.cells.active_lg					8.85E-16	+	8.85E-16	+			8.85E-16	+	8.85E-16	+					4	8.85E-16	8.85E-16
83	NK.cells.active_lg	Cancer																					
84	Epithelial	pM0											3.92E-05	-	3.92E-05	-					2	3.92E-05	3.92E-05
84	pM0	Epithelial																					

Table S3. List of edges and corresponding arcs, whether an arc was predicted to promote (+) or inhibit (-) the target node, and the strength inferred using the different structure learning algorithms in analyzing the features present in primary melanoma tissue samples in the TCGA SKCM dataset. Rows highlighted in green indicate arcs that were included in the consensus seed network, yellow indicate that the directionality was unclear, and red indicate arcs included in the “blacklist". The edge numbers correspond to the x-axis in Figure S2, panel B.

Edge No	from	to	mmpc CorSign	hiton CorSign	pc_stable strength	pcstable CorSign	gs strength	gs CorSign	iamb strength	iamb CorSign	iambfdr strength	iambfdr CorSign	tabu strength	tabu CorSign	hc strength	hc CorSign	mmhc strength	mmhc CorSign	rsmax2 strength	rsmax2 CorSign	Hit Count	Min strength	Max strength
1	CCN4	Mesenchymal	+	+	3.83E-06	+			1.14E-06	+											4	1.14E-06	3.83E-06
1	Mesenchymal	CCN4	+	+																			
2	pM0	pM1	-	-											9.96E-07	-	9.96E-07	-	9.96E-07	-	9	1.59E-34	9.96E-07
2	pM1	pM0	-	-	1.59E-34	-	1.59E-34	-	1.59E-34	-	1.59E-34	-											
3	pM0	pM2	-	-											6.36E-46	-	2.04E-18	-	2.04E-18	-	9	6.36E-46	2.04E-18
3	pM2	pM0	-	-	6.36E-46	-	6.36E-46	-	6.36E-46	-	6.36E-46	-											
4	CAF_lg	Mesenchymal	+	+											3.26E-18	+					9	3.26E-18	3.26E-18
4	Mesenchymal	CAF_lg	+	+	3.26E-18	+	3.26E-18	+	3.26E-18	+	3.26E-18	+					3.26E-18	+	3.26E-18	+			
5	Cancer	CCN4																			0	1.00E+00	1.00E+00
5	CCN4	Cancer																					
6	NK.cells.active_lg	NK.cells.rest_lg	-	-	2.37E-09	-									2.37E-09	-					9	2.01E-10	2.37E-09
6	NK.cells.rest_lg	NK.cells.active_lg	-	-			2.01E-10	-	2.01E-10	-	2.01E-10	-					2.01E-10	-	2.01E-10	-			
7	B.cells.ive_lg	T.cells.CD8_lg																			0	1.00E+00	1.00E+00
7	T.cells.CD8_lg	B.cells.ive_lg																					
8	CAF_lg	proliferation																			0	1.00E+00	1.00E+00
8	proliferation	CAF_lg																					
9	Endothelial.cells_lg	proliferation																			0	1.00E+00	1.00E+00
9	proliferation	Endothelial.cells_lg																					
10	Cancer	Epithelial																			0	1.00E+00	1.00E+00
10	Epithelial	Cancer																					
11	Cancer	proliferation																			0	1.00E+00	1.00E+00
11	proliferation	Cancer																					
12	CD4Tcell_sc_lg	Macrophages_sc_lg	-	-																	8	1.96E-06	2.61E-04
12	Macrophages_sc_lg	CD4Tcell_sc_lg	-	-	2.61E-04	-	2.61E-04	-	2.61E-04	-					1.96E-06	-	2.61E-04	-	2.61E-04	-			
13	pM1	T.cells.CD8_lg																			0	1.00E+00	1.00E+00
13	T.cells.CD8_lg	pM1																					
14	Macrophages_sc_lg	T.cells.CD8_lg	+	+	2.11E-05	+	1.43E-07	+	1.43E-07	+	1.43E-07	+			1.23E-11	+	1.43E-07	+	2.11E-05	+	9	1.23E-11	2.11E-05
14	T.cells.CD8_lg	Macrophages_sc_lg	+	+																			
15	CD4Tcell_sc_lg	Endothelial.cells_lg																			0	1.00E+00	1.00E+00
15	Endothelial.cells_lg	CD4Tcell_sc_lg																					
16	CAF_lg	T.cells.CD8_lg													1.58E-05	-					1	1.58E-05	1.58E-05
16	T.cells.CD8_lg	CAF_lg																					
17	B.cells.ive_lg	CD4Tcell_sc_lg																			0	1.00E+00	1.00E+00
17	CD4Tcell_sc_lg	B.cells.ive_lg																					
18	Epithelial	Neutrophils_lg																			0	1.00E+00	1.00E+00
18	Neutrophils_lg	Epithelial																					
19	Cancer	Endothelial.cells_lg																			0	1.00E+00	1.00E+00
19	Endothelial.cells_lg	Cancer																					
20	B.cells.ive_lg	Endothelial.cells_lg																			0	1.00E+00	1.00E+00
20	Endothelial.cells_lg	B.cells.ive_lg																					
21	Cancer	pM2																			0	1.00E+00	1.00E+00
21	pM2	Cancer																					
22	Cancer	Mesenchymal																			0	1.00E+00	1.00E+00
22	Mesenchymal	Cancer																					
23	pM2	proliferation																			0	1.00E+00	1.00E+00
23	proliferation	pM2																					
24	Neutrophils_lg	pM2																			0	1.00E+00	1.00E+00
24	pM2	Neutrophils_lg																					
25	Macrophages_sc_lg	Neutrophils_lg																			0	1.00E+00	1.00E+00
25	Neutrophils_lg	Macrophages_sc_lg																					
26	CAF_lg	pM1																			0	1.00E+00	1.00E+00
26	pM1	CAF_lg																					
27	pM1	pM2													9.70E-30	-					1	9.70E-30	9.70E-30
27	pM2	pM1																					
28	NK.cells.rest_lg	T.cells.CD8_lg																			0	1.00E+00	1.00E+00
28	T.cells.CD8_lg	NK.cells.rest_lg																					
29	CD4Tcell_sc_lg	NK.cells.active_lg																			0	1.00E+00	1.00E+00
29	NK.cells.active_lg	CD4Tcell_sc_lg																					
30	CD4Tcell_sc_lg	pM1																			0	1.00E+00	1.00E+00
30	pM1	CD4Tcell_sc_lg																					
31	Cancer	pM1																			0	1.00E+00	1.00E+00
31	pM1	Cancer																					
32	Cancer	pM0																			0	1.00E+00	1.00E+00
32	pM0	Cancer																					
33	Endothelial.cells_lg	Mesenchymal	+	+	1.28E-07	+															8	3.73E-08	1.28E-07
33	Mesenchymal	Endothelial.cells_lg	+	+			3.73E-08	+	3.73E-08	+					3.73E-08	+	3.73E-08	+	3.73E-08	+			
34	B.cells.ive_lg	pM1																			0	1.00E+00	1.00E+00
34	pM1	B.cells.ive_lg																					
35	CAF_lg	Macrophages_sc_lg	+	+	1.45E-08	+	1.45E-08	+	1.45E-08	+	1.45E-08	+			1.02E-11	+	1.45E-08	+	1.45E-08	+	9	1.02E-11	1.45E-08
35	Macrophages_sc_lg	CAF_lg	+	+																			
36	B.cells.ive_lg	Cancer																			0	1.00E+00	1.00E+00
