Multiplexed Imaging Analysis of the Tumor-Immune Microenvironment Reveals Predictors of Outcome in Triple-Negative Breast Cancer

Aalok Patwa[1*], Rikiya Yamashita[1,2*], Jin Long[2], Michael Angelo[3], Leeat Keren[4], Daniel L. Rubin[1,2**]

1. Department of Biomedical Data Science, Stanford University
2. Center for Artificial Intelligence in Medicine and Imaging, Stanford University
3. Department of Pathology, Stanford University
4. Department of Molecular Cell Biology, Weizmann Institute of Science

* Equal contribution

** Corresponding author: dlrubin@stanford.edu

Abstract

Triple-negative breast cancer (TNBC), the poorest-prognosis breast cancer subtype, lacks clinically approved biomarkers for patient risk stratification, treatment management, and immunotherapies. Prior literature has shown that interrogation of the tumor-immune microenvironment (TIME) may be a promising approach for the discovery of novel biomarkers that can fill these gaps. Recent developments in high-dimensional tissue imaging technology, such as multiplexed ion beam imaging (MIBI), provide spatial context to protein expression in the TIME, opening doors for in-depth characterization of cellular processes. We developed a computational pipeline for the robust examination of the TIME using MIBI. We discover that profiling the functional proteins involved in cell-to-cell interactions in the TIME predicts recurrence and overall survival in TNBC. The interactions between CD45RO and Beta Catenin and CD45RO and HLA-DR were the most relevant for patient stratification. We demonstrated the clinical relevance of the immunoregulatory proteins PD-1, PD-L1, IDO, and Lag3 by tying their interactions to recurrence and survival. Multivariate analysis revealed that our methods provide additional prognostic information compared to clinical variables. Our novel computational pipeline produces interpretable results, and is generalizable to other cancer types.

Introduction

Triple-negative breast cancer (TNBC) is a subtype of breast cancer that is negative for estrogen receptor, progesterone receptor, and human epidermal growth factor receptor 2. Representing an estimated 10-20% of breast cancers, it is characterized by aggressive behavior, including earlier onset, larger tumor size, and a more advanced grade[1,2]. TNBC is the subtype of breast cancer with the poorest prognosis [3], having a lower chance of survival[4,5] and higher risk of recurrence, especially within a short timeframe[6,7]. The absence of common breast cancer hormonal targets as well as high heterogeneity among TNBC tumors makes treatment management and therapy targeting difficult, creating a need for more advanced interrogation of cellular processes within TNBC tumors[8]. Currently administered treatments, such as checkpoint inhibitors, only provide benefit to a small proportion of treated patients and are associated with high cost and toxicity[9]. Their effectiveness is limited, necessitating further interrogation of cancer-cell cues, factors in the tumor microenvironment, and host-related influences[10]. There is a need to identify additional biomarkers of TNBC to aid prognosis and identify new immunotherapies[11–13]. Identifying predictors of recurrence and survival in TNBC patients would serve two critical purposes. First, it would allow for patient risk stratification, which would lead to more specific treatments and spare patients from unnecessary aggressive therapies[14]. Second, it would identify novel biomarkers as potential new therapeutic targets.

The tumor-immune microenvironment (TIME) is a dynamic system comprising cancer cells, immune cells, and the surrounding extracellular matrix and vasculature[15]. The TIME is modulated by the expression and secretion of signaling molecules and other proteins that contribute to angiogenesis, immune suppression, as well as the coordination of the immune response[16]. Previous research has sought to discover the features of the TIME that are tumor-promoting or tumor-rejecting using transcriptomic and proteomic data [17–20].

However, until recently, conventional histological techniques lacked the ability to measure the expression of a multitude of proteins at subcellular resolution while preserving spatial information[21,22]. Advancements in high-dimensional multiplexed imaging, such as through the development of multiplexed ion beam imaging (MIBI), have allowed for more direct interrogation of the TIME[23] while boosting standardization and reproducibility of results[24]. MIBI uses secondary ion mass spectrometry to image antibodies tagged with isotopically pure elemental reporters[25]. MIBI is compatible with formalin-fixed paraffin-embedded (FFPE) tissue samples, the foremost preservation method of solid tissue in routine clinical pathology. MIBI enables in-depth analysis of the interactions of the TIME, measuring the expression of more than 40 proteins simultaneously while preserving spatial information[26] and avoiding spectral overlap[27] and autofluorescence[28].

This study builds on the work of Keren et al[22], who found structure in the composition and spatial organization of the TIME. TIME architecture was broadly classified as immune cold, mixed, or compartmentalized. The cold/compartmentalized architecture distinction was linked to survival.

However, besides this distinction, it is still unknown whether features of the TIME are linked to recurrence and overall survival. Recurrence, a key outcome for patient management, has not been studied. Although previous work has studied survival, there is still a need to study more granular features of the TIME. Examining the function of individual proteins can add prognostically relevant information[29], but further research is needed to investigate interactions at subcellular resolution[30].

This study aims to uncover features of the TIME that are associated with recurrence and overall survival by analyzing MIBI scans of TNBC tissue[22,25]. The primary focus is to profile the proteins involved in cell-to-cell interactions and establish a link between the spatial organization of cells with varying expression patterns clinical outcome. We examine interactions involving functional proteins and immunoregulatory proteins in particular. As corollary aims, we demonstrate association between protein co-expression patterns and recurrence / survival, examine proteins whose overall expression is associated with recurrence / survival, and test associations between immune composition and recurrence / survival. We also propose novel biomarkers captured by MIBI that show potential as therapeutic targets.

Results

Overview of Results

We examined multiple features of the TIME to identify which features were most predictive of recurrence and survival in our cohort. We found that low-level features, namely cell type prevalence and protein expression, did not predict recurrence or survival. We then examined higher level features with spatial context. We found that protein co-expression, cell-to-cell interactions of functional proteins, as well as cell-to-cell interactions of immunoregulatory proteins, a subset of the functional proteins, held predictive power. With this, we performed multivariate analysis to find the most prognostically relevant. The immunoregulatory protein interaction features were the most predictive of recurrence, and the functional protein features were the most predictive of survival. Our findings show that quantifying cell-to-cell interactions holds potential for patient risk stratification, particularly because they provided more prognostic information than overall protein expression or cell type prevalence. Ultimately, we provide a computational pipeline that interrogates the TIME and uncovers association between spatial organization and two clinical outcomes.

Patient Population

Our study examines 38 triple-negative breast cancer (TNBC) patients, a subset of the 41 patients examined by Keren et al[22]. Formalin fixed, paraffin embedded (FFPE) slides of breast tissue were taken from patients and scanned using MIBI, and subsequently segmented to demarcate cell boundaries[22]. Patient data regarding age, tumor grade, stage, cancer site, and clinical outcome – recurrence and overall survival (OS) -- were also gathered (Table 1).

Characteristic		**Overall**
# of Patients		38
Age		54 (15)
Grade	**1**	1 (2.7)
	2	5 (13.2)
	3	29 (76.3)
	4	2 (5.3)
	Missing	1 (2.6)
Cells per image		5006 (1527)

Table 1. Patient cohort characteristics. Age and cells per image are represented by mean (standard deviation). Grade and stage are represented by count (% of total).

Dataset

MIBI scans produce images of protein expression from FFPE tissue, where each image has 44 channels; each channel conveys the expression of a certain protein on the tissue sample (Figure 1a). A computational pipeline was used to segment individual cells from the images. The resulting segmented images present the location of various cell types on the slide (Figure 1b).