36	Cancer	B.cells.ive_lg																					
37	CAF_lg	pM2																			0	1.00E+00	1.00E+00
37	pM2	CAF_lg																					
38	Mesenchymal	pM0																			0	1.00E+00	1.00E+00
38	pM0	Mesenchymal																					
39	Endothelial.cells_lg	pM1																			0	1.00E+00	1.00E+00
39	pM1	Endothelial.cells_lg																					
40	CCN4	proliferation																			0	1.00E+00	1.00E+00
40	proliferation	CCN4																					
41	Cancer	T.cells.CD8_lg																			0	1.00E+00	1.00E+00
41	T.cells.CD8_lg	Cancer																					
42	Macrophages_sc_lg	pM0																			0	1.00E+00	1.00E+00
42	pM0	Macrophages_sc_lg																					
43	NK.cells.rest_lg	pM0																			0	1.00E+00	1.00E+00
43	pM0	NK.cells.rest_lg																					
44	Mesenchymal	Neutrophils_lg																			0	1.00E+00	1.00E+00
44	Neutrophils_lg	Mesenchymal																					
45	CCN4	Macrophages_sc_lg																			0	1.00E+00	1.00E+00
45	Macrophages_sc_lg	CCN4																					
46	Cancer	NK.cells.rest_lg																			0	1.00E+00	1.00E+00
46	NK.cells.rest_lg	Cancer																					
47	CD4Tcell_sc_lg	Epithelial																			1	1.36E-02	1.36E-02
47	Epithelial	CD4Tcell_sc_lg													1.36E-02	+							
48	B.cells.ive_lg	NK.cells.rest_lg																			0	1.00E+00	1.00E+00
48	NK.cells.rest_lg	B.cells.ive_lg																					
49	Epithelial	Mesenchymal																			0	1.00E+00	1.00E+00
49	Mesenchymal	Epithelial																					
50	Endothelial.cells_lg	Neutrophils_lg																			0	1.00E+00	1.00E+00
50	Neutrophils_lg	Endothelial.cells_lg																					
51	Macrophages_sc_lg	proliferation													1.72E-04	-					1	1.72E-04	1.72E-04
51	proliferation	Macrophages_sc_lg																					
52	Mesenchymal	T.cells.CD8_lg																			0	1.00E+00	1.00E+00
52	T.cells.CD8_lg	Mesenchymal																					
53	CCN4	NK.cells.active_lg																			0	1.00E+00	1.00E+00
53	NK.cells.active_lg	CCN4																					
54	CAF_lg	Epithelial																			0	1.00E+00	1.00E+00
54	Epithelial	CAF_lg																					
55	B.cells.ive_lg	CAF_lg																			0	1.00E+00	1.00E+00
55	CAF_lg	B.cells.ive_lg																					
56	NK.cells.rest_lg	proliferation																			0	1.00E+00	1.00E+00
56	proliferation	NK.cells.rest_lg																					
57	B.cells.ive_lg	Epithelial																			0	1.00E+00	1.00E+00
57	Epithelial	B.cells.ive_lg																					
58	B.cells.ive_lg	CCN4																			0	1.00E+00	1.00E+00
58	CCN4	B.cells.ive_lg																					
59	Cancer	Macrophages_sc_lg																			0	1.00E+00	1.00E+00
59	Macrophages_sc_lg	Cancer																					
60	CAF_lg	Endothelial.cells_lg																			0	1.00E+00	1.00E+00
60	Endothelial.cells_lg	CAF_lg																					
61	CCN4	pM0																			0	1.00E+00	1.00E+00
61	pM0	CCN4																					
62	Endothelial.cells_lg	T.cells.CD8_lg																			0	1.00E+00	1.00E+00
62	T.cells.CD8_lg	Endothelial.cells_lg																					
63	Macrophages_sc_lg	Mesenchymal																			0	1.00E+00	1.00E+00
63	Mesenchymal	Macrophages_sc_lg																					
64	B.cells.ive_lg	Macrophages_sc_lg																			0	1.00E+00	1.00E+00
64	Macrophages_sc_lg	B.cells.ive_lg																					
65	Mesenchymal	pM2																			0	1.00E+00	1.00E+00
65	pM2	Mesenchymal																					
66	NK.cells.active_lg	pM0													3.16E-05	-					1	3.16E-05	3.16E-05
66	pM0	NK.cells.active_lg																					
67	B.cells.ive_lg	Mesenchymal																			0	1.00E+00	1.00E+00
67	Mesenchymal	B.cells.ive_lg																					
68	CCN4	Endothelial.cells_lg																			0	1.00E+00	1.00E+00
68	Endothelial.cells_lg	CCN4																					
69	CAF_lg	Cancer																			0	1.00E+00	1.00E+00
69	Cancer	CAF_lg																					
70	Neutrophils_lg	pM0																			0	1.00E+00	1.00E+00
70	pM0	Neutrophils_lg																					
71	CCN4	Neutrophils_lg																			0	1.00E+00	1.00E+00
71	Neutrophils_lg	CCN4																					
72	Neutrophils_lg	pM1																			0	1.00E+00	1.00E+00
72	pM1	Neutrophils_lg																					
73	Macrophages_sc_lg	NK.cells.rest_lg																			0	1.00E+00	1.00E+00
73	NK.cells.rest_lg	Macrophages_sc_lg																					
74	Epithelial	pM2																			0	1.00E+00	1.00E+00
74	pM2	Epithelial																					
75	Endothelial.cells_lg	pM0																			0	1.00E+00	1.00E+00
75	pM0	Endothelial.cells_lg																					
76	CCN4	NK.cells.rest_lg	+	+	1.57E-02	+	1.32E-03	+	1.32E-03	+					1.57E-02	+	1.32E-03	+	1.32E-03	+	8	1.32E-03	1.57E-02
76	NK.cells.rest_lg	CCN4	+	+																			
77	NK.cells.active_lg	T.cells.CD8_lg		+	1.87E-03	+													1.87E-03	+	3	1.87E-03	1.87E-03
77	T.cells.CD8_lg	NK.cells.active_lg		+																			
78	CAF_lg	CD4Tcell_sc_lg													5.54E-03	+					1	5.54E-03	5.54E-03
78	CD4Tcell_sc_lg	CAF_lg																					
79	CAF_lg	CCN4																			1	4.75E-06	4.75E-06
79	CCN4	CAF_lg													4.75E-06	+							
80	Endothelial.cells_lg	Epithelial													1.01E-02	+					1	1.01E-02	1.01E-02
80	Epithelial	Endothelial.cells_lg																					
81	Macrophages_sc_lg	NK.cells.active_lg																			1	1.29E-04	1.29E-04
81	NK.cells.active_lg	Macrophages_sc_lg													1.29E-04	+							
82	Macrophages_sc_lg	pM1																			1	1.64E-02	1.64E-02
82	pM1	Macrophages_sc_lg													1.64E-02	+							
83	Cancer	NK.cells.active_lg																			0	1.00E+00	1.00E+00
83	NK.cells.active_lg	Cancer																					
84	Epithelial	pM0																			0	1.00E+00	1.00E+00
84	pM0	Epithelial																					

Table S4. Proliferation metrics associated CD4$^+$ and CD8$^+$ T cells stimulated in vitro in different conditions. DiI: fraction diluted; PF: Precursor frequency, %dividing cells; PI: Proliferation index; and SD^D: proliferation variance. Summary statistics were calculated from three biological replicates and represented as mean (standard deviation). Statistical significance was assessed using type III repeated measures ANOVA, where * indicates a p-value < 0.05.

	Live CD4+ T cells				Live CD8+ T cells			
Experimental Conditions	DiI	PF	PI	SD^D	DiI	PF	PI	SD^D
AP beads + rCCN4	0.670 (0.012)	0.392 (0.014)	1.407 (0.033)	0.274	0.983 (0.003)	0.851 (0.014)	2.655 (0.045)	0.103
AP beads + CCN4 KO TCM	0.472* (0.008)	0.221* (0.003)	1.404 (0.023)	0.274	0.914* (0.015)	0.715* (0.032)	1.729* (0.044)	0.205
AP beads + WT TCM	0.552* (0.047)	0.282* (0.038)	1.403 (0.031)	0.272	0.920* (0.020)	0.704* (0.044)	1.923* (0.052)	0.189
AP beads	0.655 (0.043)	0.366 (0.035)	1.473 (0.060)	0.255	0.982 (0.002)	0.841 (0.026)	2.756 (0.187)	0.106
No stimulation	0.046 (0.025)	0.016 (0.013)	1.763 (1.056)	0.371	0.062 (0.018)	0.008 (0.004)	2.520 (0.441)	0.229

Table S5. List of fluorophore-conjugated anti-mouse antibodies using to quantify cell subsets by flow cytometry.

Marker	Clone	Fluorophore	Manufacturer
LIVE/DEAD Fix	--	Violet/Pacific Blue	Invitrogen
CD45	30-F11	BB515	BD Biosciences #564590
CD3e	500A2	Alexa Fluor 700	BioLegend #152316
CD4	GK1.5	APC-Cy7	BD Biosciences #552051
CD8a	REA601	APC	Miltenyi 130-109-248
CD161 (NK-1.1)	PK136	APC-Cy7	BioLegend #108723
CD45R/B220	RA3-6B2	APC	BioLegend #103212
CD49b	DX5	PerCP/Cy5.5	Biolegend #108915
CD11b	M1/70	PerCP/Cy5.5	eBioscience #45-0112-80
CD11c	N418	PE	eBioscience #12-0114-81
F4/80	BM8	APC-Cy7	BioLegend #123117
Ly-6G/Ly-6C (Gr-1)	RB6-8C5	APC	BioLegend #108412
CD279 (PD-1)	REA802	PE	BioLegend #135205
I-A/I-E (MHC-II)	M5/114.15.2	Alexa Fluor 700	BioLegend #107622

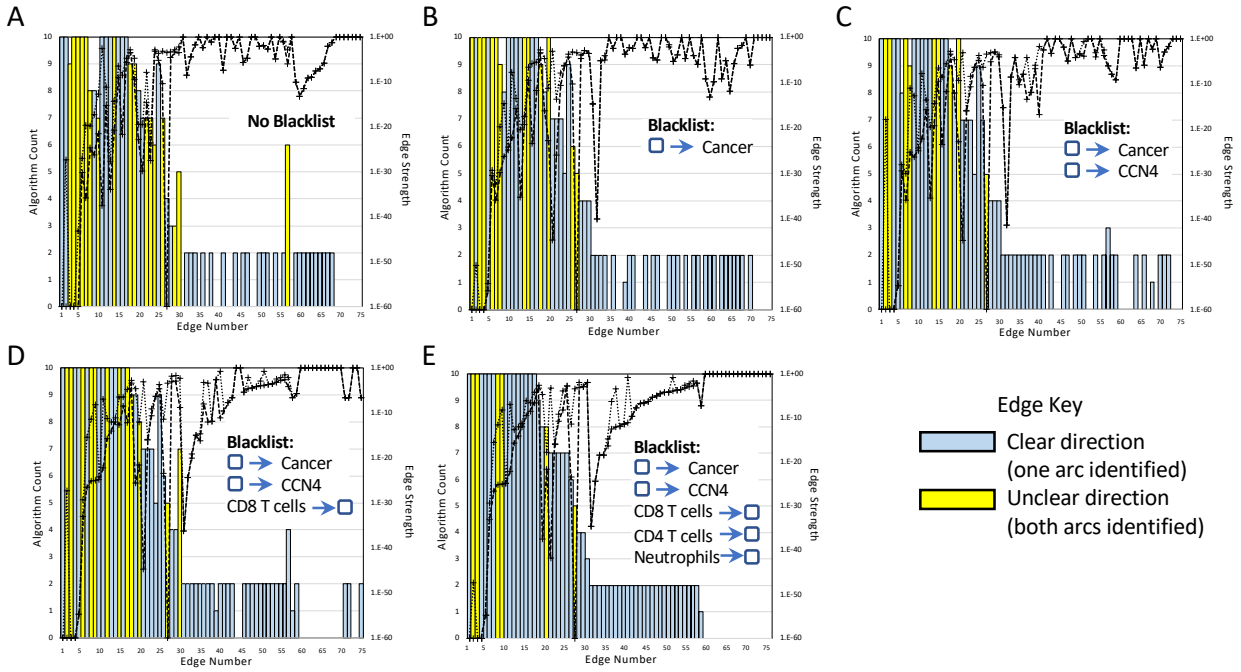
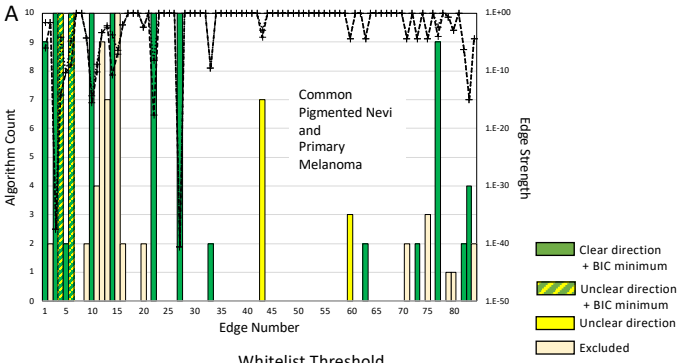
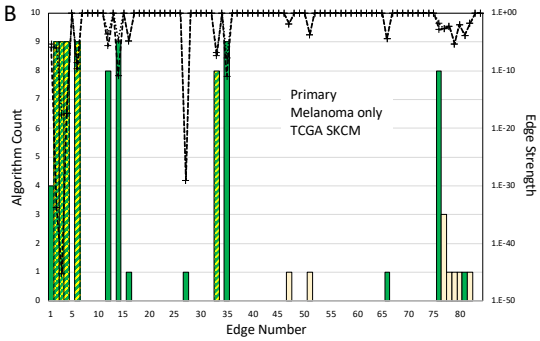