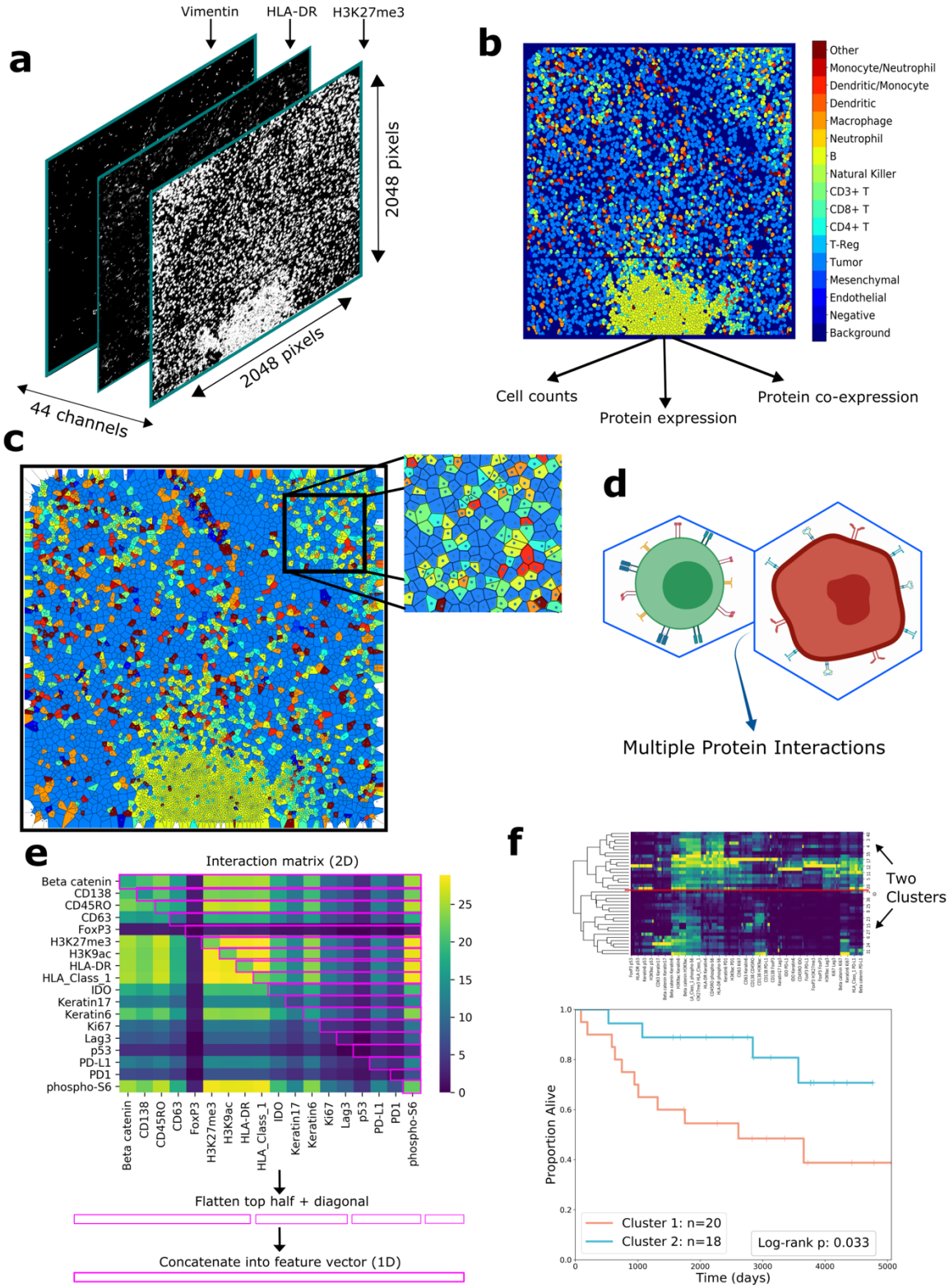


Figure 1. Overview of computational pipeline. a Drawing of TIFFs created by MIBI scans. The TIFFs have dimensions 2048x2048 pixels with 44 channels, one for each protein. Each pixel in the image at each channel conveys the concentration of that protein at that location. **b** Color-mapped image of cell segmentation performed on MIBI TIFFs. The image now has one channel with dimensions 2048x2048. Each cell has its own connected

component with a unique grayscale value from 0-16 representing which cell type it is (see color bar on right). From this, I extract cell counts, measure protein expression, and quantify co-expression. **c** Voronoi representation of the segmented image in 1b. Each cell has its own region and borders a finite number of other cells. **d** Using Voronoi diagrams, I analyze interactions between neighboring cells. (Drawing created from Biorender.com) **e** An interaction matrix is computed for each patient, with the entry at row A and column B representing the number of times a cell positive for protein A was adjacent to a cell positive for protein B (top). The top half of the matrix, including the diagonal, is selected as shown with the purple rectangles. These rectangles are then flattened to form one feature vector for each patient. **f** Interaction features are used to cluster patients (top), and the two clusters are compared with regard to outcomes (bottom).

Immune composition of cells is not associated with recurrence and survival.
We examined whether the prevalence of certain cell populations in the TIME was associated with recurrence and survival. We computationally measured the number of cells of each cell type in each patient and represented that number as a proportion of the total number of cells in that patient's sample. We then performed univariate Cox regression to determine whether each cell type's proportion was related to recurrence and overall survival.

After Benjamini-Hochberg correction[31], there were no cell types whose coefficients were significant for either recurrence or overall survival. However, some information can still be extracted from the hazard ratios shown (Table 2a and 2b).

Cell Type	Coefficient	Hazard Ratio	Coefficient P	BH-Corrected FDR
Monocyte/Neutrophil	-0.372	0.689	0.093	0.647
CD8+ T	-0.072	0.93	0.154	0.647
Macrophage	-0.053	0.949	0.182	0.647
Mesenchyme	0.037	1.038	0.183	0.647
CD4+ T	-0.07	0.933	0.244	0.647
Tumor	0.011	1.011	0.286	0.647
Natural Killer	-0.867	0.42	0.329	0.647
Dendritic/Monocyte	-0.118	0.888	0.345	0.647
Endothelial	0.103	1.109	0.46	0.767
B	-0.013	0.988	0.68	0.899
Neutrophil	-0.046	0.955	0.682	0.899
Dendritic	0.022	1.022	0.739	0.899
CD3+ T	0.039	1.04	0.779	0.899
Other	0.004	1.004	0.942	0.971
Regulatory T	-0.009	0.991	0.971	0.971

Table 2a. Immune composition Cox regression results for recurrence. There is no association between cell prevalence and recurrence in the cohort.

Cell Type	Coefficient	Hazard Ratio	Coefficient P	BH-Corrected FDR
Dendritic/Monocyte	-0.292	0.747	0.076	0.435
Tumor	0.018	1.018	0.100	0.435
Monocyte/Neutrophil	-0.331	0.718	0.121	0.435
Other	-0.159	0.853	0.124	0.435
Macrophage	-0.066	0.936	0.145	0.435
Mesenchyme	-0.047	0.954	0.308	0.647
CD3+ T	-0.169	0.845	0.315	0.647
Dendritic	0.048	1.05	0.41	0.647
Neutrophil	-0.119	0.888	0.429	0.647

Natural Killer	0.357	1.43	0.431	0.647
Endothelial	0.102	1.107	0.505	0.689
CD4+ T	-0.022	0.978	0.628	0.692
B	-0.015	0.985	0.644	0.692
Regulatory T	-0.129	0.879	0.646	0.692
CD8+ T	0.006	1.006	0.88	0.880

Table 2b. Immune composition Cox regression results for survival. There is no association between cell prevalence and survival in the cohort.

Single-cell expression levels of functional proteins are not associated with recurrence and survival.

We examined whether the expression of functional proteins in the cells of the tissue samples was predictive of recurrence and survival (Figure 2a). We calculated the per-pixel expression levels of each protein in each patient. The histograms of expression for several proteins is shown in Figure 2b, and the histograms for all proteins is shown in Figure S1. For this analysis, we included only functional proteins. These proteins stand in contrast to proteins used for lineage assignment; their expression is modulated according to the functional state of the cell. The distinction is shown in Table S1.

There were no functional proteins whose Beta coefficients had significant p-values after Benjamini-Hochberg correction for either recurrence (Table S2a) or overall survival (Table S2b).

Within this cohort, the expression of functional proteins did not hold predictive power. As such, we decided to move away from macro-level interrogation of the TIME, opting to add spatial context to our analysis by quantifying protein co-expression and cell-to-cell interactions.

Co-expression of functional proteins in patients' cells is associated with recurrence and survival.

We sought to develop a computational pipeline to determine whether environmentally derived signals have pleiotropic effects in driving overall localization and coordination of immune inhibition[22]. We created co-expression matrices for each patient to summarize the number of times pairs of proteins were co-expressed in a patient to interrogate immune coordination (Figure 2c).

The co-expression matrices provide more information about the phenotypes of the cells present in each patient, placing the expression of proteins in a single-cell context. We used the co-expression information as features to describe each patient. Patients were grouped by hierarchical clustering, and the tree was cut to form two patient clusters (Figure 2d). Our choice to select two clusters in this analysis, as well as all hierarchical clustering analysis, was motivated by silhouette score analysis[32], which showed that division into two clusters would maximize inter-cluster dissimilarity. The results of the silhouette score analysis can be seen in Table S3. The Kaplan-Meier curves of these two patient clusters diverged in the case of both recurrence (log-rank p=0.053) and overall survival (log-rank p=0.094) (Figure 2e). We also tested patient stratification when three clusters were chosen. The log-rank p-value for recurrence was 0.093 and 0.222 for survival (Figure S2).