Fig. S1. Impact of arcs included in blacklist on consensus edges in the seed network. Using the TCGA BRCA dataset, a consensus seed network was generated multiple times using a blacklist that was progressively increased from an empty list (A) to a list that included cancer (B-E) and CCN4 (C-E) as tail nodes and CD8 T cells (D,E) and CD4 T cells and neutrophils (E) as head nodes. As the number of arcs included in the blacklist was increased, the number of edges with unclear direction was decreased. Seventeen edges had an unclear direction without specifying a blacklist (A), while only 6 edges had an unclear direction in the final blacklist (E). Edges were ordered based on the number of algorithms that detected either arc were enriched (bar graph - left axis) and the strength of enrichment (dotted lines - right axis). The lines associated with the strength of enrichment represent the minimum (dashed line) and maximum (dotted line) values obtained by the different algorithms for each edge. Coloring of bar graph indicates whether the direction associated with an edge was clear (gray - i.e., only one arc was found to be significant among the algorithms) or unclear (yellow - i.e., both arcs were found to be significant among the algorithms).





Whitelist Threshold

Node	Empty Graph	< 0	< 1E-20	< 1E-15	< 1E-10	< 1E-05	< 1E-03	All
	BIC / Number of Parental Arcs							
Cancer	**35.3 / 0**	35.3 / 0	35.3 / 0	35.3 / 0	35.3 / 0	35.3 / 0	35.3 / 0	35.3 / 0
CCN4	15.6 / 0	15.6 / 0	15.6 / 0	15.6 / 0	**12.9 / 1**	12.9 / 1	12.9 / 1	12.9 / 1
B Cells naive	**17.4 / 0**	17.4/ 0	17.4 / 0	17.4 / 0	17.4 / 0	17.4 / 0	17.4 / 0	18.7 / 1
CAF	11.5 / 0	**10.0 / 1**	10.0 / 1	10.0 / 1	10.0 / 1	10.0 / 1	10.0 / 1	11.9 / 2
CD4 T cells	**15.4 / 0**	15.8 / 2	15.8 / 2	15.8 / 2	15.8 / 2	15.8 / 2	15.8 / 2	15.8 / 2
CD8 T cells	22.3 / 0	18.5 / 3	**17.0 / 2**	17.0 / 2	17.0 / 2	17.0 / 2	17.0 / 2	19.6 / 4
Endothelial cells	**16.8 / 0**	16.8 / 0	16.8 / 0	16.8 / 0	16.8 / 0	16.8 / 0	16.8 / 0	16.8 / 0
Epithelial	14.8 / 0	**10.0 / 1**	10.0 / 1	10.0 / 1	10.0 / 1	10.0 / 1	11.9 / 2	11.9 / 2
Macrophages	19.2 / 0	19.2 / 0	19.2 / 0	19.2 / 0	19.2 / 0	19.2 / 0	**17.6 / 2**	17.6 / 2
Mesenchymal	19.5 / 0	13.4 / 2	13.4 / 2	13.4 / 2	13.4 / 2	**13.1 / 3**	13.1 / 3	13.1 / 3
Neutrophils	**19.2 / 0**	19.2 / 0	19.2 / 0	19.2 / 0	19.2 / 0	19.2 / 0	19.6 / 1	19.6 / 1
NK cells active	17.8 / 0	17.8 / 0	17.8 / 1	**13.8 / 2**	13.8 / 2	13.8 / 2	13.8 / 2	13.8 / 2
NK cells resting	**13.1 / 0**	15.9 / 2	13.1 / 0	13.1 / 0	13.1 / 0	13.1 / 0	13.1 / 0	13.1 / 0
pM0	**11.2 / 0**	11.2 / 0	11.2 / 0	11.2 / 0	11.2 / 0	11.2 / 0	11.4 / 1	11.4 / 1
pM1	19.2 / 0	19.2 / 0	19.2 / 0	19.2 / 0	19.2 / 0	**16.3 / 1**	16.3 / 1	18.4 / 2
pM2	21.6 / 0	**7.7 / 2**	7.7 / 2	7.7 / 2	7.7 / 2	7.7 / 2	7.7 / 2	7.7 / 2
Proliferation	**16.8 / 0**	16.8 / 0	16.8 / 0	16.8 / 0	16.8 / 1	16.8 / 1	17.0 / 2	17.0 / 2

Whitelist Threshold

Node	Empty Graph	< 0	< 1E-20	< 1E-15	< 1E-10	< 1E-05	< 1E-03	All
	BIC / Number of Parental Arcs							
Cancer	**0.0 / 0**	0.0 / 0	0.0 / 0	0.0 / 0	0.0 / 0	0.0 / 0	0.0 / 0	0.0 / 0
CCN4	**10.9 / 0**	10.9 / 0	10.9 / 0	10.9 / 0	10.9 / 0	10.9 / 0	10.9 / 0	10.9 / 0
B Cells naive	**28.5 / 0**	28.5/ 0	28.5 / 0	28.5 / 0	28.5 / 0	28.5 / 0	28.5 / 0	28.5 / 0
CAF	20.2 / 0	**15.8 / 1**	15.8 / 1	15.8 / 1	15.8 / 1	20.9 / 1	20.9 / 1	20.9 / 1
CD4 T cells	35.9 / 0	**32.3 / 1**	32.3 / 1	32.3 / 1	32.3 / 1	32.3 / 1	32.3 / 1	33.6 / 3
CD8 T cells	20.8 / 0	**18.9 / 1**	18.9 / 1	18.9 / 1	18.9 / 1	18.9 / 1	19.5 / 2	21.4 / 3
Endothelial cells	25.0 / 0	**22.8 / 1**	22.8 / 1	22.8 / 1	22.8 / 1	22.8 / 1	22.8 / 1	25.0 / 0
Epithelial	**16.0 / 0**	16.0 / 0	16.0 / 0	16.0 / 0	16.0 / 0	16.0 / 0	16.0 / 0	18.0 / 1
Macrophages	22.2 / 0	19.6 / 1	19.6 / 1	19.6 / 1	19.6 / 1	19.6 / 1	**19.0 / 2**	20.4 / 3
Mesenchymal	20.3 / 0	20.3 / 0	20.3 / 0	20.3 / 0	20.3 / 0	**16.6 / 2**	16.6 / 2	16.6 / 2
Neutrophils	**25.6 / 0**	25.6 / 0	25.6 / 0	25.6 / 0	25.6 / 0	25.6 / 0	25.6 / 0	25.6 / 0
NK cells active	32.4 / 0	**25.5 / 1**	25.5 / 1	25.5 / 1	25.5 / 1	25.5 / 1	25.5 / 1	25.5 / 1
NK cells resting	33.2 / 0	**31.9 / 1**	31.9 / 1	31.9 / 1	31.9 / 1	31.9 / 1	33.2 / 0	31.9 / 1
pM0	23.9 / 0	23.9 / 0	23.9 / 0	23.9 / 0	23.9 / 0	23.9 / 0	**23.6 / 1**	23.6 / 1
pM1	19.1 / 0	**17.6 / 1**	17.6 / 1	17.6 / 1	17.6 / 1	17.6 / 1	17.6 / 1	17.6 / 1
pM2	20.9 / 0	15.0 / 1	**10.8 / 2**	10.8 / 2	10.8 / 2	10.8 / 2	10.8 / 2	10.8 / 2
Proliferation	**16.8 / 0**	16.8 / 0	16.8 / 0	16.8 / 0	16.8 / 0	16.8 / 0	17.0 / 1	17.0 / 1

Fig. S2. Summary of the evidence obtained from two melanoma datasets supporting the consensus edges in the seed network. Analysis of datasets containing samples from both common pigmented nevi and primary melanoma (A) and from only primary melanoma (B). For each edge listed in Tables S2 and S3, the number of algorithms that detected than an edge was enriched (bar graph - left axis) and the strength of enrichment (dotted lines - right axis) are shown. The lines associated with the strength of enrichment represent the minimum (dashed line) and maximum (dotted line) values obtained by the different algorithms for each edge. Bar graph coloring indicates whether an edge was significantly enriched with a clear direction and included in the set of arcs associated with the BIC minimum (green), significantly enriched without a clear direction but included in the set of arcs associated with the BIC minimum (yellow/green), significantly enriched but without a clear direction (yellow), or excluded from the consensus seed network list (tan).

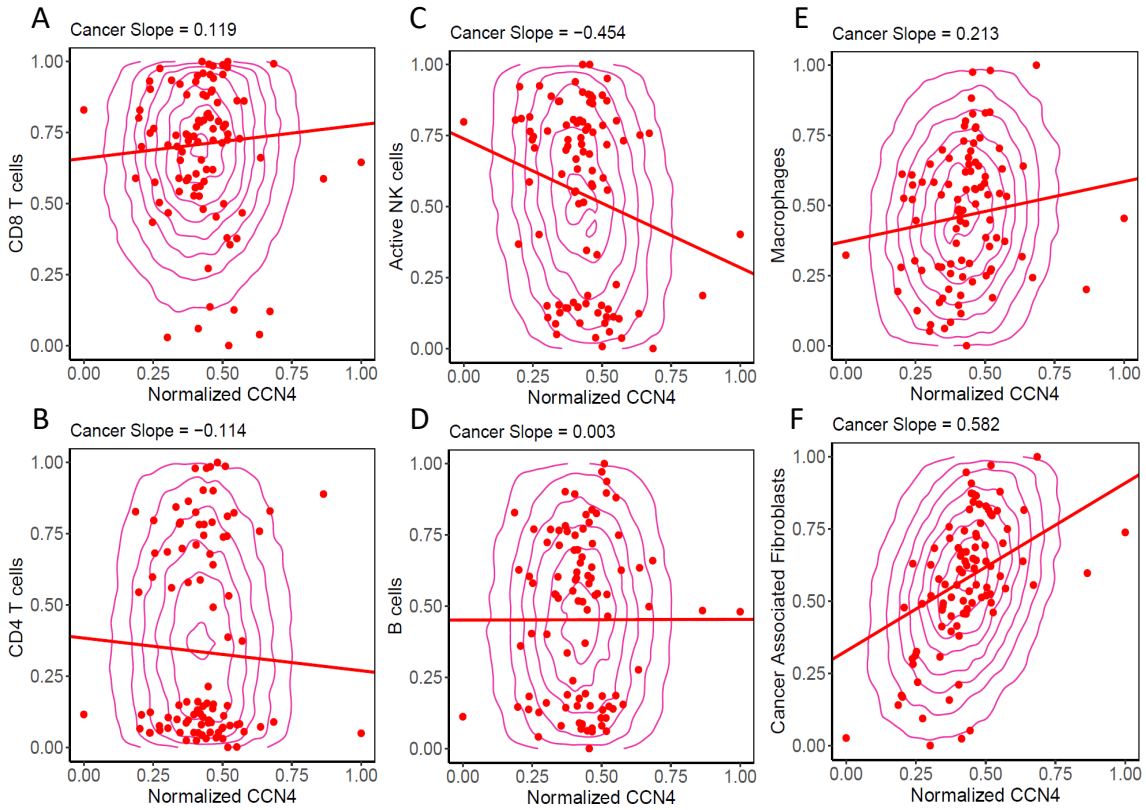


Fig. S3. Conditional probability query of the SKCM DAG compared against digital cytometry estimates obtained from experimental data. Experimental samples obtained from primary melanoma tissue are shown as red dots. Samples of the conditional probability model for Cancer > 0.95 (red contours) for CD8 T cells (A), CD4 T cells (B), active NK cells (C), B cells (D), Macrophages (E) and Cancer Associated Fibroblasts (F). Linear trend lines are superimposed on the conditional probability samples.