We assessed the relative importance of individual co-expression features using random forest variable importance. Feature importance was scored using the mean decrease in Gini Index. The four most important co-expression features were CD45RO + H3K27me3 (score=0.822), CD45RO + H3K9ac (score=0.767), CD45RO + HLA Class 1 (score=0.646), and HLA-DR + IDO (score=0.604). These results show that calculating the co-expression of proteins, namely the combinations listed above, can aid patient stratification. Interestingly, CD45RO has been discussed in the literature as having importance in anti-tumor immunity[33,34]. Its co-expression with HLA Class 1, an antigen used to promote cytotoxic T cell activation, is aligned with existing literature on melanoma[35], and may evidence coordination between memory T cells and cytotoxic T cells in cancer.

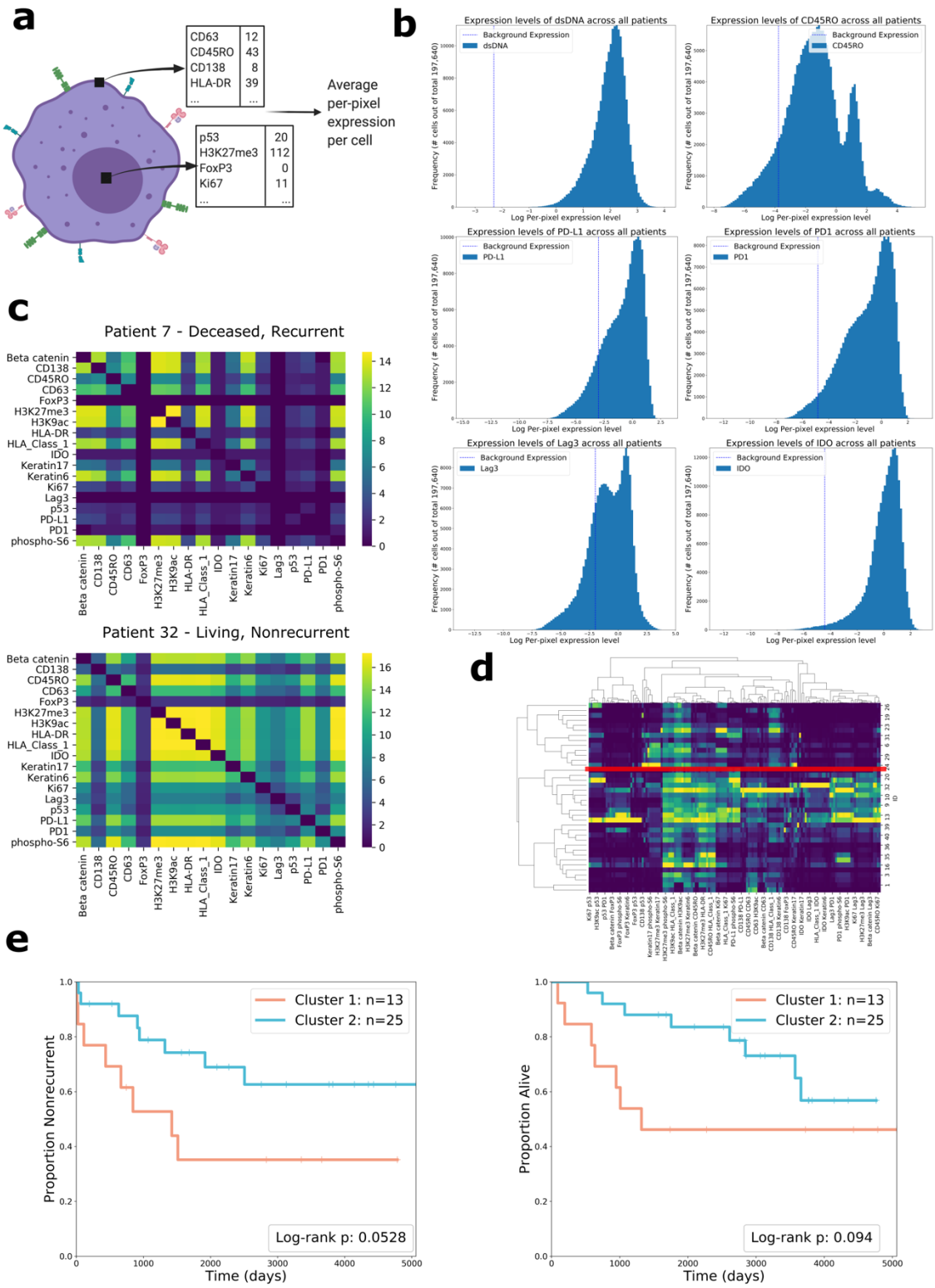


Figure 2. Quantification and analysis of protein expression. a Drawing showing how protein expression is calculated. The black squares each represent one pixel in the image. Expression levels are measured for each pixel in the cell and then summed across all pixels in the cell. The resulting number is divided by the size of the cell (in pixels), resulting in the average per-pixel expression level of the cell for each protein. (Drawing created from Biorender.com) **b** Histograms showing the distributions of log per-pixel expression levels for several relevant

proteins. Per-pixel expression in the background channel (the positivity threshold) is shown with the vertical dotted line. **c** Heatmaps showing the cube root of co-expression of pairs of functional proteins in two different patients. The colorbar also shows the cube root, so color value 16 indicates 16^3 instances of co-expression. **d** Clustermap showing flattened features for all 38 patients. Two clusters were chosen from the dendrogram. The red line shows the way that the two clusters were separated. **e** Kaplan-Meier curves comparing clusters formed from co-expression features for recurrence (left) and overall survival (right). Log-rank test p-value shown in plot legend.

Cell-to-cell interactions contain prognostically relevant information.

We examined cell-to-cell interactions by creating Voronoi diagrams out of the segmented MIBI images (Figure 1c). Voronoi polygons have been used previously to define spatial organization and cellular morphology overall[28,36]. Each cell's Voronoi polygon is created from the location of its centroid; its polygon will border some number of polygons from other cells[37]. These borders can be used to model cell-to-cell interactions (Figure 1d); cells whose polygons share a border can be considered adjacent (Figure 3a). Due to the geometry of the Voronoi tessellation algorithm, polygons will only border their immediate neighbors, which restricts the area of influence of a certain cell to the cells that are closest nearby.

We then created interaction matrices to summarize all of the interactions in a patient. Therefore, the entry in the matrix at row A and column B represents the number of times a cell positive for protein A was adjacent to a cell positive for protein B (Figure 3b).

Functional protein interaction data was used as features for hierarchical clustering, resulting in two clusters, with 17 patients in Cluster 1 and 21 patients in Cluster 2 (Figure 3c). The Kaplan-Meier curves comparing these two patient clusters diverged significantly for recurrence (log-rank p=0.065) and overall survival (log-rank p=0.033) (Figure 3d).

Our method of quantifying cell-to-cell interactions reveals that the spatial proximity of functional proteins contains valuable prognostic information; the number of these interactions can be used as features to cluster patients into groups with significantly different outcomes.

By contrast, quantifying interactions between lineage proteins does not hold predictive power. Hierarchical clustering on features of lineage protein interactions did not result in clusters that differed in recurrence and survival outcome significantly (Figure S3).

A drawing comparing the clusters formed from clustering on functional protein interaction features to the morphology distinction performed by Keren et al is shown in Figure S5.