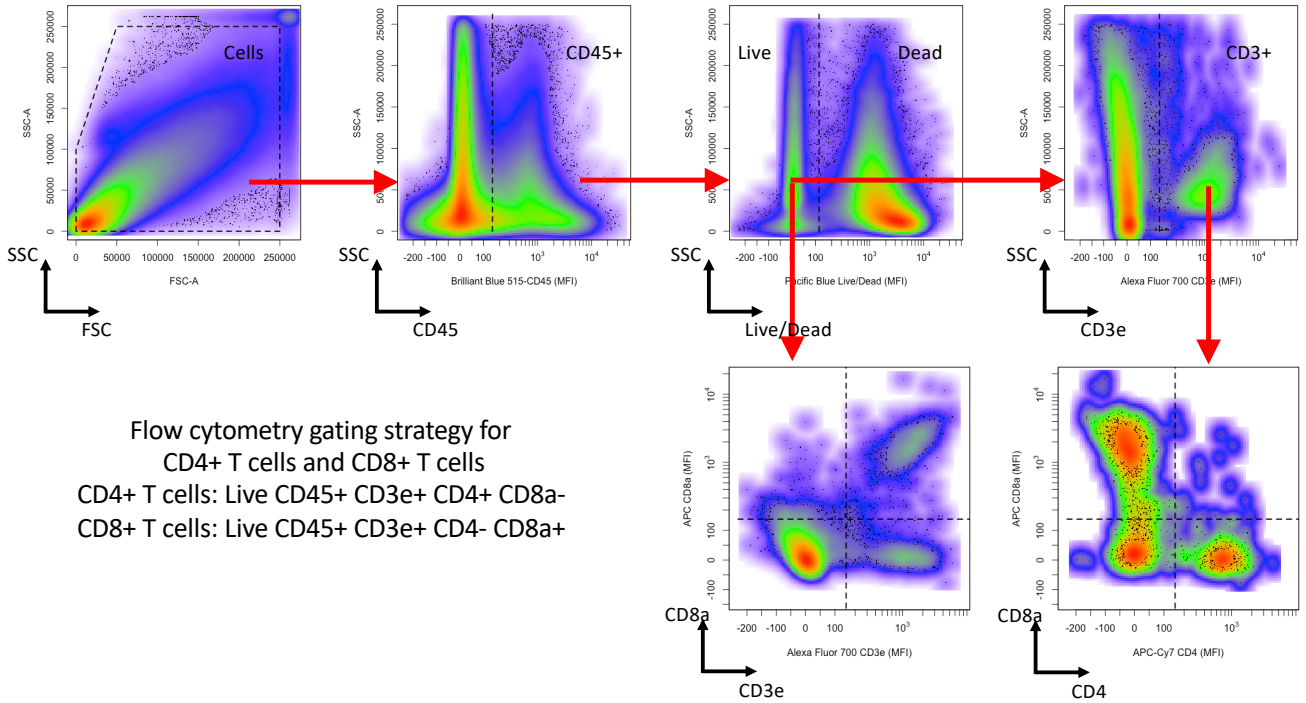


Fig. S4. Flow cytometry gating strategy for T cells. CD45 staining versus side scatter area was used to gate for CD45+ cells. Live Dead Pacific Blue staining versus side scatter area was used to then gate for Live CD45$^+$ cells, which were then gated based on CD3e$^+$ expression. Live CD45$^+$ CD3e$^+$ cells were further subdivided into CD8$^+$ T cells (live CD8$^+$ CD3e$^+$ CD45$^+$ cells), CD4 T cells (live CD4$^+$ CD3e$^+$ CD45$^+$ cells), and double negative T cells (live CD8$^-$ CD4$^-$ CD3e$^+$ CD45$^+$ cells).

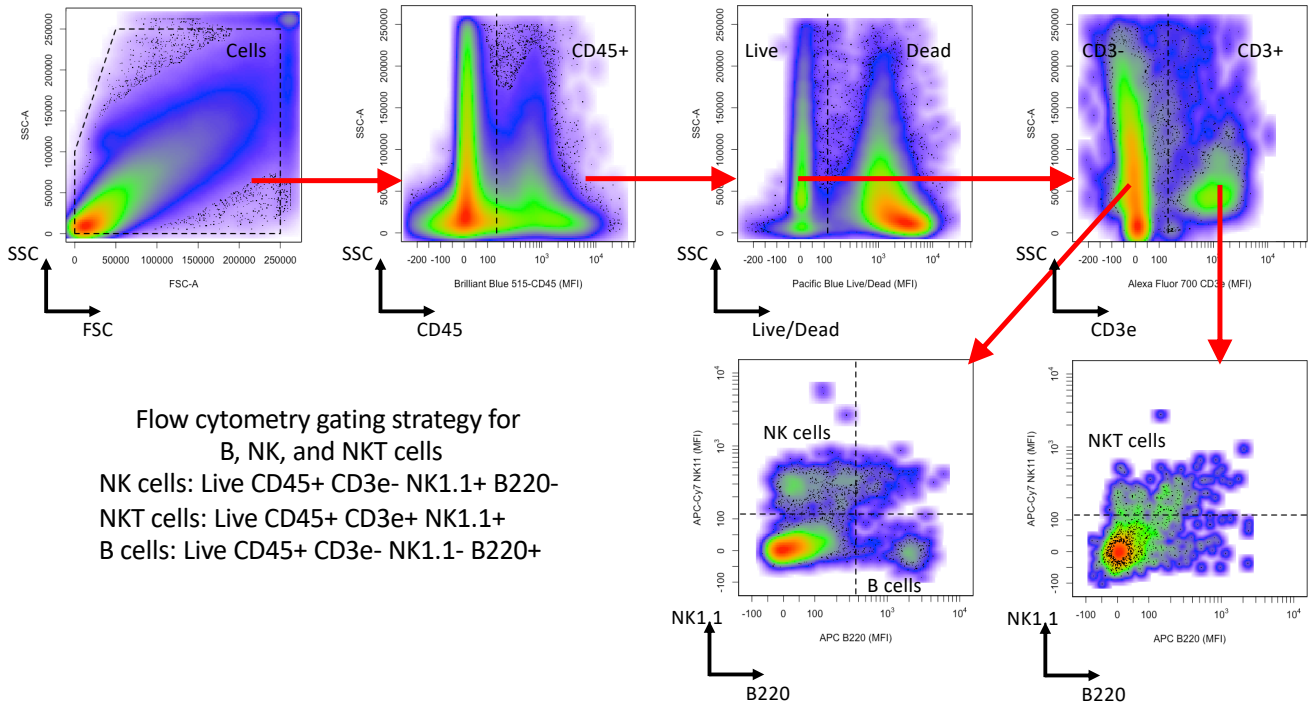


Fig. S5. Flow cytometry gating strategy for B, NK, and NKT cells. CD45 staining versus side scatter area was used to gate for CD45+ cells. Live Dead Pacific Blue staining versus side scatter area was used to gate for Live CD45$^+$ cells, which were then subdivided into B cells (live NK1.1$^-$ B220$^+$ CD3$^-$ CD45$^+$ cells), NK cells (live NK1.1$^+$ B220$^-$ CD3$^-$ CD45$^+$ cells), and NKT cells (live NK1.1$^+$ CD3e$^+$ CD45$^+$ cells).

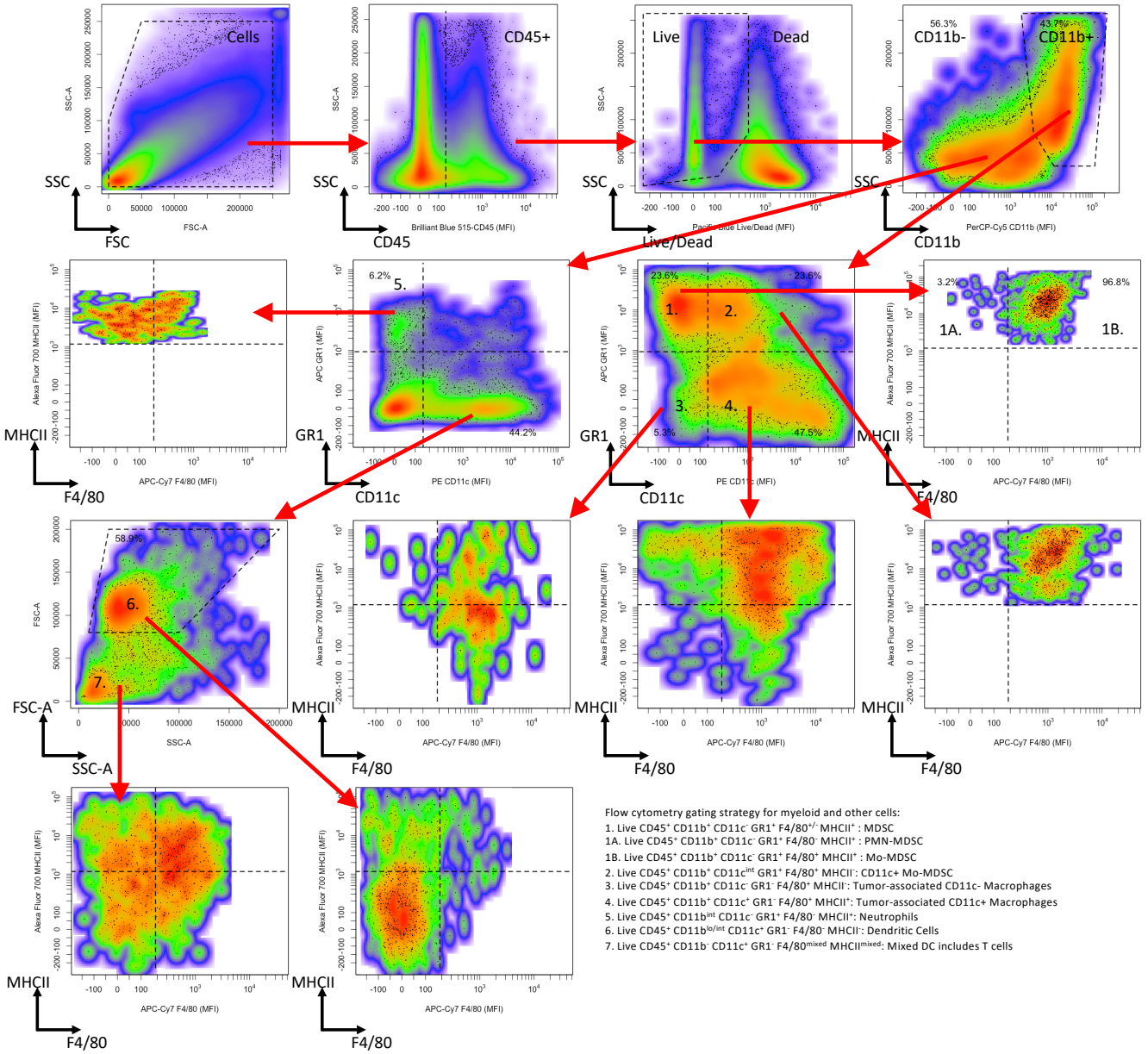


Fig. S6. Flow cytometry gating strategy for Tumor associated neutrophils and myeloid cell subsets. CD45 staining versus side scatter area was used to gate for CD45$^+$ cells. Live Dead Pacific Blue staining versus side scatter area was used to gate for Live CD45$^+$ cells, which were then subdivided into subsets based on CD11b staining followed by Gr1 versus CD11c staining. From the CD11b$^+$ gate, myeloid-derived suppressor cells (MDSC) (live CD45$^+$ CD11b$^+$ Gr1$^+$ cells) were subdivided into CD11c$^{int/+}$ MDSC (F4/80$^+$ MHCII$^+$) and CD11c$^-$ MDSC (F4/80^{mixed} MHCII$^+$). Also from the CD11b$^+$ gate, macrophages (live Gr1$^-$ F4/80$^+$ CD11b$^+$ CD45$^+$ cells) were subdivided into tumor-associated CD11c$^+$ (CD11c$^{int/+}$ MHCII^{hi}) and CD11c$^-$ (CD11c$^-$ MHCII^{lo})subsets. The CD11b$^-$ subset included tumor-associated neutrophils (TAN) (Gr1$^+$ CD11c$^-$ CD11b^{int} MHCII^{hi} F4/80$^-$) and dendritic cells (Gr1$^-$ CD11c$^+$ CD11b$^{lo/int}$ FSC-A^{hi} MHCII^{lo} F4/80$^-$).