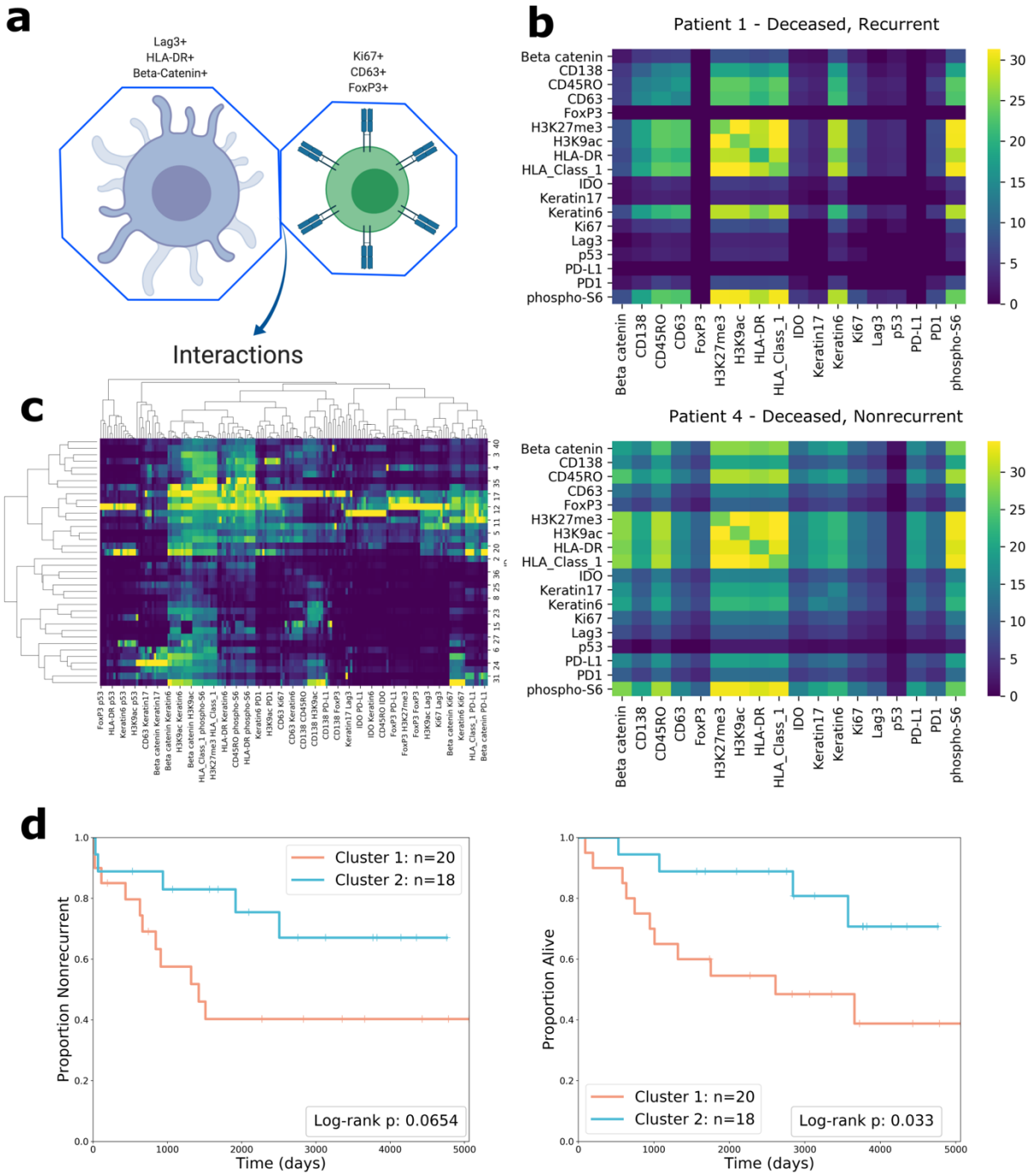


Figure 3. Analysis of cell-to-cell interactions. a Drawing showing how interactions are analyzed to find which combinations of proteins are involved in the interaction. The interaction is characterized by the adjacency of the two Voronoi polygons. Each cell involved in the interaction has a unique protein expression pattern, resulting in complex interactions shown in the bottom arrow. (Drawing created from Biorender.com) **b** Heatmaps showing the cube root of the number of interactions between pairs of functional proteins in two patients. The entry at row A and column B in the heatmap represents the cube root of the number of times that a cell positive for protein A was adjacent to a cell positive for protein B in that patient's MIBI image. Pairs who had zero interactions are excluded from the plot. **c** Clustermap of patients' functional protein interactions features. **d** Kaplan-Meier curves of recurrence (left) and overall survival (right) comparing clusters formed from interaction features. Log-rank test p-values are shown in the plot legend.

Interactions between immunoregulatory proteins predict recurrence and survival.
We further examined a subset of functional proteins, the immunoregulatory proteins: PD-1, PD-L1, IDO, and Lag3, which are in consideration as immunotherapy targets[9,38–42]. Prior research did not answer whether interactions involving these four proteins are associated with recurrence and survival which would be valuable information in their designation as immunotherapy targets. We sought to answer this question.

To answer this question, we quantified spatial interactions between cells expressing immunoregulatory proteins, excluding all other proteins from the analysis (Figure 4a). Similarly to previous analysis, the counts of interactions were used as features to cluster patients (Figure 4b). The Kaplan-Meier curves of the clusters formed from this analysis diverged significantly when assessing recurrence (log-rank p=0.0058) and overall survival (log-rank p=0.10) (Figure 4c). We also tested patient stratification when three clusters were chosen. The log-rank p-value for recurrence was 0.020 and 0.173 for survival (Figure S4).

Ablation analyses reveals predictive groups of features.
We further examined the cell-to-cell interaction data by performing multiple ablation analyses.

For the first ablation analysis, we examined individual functional proteins one-at-a-time, including only the interactions that involved this protein as features (Figure 4d). For example, when examining the interactions of PD-1, we constructed feature vectors to include PD-1/Lag3 interactions, PD-1/Ki67 interactions, PD-1/PD-L1 interactions, and so on. In the case of recurrence, several proteins had interactions who were predictive: IDO (p=0.008), HLA Class 1 (p=0.011), H3K27me3 (p=0.011), Beta Catenin (p=0.023), PD-1 (p=0.060), and CD45RO (p=0.098). Phospho-S6 and Lag3 had interactions which differed across overall survival groups with p-values 0.041 and 0.052, respectively.

We also examined "homotypic" interactions – interactions between the same protein. Homotypic interactions are found in the diagonal of the interaction matrix – they represent the number of times that a cell positive for protein A was adjacent to a cell positive for protein A (Figure 4e). This information communicates the spatial proximity of cells with similar expression patterns. We used all of the homotypic interactions of functional proteins (the entire diagonal) as features for each patient and repeated the clustering analysis. The Kaplan-Meier curves diverged significantly, with a log-rank p of 0.064 for recurrence and 0.027 for overall survival, indicating that homotypic interactions contain valuable information in particular that can be used to predict recurrence.

We calculated the importance of interaction features by fitting a random forest model with interactions as predictors and cluster assignments as the response variable. Feature importance was scored using the mean decrease in Gini Index. The highest-importance feature was the Beta Catenin + CD45RO interaction feature (score=0.794), followed by CD45RO + HLA-DR (score=0.738), PD-1 + CD45RO (score=0.716), PD-1 + H3k27me3 (score=0.709), Lag3 + CD45RO (score=0.706), IDO + PD-1 (score=0.694), and Lag3 + PD-1 (score=0.647). Within these 7 interaction features, CD45RO appears 4 times, PD-1 appears 4 times, and Lag3 appears twice. These results point to these interaction features being particularly useful for patient

stratification; they contributed the most to clustering, and the resulting clusters differed significantly in terms of recurrence and survival.