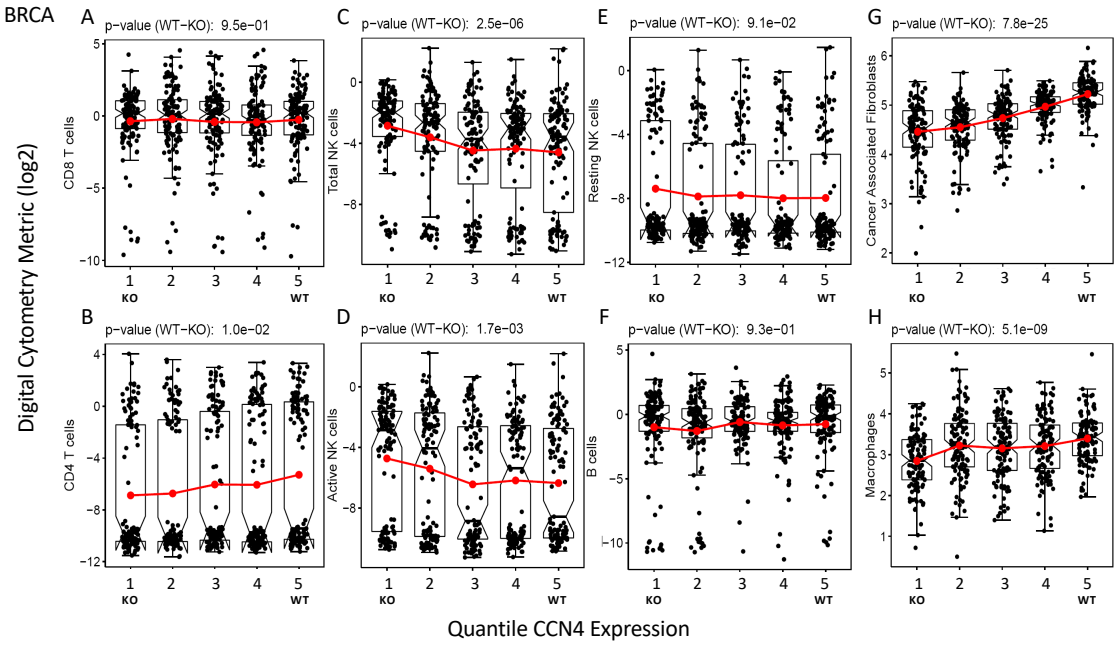


Fig. S7. Distribution in digital cytometry features extracted from the BRCA dataset stratified into quantiles based on CCN4 expression. The mean value for quantile 1 was used to represent CCN4 KO phenotype while quantile 5 was used to represent WT phenotype.

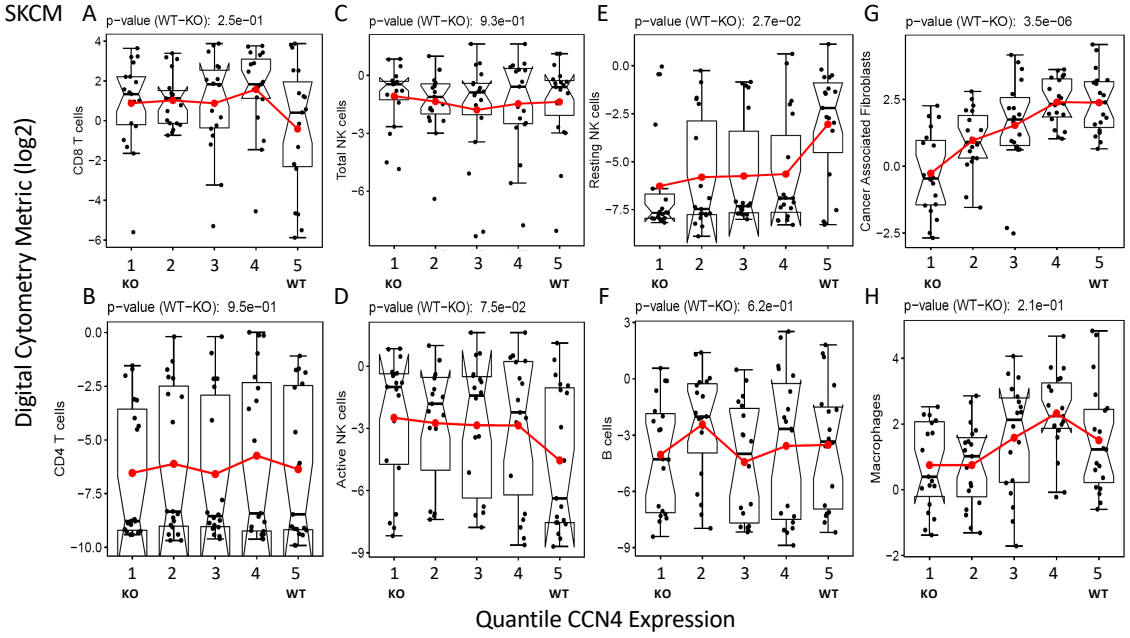


Fig. S8. Distribution in digital cytometry features extracted from the SKCM dataset stratified into quantiles based on CCN4 expression. The mean value for quantile 1 was used to represent CCN4 KO phenotype while quantile 5 was used to represent WT phenotype.

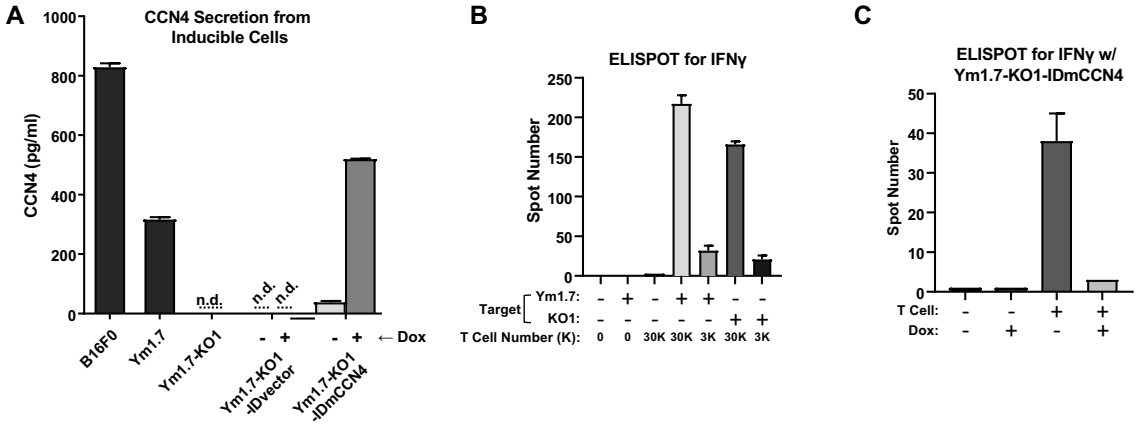


Fig. S9. Control experiments related to ELISpot assay using an inducible CCN4 YUMM1.7 cell line. (A) CCN4 secretion, measured with ELISA, from CCN4-inducible cells in conditioned media in the presence or absence of 0.5 μg/ml doxycycline (n.d. = not detected). (B) ELISpot for IFNγ release with different target cells and different amount of effector CD8$^+$ T cells (In vivo activated CD8$^+$ T cells against YUMM1.7 (Ym1.7)). (C) ELISpot for IFNγ with CCN4-inducible cells as targets using in vivo activated CD8+ T cell against YUMM1.7. Results shown as mean $\pm$ S.D. for three biological replicates.