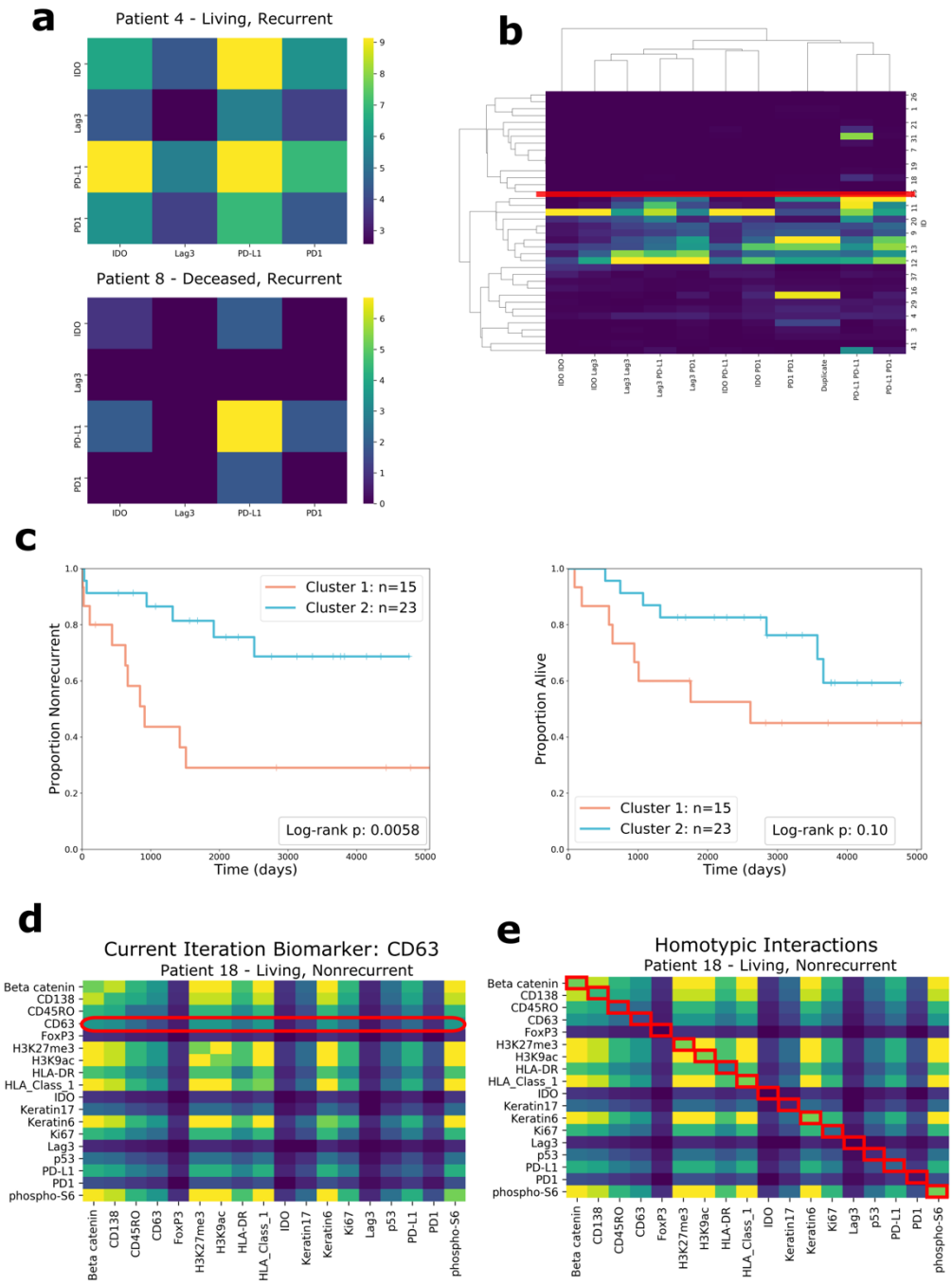


Figure 4. Analysis of subsets of interactions. a Heatmaps of the interaction matrices of solely immunoregulatory proteins (IDO, Lag3, PD-L1, PD-1) for two patients, whose outcomes are shown. **b** Clustermap of patient's immunoregulatory protein interactions features. The place at which the dendrogram was split is indicated with a red line. **c** Kaplan-Meier curves for recurrence (left) and survival (right) comparing clusters formed from solely immunoregulatory protein interactions. Log-rank test p-values are shown in the plot legends. **d** Diagram showing how the interactions of individual proteins are evaluated through ablation analysis one-at-a-time. The only

interactions included as features are the ones that involve a specific protein. The diagram gives the example of CD63. **e** Diagram showing the set of homotypic interactions. As shown by the red boxes, only the entries in the diagonal are included as features.

Multivariate analysis reveals features with independent predictive power for recurrence and survival.

To assess the prognostic importance of the features we identified, we fitted three Cox Proportional Hazard models, each of which included one of the cluster variables, two clinical variables (grade and age) and the immune architecture distinction described by Keren et al. We then measured hazard ratios and performed the log-likelihood ratio test to determine whether each cluster variable added additional prognostic information.

This analysis revealed that each cluster contained additional prognostic information when compared to age and grade for at least one clinical outcome. The protein co-expression cluster contained independent prognostic information for recurrence (HR=0.39, log-likelihood p=0.077). The functional proteins interaction cluster contained independent prognostic information for survival (HR=0.29, log-likelihood p=0.034). The immunoregulatory proteins interaction cluster contained independent prognostic information for both recurrence (HR=0.27, log-likelihood p=0.018) and survival (HR=0.36, log-likelihood p=0.052).

These results suggest that our computational pipeline was able to extract features that could provide additional clinical relevance when compared when clinical variables and previously determined characteristics of the TIME. Next, we assessed the relative predictive ability between each of the cluster variables. To do this, we fit random forests with five predictors: the three cluster variables and the two clinical variables. We then measured variable importance by calculating Shapley values[43] and overall goodness-of-fit using Harrell's c-index[44].

The random forest analysis corroborated our results from the multivariate Cox regression analysis. The immunoregulatory protein interactions cluster was the most predictive feature for recurrence (Figure 5a), and the functional protein interactions cluster was the most predictive feature for survival (Figure 5b). These features were more important than tumor grade and age. The c-index for the recurrence model was 0.718, and the c-index for the survival model was 0.731.

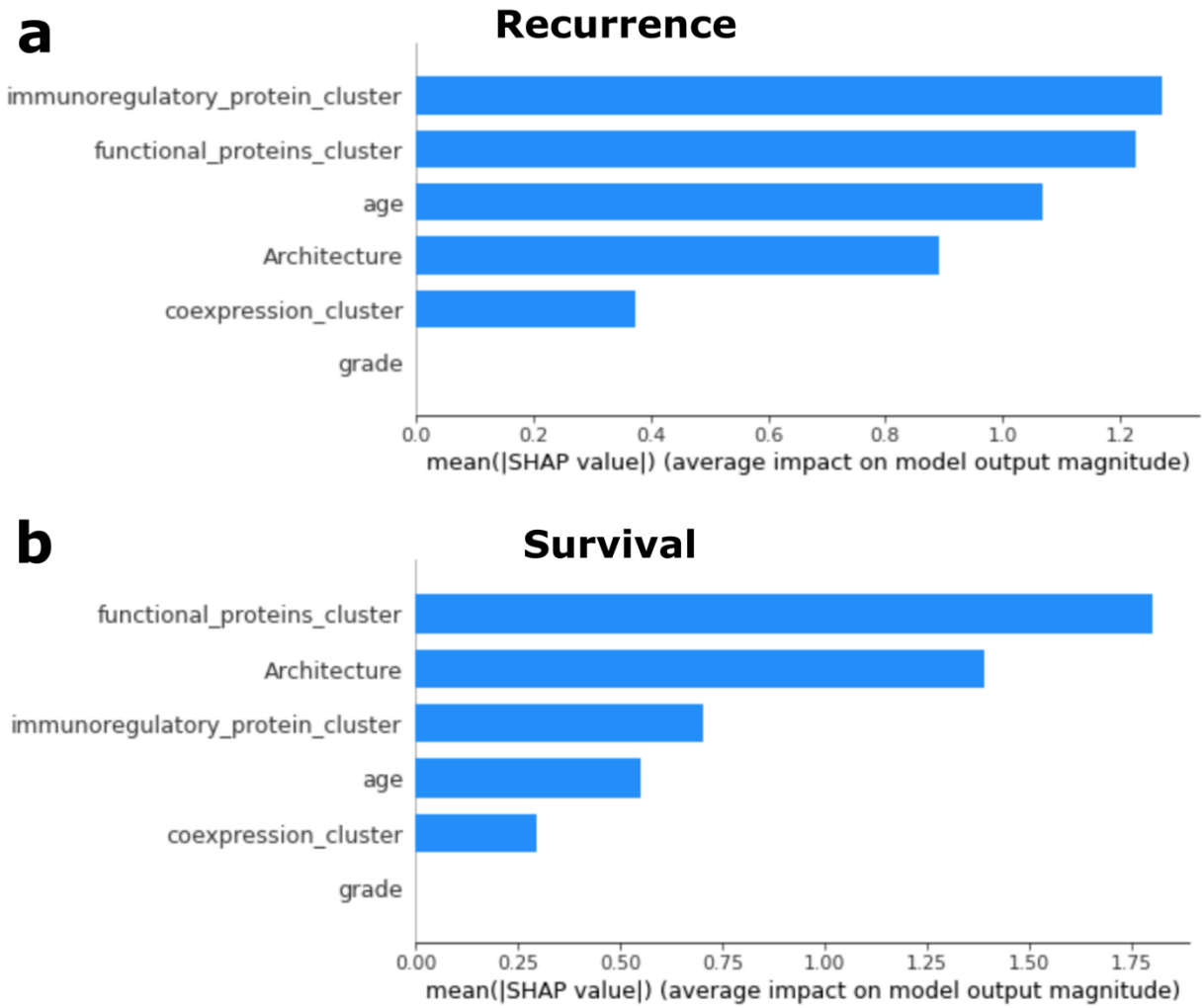


Figure 5. Random forest variable importance. a Bar plot showing the mean Shap value for each variable in a random forest predicting recurrence. **b** Bar plot showing the mean Shap value for each variable in a random forest predicting survival.

Discussion

Triple-negative breast cancer (TNBC) is the most aggressive breast cancer subtype, with a higher risk of recurrence and lower probability of survival. It lacks clinically approved biomarkers for patient risk stratification, making treatment planning and management difficult. Furthermore, there is a need to identify molecules for novel immunotherapies. In this study, we improve understanding of TNBC by analyzing the tumor-immune microenvironment (TIME) using multiplexed ion-beam imaging (MIBI).

This study makes significant contributions to the study of TNBC in three ways. First, we identify novel predictors of recurrence and overall survival in TNBC, demonstrating that the information contained within cellular interactions and protein co-expressions can aid patient stratification and therapeutic design, as proven through evaluation of patient groups, statistical tests, and predictive

modeling. Second, we demonstrate that the immune composition of the TIME does not always hold predictive power, and should therefore be examined with caution. Third, we present a computational pipeline for the robust interrogation of the TIME that produces interpretable and conclusive results, making it viable in a clinical setting.

Our primary aim was to understand cell-to-cell interactions in the TIME for the purpose of patient risk stratification and treatment management. Our findings show that profiling cell-to-cell interactions involving functional proteins is a technique for prognosis. The two most important interaction pairs were CD45RO + Beta Catenin and CD45RO + HLA-DR, a finding corroborates underlying biology. CD45RO marks memory T cells, which have been shown to mediate anti-tumor immunity[33,34]. Beta Catenin is expressed on tumor cells primarily[45,46], so its interaction with CD45Ro evidences the anti-tumor actions of memory T cells. HLA-DR is expressed on antigen-presenting cells[47], so its interaction with CD45RO+ cells evidences coordination between different immune cells to suppress tumor growth. Further, we found that "homotypic" interactions – interactions between two cells involving the same protein – hold predictive power. This indicates a coordinated immune response characterized by the localized enrichment of functional proteins. Multivariate analysis revealed that the interactions of functional proteins contained independent prognostic information for survival, even when compared to clinical variables like tumor grade and age.

Our methods differ from the previous analysis of these data in several ways. Keren et al calculated interaction matrices by defining a distance of 39 micrometers to establish adjacent cells[22]; however, the features within these interaction matrices did not result in patient clusters that differed significantly with respect to clinical outcome. This may suggest that using a set distance for adjacency is of insufficient spatial resolution to differentiate microenvironments. Our analysis also used a much lower threshold for cell protein positivity, reasoned by the expression distributions shown in Figure 2b. This lower threshold may have improved the detection of important interactions. Voronoi diagrams and Delaunay triangulation have been used previously to define and examine cellular neighborhoods in colorectal cancer[28,48]. In contrast, we use Voronoi diagrams to examine protein expression in pairwise cellular interactions, rather than larger neighborhoods. We then use these pairwise cellular interactions to explain higher levels of abstraction, building interaction matrices to summarize the TIME overall.

We further analyzed four immunoregulatory proteins – IDO, Lag3, PD-1, and PD-L1 – which are all currently in consideration as immunotherapy targets[9,38–42]. We found that the expression profiles of these proteins were prognostically relevant, suggesting that these proteins play a large role in modulating tumor progression. A host of literature has described the importance of these proteins in TIME processes[9,29,38,49], but only a small subset of such literature examines them in the context of paired cellular interactions. Interestingly, the individual expression levels of these proteins were not prognostically relevant; after Benjamini-Hochberg adjustment, none of the proteins had expression levels significantly associated with recurrence or survival. However, the cluster variable formed from their interactions was highly predictive of recurrence, as shown by multivariate analysis. It contained independent predictive power when compared to tumor grade, age, and the architecture distinction determined by Keren et al.

We found that the co-expression profile of functional proteins in patients' cells is a predictor of recurrence and survival, and that it contains independent prognostic information for recurrence. The four most important co-expressions were CD45RO + H3K27me3, CD45RO + H3K9ac, CD45RO + HLA Class 1, and HLA-DR + IDO. These results point to highly specific cellular phenotypes, a trademark of a complex TIME[22,50,51]. Our computational pipeline presents an efficient, interpretable way to calculate co-expressions and use them to predict recurrence and survival.

Our methodology allowed for analysis of the cell types present in the TIME as a whole, providing a macro-level view of immune coordination. Our findings indicate that caution should be exerted when using immune composition as a biomarker in clinical settings. After Benjamini-Hochberg adjustment, there were no cell types with significant prognostic value. This includes the monocyte/neutrophil cell type[52], the dendritic cell/monocyte cell type[53], natural killer cells[54,55], CD8+ T cells[24,56–60], macrophages[61], B cells[62], CD4+ T cells[63,64], and CD3+ T cells[65,66].

The subcellular resolution achieved by MIBI allowed us to quantify the expression of individual molecules on a single-cell basis. After Benjamini-Hochberg adjustment, there were no proteins significantly associated with outcome. This suggests that adding spatial context to the TIME is necessary to reveal hidden prognostic information.

A limitation of our work is that our results are derived from a sample of only 38 TNBC patients that were treated at Stanford hospital from 2002 to 2015 – future work is needed to validate these results on a larger cohort of patients. In addition, this study was retrospective and performed with patients at a single institution. Our cell type classifications were found computationally, derived only from the expression of molecules that were a part of our chosen assay – future work should repeat this analysis using other biologically relevant molecules.

Nonetheless, this study presents a computational pipeline for the robust interrogation of multiple features of the TIME. Our methods produce interpretable results, which may make them beneficial in therapeutic design[67]. Furthermore, they can be applied to other cancer types, as they are generalizable to any MIBI scan.

Methods

Patient Population and Dataset

Our study examined 38 TNBC patients who were treated at Stanford Hospital from 2002-2015, a subset of the cohort examined by Keren et al[22]. Although the original cohort contained 41 TNBC patients, 3 of the patients were unusable for our analysis. Patients 22 and 38 lacked recurrence outcomes, and Patient 30's images were corrupted. These patients had no special type, with estrogen receptor and progesterone receptor positivity less than 1% and HER2 unamplified. 1mm cores were taken from each patient and H&E stained. All samples were then stained with an antibody mix and scanned using MIBI-TOF. A computational pipeline converted the output of MIBI-TOF into images.

The dataset included two separate sets of 2048 x 2048 pixel images, representing a region of 800^2 square micrometers. The first set of images are 44-channel TIFFs that represent protein expression levels, where each patient has one TIFF. Each channel in the TIFF corresponds to one of the 44 molecules profiled in the study. Of the 44 proteins, 36 were biological macromolecules, such as double-stranded DNA or IDO, and 8 were elemental reporters. Each pixel in the image has a value representing the expression of the protein in that location. The second image set contained 38 grayscale segmentations of cells in the patient's sample. Protein expression data was used to classify segmented cells into appropriate cell types (CD8+ T lymphocyte, B lymphocyte, endothelial, etc.). Patient data regarding age, tumor grade, stage, and recurrence and survival outcomes were also gathered.

Analysis of Cell Prevalence

We examined whether the cellular composition of the TIME was associated with recurrence and survival. We quantified the number of cells of each cell type in each patient. To isolate specific cell types at a time, we created binary masks of each grayscale value to isolate each cell type. Then, we found the number of connected components in each mask, which gave me the number of cells of each type. After noticing large variation in the total number of cells per patient, we divided each patient's cell type count by the total number of cells in their TIME to control for this lurking variable.

Univariate Cox regression was then performed for each cell type to assess its association with recurrence and survival. Regression coefficient p-values were adjusted using the Benjamini-Hochberg method[31].

Single-Cell Protein Expression

We examined whether the expression levels of functional proteins within patients' TIME was predictive of recurrence and survival. For this analysis, we analyzed functional proteins, which modulate the activity of the cells in the TIME. These proteins stand in contrast to proteins used for lineage assignment; their expression is related to the functional state of the cell.

We labelled connected components and created a binary mask of each component to isolate the space taken up by each cell. We then applied this mask to the MIBI protein expression images, summing the value in each channel of the TIFF for each pixel in the mask. This creates a 44-length vector of protein expression per cell. We realized that per-cell expression levels are dependent on cell size, so we divided the 44-length expression vector by the size (in pixels) of the cell. This leaves a 44-length vector representing average per-pixel protein expression for a certain cell.

We calculated protein positivity thresholds from the expression levels of the image background, which lacks cells and therefore can act as a negative control. We calculated total protein expression in all background pixels in all patients and divided these values by the total number of background pixels across all patients (~67,000,000 pixels). We used each protein's threshold value to determine whether a cell was positive for a certain protein. Then, we counted the number of cells in each patient that was positive for each protein and divided by the total number of cells in the patient. The resulting number described, "What proportion of cells in this patient are positive for this protein?" For example, 100% of a patient's cells would be positive for DNA, but only 30% would be positive for PD-1. Univariate Cox regression was used to determine

association between protein expression proportions and clinical outcomes. Coefficient p-values were adjusted using the Benjamini-Hochberg method.

Individual Cell Co-expression

The co-expression of proteins in a single cell reveals functional status and immune coordination. We assessed the association between co-expression of functional proteins and recurrence and survival. We had previously measured single-cell protein expression and determined a protein positivity threshold, information was used in this analysis. We defined co-expression as an instance in which an individual cell is positive for a pair of proteins. For example, if a particular cell is positive for IDO, Lag3, and PD-1, it would have 3 instances of co-expression: IDO/Lag3, IDO/PD-1, and Lag3/PD-1. We constructed an 18x18 co-expression matrix for each patient to summarize the number of cells that co-expressed each pair of proteins. Because these matrices were symmetrical (a co-expression of IDO/Lag3 is the same as a co-expression of Lag3/IDO), we divided the matrix in two across the diagonal and flattened the top half to create a feature vector for each patient. To eliminate lurking variables, we standard scaled features across patients. We then performed hierarchical clustering to segment patients according to these features.

Silhouette analysis[32] revealed that choosing two clusters would lead to the optimal segmentation, so we cut the dendrogram into two distinct clusters and compared the two groups using the log-rank test. Then, to assess the importance of individual co-expression features, we fit a random forest with all of the co-expression pairs as predictors and the cluster assignment as the response. We assessed variable importance using mean decrease in Gini index.

Voronoi Tessellation

Analyzing cell-to-cell interactions requires a method of defining cell adjacencies. We used Voronoi diagrams to model cellular adjacencies within the TIME. Voronoi tessellation divides a planar space into a number of regions such that each point in the plane has its own region in the tessellation[37]. The sides of each Voronoi polygon are constructed to bisect two input points. Therefore, each line segment in the Voronoi tessellation represents the borders between two input points. Voronoi diagrams have been applied to single-cell imaging technology in the past, specifically for visualizing the spatial organization of colorectal cancer cells[28,36]. Due to the geometry of the Voronoi tessellation algorithm, polygons will only border their immediate neighbors.

We labelled connected components from the cell segmentation images to find each cell's centroid, which was then used to create Voronoi diagrams for each centroid. Therefore, every cell in the original cell segmentation images has a corresponding Voronoi diagram. We considered cells with bordering Voronoi regions to be adjacent, and therefore interacting. This created a reliable foundation for upstream analysis.

Cell-to-cell Interaction Analysis

We used the borders created by Voronoi diagrams to iterate over all cell adjacencies in each MIBI image. Each adjacency represented an individual interaction between two cells. We constructed two lists: List 1 contained the names of the proteins that Cell 1 was positive for, and List 2 contained the names of the proteins that Cell 2 was positive for. We took the Cartesian

product of the two lists to find all of the combinations of proteins present in this interaction. For example, if Cell 1 was positive for PD-L1 and Lag3, and Cell 2 was positive for PD-1 and IDO, then we would count the following interactions: PD-L1 + PD-1, PD-L1 + IDO, Lag3 + PD-1, and Lag3 + IDO. These interactions would be tallied in the overall interaction matrix for each patient, in which row A and column B represents the number of times a cell positive for protein A was adjacent to a cell positive for protein B. For this analysis, we only counted interactions between functional proteins, excluding proteins used for lineage assignment.

We selected the top half of the symmetric matrix and flattened it to create feature vectors for each patient. Hierarchical clustering was performed and the dendrogram was cut to produce two clusters based on silhouette score analysis. These two clusters were compared using Kaplan-Meier curves, the log-rank test, and Cox regression. To assess the importance of individual interactions, we fit a random forest with all interactions as predictors and the cluster assignment as the response. We measured variable importance using mean decrease in Gini index.

Multivariate Analysis

To assess whether the features identified by our computational pipeline contained independent prognostic information, we performed multivariate Cox regression. We fit three Cox Proportional Hazard models, each of which contained one of the three cluster variables identified by our study (protein co-expression, functional protein interactions, immunoregulatory protein interactions), two clinical variables (tumor grade and age), and the immune architecture distinction determined by Keren et al. We found the hazard ratios of each cluster variable and used the log-likelihood ratio test to determine whether the variables contained additional prognostic information.

We also fit random forests to measure relative variable importance. We included all six variables in the random forests and calculated Shapley values to get stable estimates of variable importance. We measured overall goodness-of-fit using Harrell's c-index.

Statistical Analysis

Primary statistical analyses were performed using Python (v3.7.3, Python Software Foundation, https://www.python.org/) with the lifelines (v0.24.0), scipy (v1.4.1), seaborn (v0.10.1), and pysurvival (v0.1.2) packages.

We chose a significance level of 0.1 for statistical analysis. Our reasoning was threefold. First, the sample size of the study reduces the power of statistical significance tests and increases the probability of Type II errors. Secondly, collinearity between predictors, such as lymphocyte count, functional protein expression, and interactions, reduces the power of significance tests. Instead of eliminating certain predictors, which would sacrifice interpretability and accuracy, we compensated by choosing a higher significance level. Finally, this study is incredibly exploratory—due to a lack of research on cell-to-cell interactions in the TNBC TIME, we felt introducing numerous possible predictors of clinical outcomes was more valuable. In analyses involving multiple comparisons, we adjusted p-values using the Benjamini-Hochberg method.

Data Availability

All data used in this study can be found at https://mibi-share.ionpath.com/. It comes with an easy-to-use interface that allows for easy examination of the data.

References

1. Aysola, K. *et al.* Triple Negative Breast Cancer - An Overview. *Hereditary Genet.* **2013,** (2013).
2. Anders, C. & Carey, L. A. Understanding and treating triple-negative breast cancer. *Oncology* **22,** 1233–9; discussion 1239 (2008).
3. Omarini, C. *et al.* Neoadjuvant treatments in triple-negative breast cancer patients: where we are now and where we are going. *Cancer Manag Res* **10,** 91–103 (2018).
4. Carey, L. A. *et al.* Race, breast cancer subtypes, and survival in the Carolina Breast Cancer Study. *JAMA* **295,** 2492–2502 (2006).
5. Blows, F. M. *et al.* Subtyping of breast cancer by immunohistochemistry to investigate a relationship between subtype and short and long term survival: a collaborative analysis of data for 10,159 cases from 12 studies. *PLoS Med.* **7,** e1000279 (2010).
6. Dent, R. *et al.* Triple-negative breast cancer: clinical features and patterns of recurrence. *Clin. Cancer Res.* **13,** 4429–4434 (2007).
7. Anders, C. K., Abramson, V., Tan, T. & Dent, R. The Evolution of Triple-Negative Breast Cancer: From Biology to Novel Therapeutics. *Am. Soc. Clin. Oncol. Educ. Book* **35,** 34–42 (2016).
8. Criscitiello, C., Azim, H. A., Schouten, P. C., Linn, S. C. & Sotiriou, C. Understanding the biology of triple-negative breast cancer. *Ann. Oncol.* **23 Suppl 6,** vi13-8 (2012).
9. Maleki Vareki, S., Garrigós, C. & Duran, I. Biomarkers of response to PD-1/PD-L1 inhibition. *Crit. Rev. Oncol. Hematol.* **116,** 116–124 (2017).
10. Pitt, J. M. *et al.* Resistance Mechanisms to Immune-Checkpoint Blockade in Cancer: Tumor-Intrinsic and -Extrinsic Factors. *Immunity* **44,** 1255–1269 (2016).
11. Bao, C. *et al.* Exploring specific prognostic biomarkers in triple-negative breast cancer. *Cell Death Dis.* **10,** 807 (2019).
12. Ilie, S. M., Bacinschi, X. E., Botnariuc, I. & Anghel, R. M. Potential clinically useful prognostic biomarkers in triple-negative breast cancer: preliminary results of a retrospective analysis. *Breast Cancer (Dove Med Press)* **10,** 177–194 (2018).
13. Stovgaard, E. S., Nielsen, D., Hogdall, E. & Balslev, E. Triple negative breast cancer - prognostic role of immune-related factors: a systematic review. *Acta Oncol.* **57,** 74–82 (2018).
14. Yue, Y. *et al.* Stratification of Prognosis of Triple-Negative Breast Cancer Patients Using Combinatorial Biomarkers. *PLoS One* **11,** e0149661 (2016).
15. Deng, L. *et al.* Immune Profiles of Tumor Microenvironment and Clinical Prognosis among Women with Triple-Negative Breast Cancer. *Cancer Epidemiol. Biomarkers Prev.* **28,** 1977–1985 (2019).
16. Quail, D. F. & Joyce, J. A. Microenvironmental regulation of tumor progression and metastasis. *Nat. Med.* **19,** 1423–1437 (2013).
17. Dannenfelser, R. *et al.* Data-driven analysis of immune infiltrate in a large cohort of breast cancer and its association with disease progression, ER activity, and genomic complexity. *Oncotarget* **8,** 57121–57133 (2017).

18. Aran, D., Hu, Z. & Butte, A. J. xCell: digitally portraying the tissue cellular heterogeneity landscape. *Genome Biol.* **18,** 220 (2017).
19. Ali, H. R., Chlon, L., Pharoah, P. D. P., Markowetz, F. & Caldas, C. Patterns of Immune Infiltration in Breast Cancer and Their Clinical Implications: A Gene-Expression-Based Retrospective Study. *PLoS Med.* **13,** e1002194 (2016).
20. Gruosso, T. *et al.* Spatially distinct tumor immune microenvironments stratify triple-negative breast cancers. *J. Clin. Invest.* **129,** 1785–1800 (2019).
21. Giesen, C. *et al.* Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry. *Nat. Methods* **11,** 417–422 (2014).
22. Keren, L. *et al.* A Structured Tumor-Immune Microenvironment in Triple Negative Breast Cancer Revealed by Multiplexed Ion Beam Imaging. *Cell* **174,** 1373–1387.e19 (2018).
23. Allam, M., Cai, S. & Coskun, A. F. Multiplex bioimaging of single-cell spatial profiles for precision cancer diagnostics and therapeutics. *NPJ Precis. Oncol.* **4,** 11 (2020).
24. Nearchou, I. P. *et al.* Spatial immune profiling of the colorectal tumor microenvironment predicts good outcome in stage II patients. *npj Digital Med.* **3,** 71 (2020).
25. Angelo, M. *et al.* Multiplexed ion beam imaging of human breast tumors. *Nat. Med.* **20,** 436–442 (2014).
26. Chevrier, S. *et al.* An immune atlas of clear cell renal cell carcinoma. *Cell* **169,** 736–749.e18 (2017).
27. Liu, C. C., Steen, C. B. & Newman, A. M. Computational approaches for characterizing the tumor immune microenvironment. *Immunology* **158,** 70–84 (2019).
28. Schürch, C. M. *et al.* Coordinated cellular neighborhoods orchestrate antitumoral immunity at the colorectal cancer invasive front. *BioRxiv* (2019). doi:10.1101/743989
29. Dieci, M. V. *et al.* Integration of tumour infiltrating lymphocytes, programmed cell-death ligand-1, CD8 and FOXP3 in prognostic models for triple-negative breast cancer: Analysis of 244 stage I-III patients treated with standard therapy. *Eur. J. Cancer* **136,** 7–15 (2020).
30. Stewart, R. L., Matynia, A. P., Factor, R. E. & Varley, K. E. Spatially-resolved quantification of proteins in triple negative breast cancers reveals differences in the immune microenvironment associated with prognosis. *Sci. Rep.* **10,** 6598 (2020).
31. Benjamini, Y. H. & Hochberg, Y. Controlling The False Discovery Rate - A Practical And Powerful Approach To Multiple Testing. *Journal of the Royal Statistical Society. Series B: Methodological* (1995).
32. Rousseeuw, P. J. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics* **20,** 53–65 (1987).
33. Yajima, R. *et al.* Tumor-infiltrating CD45RO(+) memory cells are associated with a favorable prognosis breast cancer. *Breast Cancer* **23,** 668–674 (2016).
34. Hu, G. & Wang, S. Tumor-infiltrating CD45RO+ Memory T Lymphocytes Predict Favorable Clinical Outcome in Solid Tumors. *Sci. Rep.* **7,** 10376 (2017).
35. Ladányi, A. *et al.* Role of the anatomic site in the association of HLA class I antigen expression level in metastases with clinical response to ipilimumab therapy in patients with melanoma. *J. Immunother. Cancer* **8,** (2020).
36. Chen, Z., Soifer, I., Hilton, H., Keren, L. & Jojic, V. Modeling Multiplexed Images with Spatial-LDA Reveals Novel Tissue Microenvironments. *J. Comput. Biol.* **27,** 1204–1218 (2020).
37. Pokojski, W. & Pokojska, P. Voronoi diagrams – inventor, method, applications. *Polish Cartographical Review* **50,** 141–150 (2018).

38. Zhu, H. *et al.* PD-1/PD-L1 counterattack alliance: multiple strategies for treating triple-negative breast cancer. *Drug Discov. Today* (2020). doi:10.1016/j.drudis.2020.07.006
39. Le Naour, J., Galluzzi, L., Zitvogel, L., Kroemer, G. & Vacchelli, E. Trial watch: IDO inhibitors in cancer therapy. *Oncoimmunology* **9,** 1777625 (2020).
40. Sun, J.-Y. *et al.* Resistance to PD-1/PD-L1 blockade cancer immunotherapy: mechanisms, predictive factors, and future perspectives. *Biomark Res* **8,** 35 (2020).
41. Fröhlich, A. *et al.* Molecular, clinicopathological, and immune correlates of LAG3 promoter DNA methylation in melanoma. *EBioMedicine* **59,** 102962 (2020).
42. Ahmed, F. S. *et al.* PD-L1 Protein Expression on Both Tumor Cells and Macrophages are Associated with Response to Neoadjuvant Durvalumab with Chemotherapy in Triple-negative Breast Cancer. *Clin. Cancer Res.* (2020). doi:10.1158/1078-0432.CCR-20-1303
43. Lundberg, S. M. & Lee, S.-I. A Unified Approach to Interpreting Model Predictions. *Advances in Neural Information Processing Systems* (2017).
44. Harrell, F. E. Evaluating the yield of medical tests. *JAMA* **247,** 2543 (1982).
45. Xu, J., Prosperi, J. R., Choudhury, N., Olopade, O. I. & Goss, K. H. β-Catenin is required for the tumorigenic behavior of triple-negative breast cancer cells. *PLoS One* **10,** e0117097 (2015).
46. Pohl, S.-G. *et al.* Wnt signaling in triple-negative breast cancer. *Oncogenesis* **6,** e310 (2017).
47. Park, I. A. *et al.* Expression of the MHC class II in triple-negative breast cancer is associated with tumor-infiltrating lymphocytes and interferon signaling. *PLoS One* **12,** e0182786 (2017).
48. Goltsev, Y. *et al.* Deep Profiling of Mouse Splenic Architecture with CODEX Multiplexed Imaging. *Cell* **174,** 968–981.e15 (2018).
49. Santoni, M. *et al.* Triple negative breast cancer: Key role of Tumor-Associated Macrophages in regulating the activity of anti-PD-1/PD-L1 agents. *Biochim. Biophys. Acta Rev. Cancer* **1869,** 78–84 (2018).
50. Chen, Y.-P. *et al.* Single-cell transcriptomics reveals regulators underlying immune cell diversity and immune subtypes associated with prognosis in nasopharyngeal carcinoma. *Cell Res.* (2020). doi:10.1038/s41422-020-0374-x
51. Romero-Cordoba, S. *et al.* Decoding Immune Heterogeneity of Triple Negative Breast Cancer and Its Association with Systemic Inflammation. *Cancers (Basel)* **11,** (2019).
52. Hagerling, C. *et al.* Immune effector monocyte-neutrophil cooperation induced by the primary tumor prevents metastatic progression of breast cancer. *Proc. Natl. Acad. Sci. USA* **116,** 21704–21714 (2019).
53. Michea, P. *et al.* Adjustment of dendritic cells to the breast-cancer microenvironment is subset specific. *Nat. Immunol.* **19,** 885–897 (2018).
54. Hu, W., Wang, G., Huang, D., Sui, M. & Xu, Y. Cancer immunotherapy based on natural killer cells: current progress and new opportunities. *Front. Immunol.* **10,** 1205 (2019).
55. Wu, S.-Y., Fu, T., Jiang, Y.-Z. & Shao, Z.-M. Natural killer cells in cancer biology and therapy. *Mol. Cancer* **19,** 120 (2020).
56. Baker, K. *et al.* Prognostic significance of CD8+ T lymphocytes in breast cancer depends upon both oestrogen receptor status and histological grade. *Histopathology* **58,** 1107–1116 (2011).
57. Vihervuori, H. *et al.* Tumor-infiltrating lymphocytes and CD8+ T cells predict survival of triple-negative breast cancer. *J. Cancer Res. Clin. Oncol.* **145,** 3105–3114 (2019).
58. Ali, H. R. *et al.* Association between CD8+ T-cell infiltration and breast cancer survival in

12,439 patients. *Ann. Oncol.* **25,** 1536–1543 (2014).
59. McIntire, P. J. *et al.* Hot Spot and Whole-Tumor Enumeration of CD8+ Tumor-Infiltrating Lymphocytes Utilizing Digital Image Analysis Is Prognostic in Triple-Negative Breast Cancer. *Clin Breast Cancer* **18,** 451–458.e1 (2018).
60. Bindea, G. *et al.* Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. *Immunity* **39,** 782–795 (2013).
61. Sousa, S. *et al.* Human breast cancer cells educate macrophages toward the M2 activation status. *Breast Cancer Res.* **17,** 101 (2015).
62. Stagg, J. & Allard, B. Immunotherapeutic approaches in triple-negative breast cancer: latest research and clinical prospects. *Ther Adv Med Oncol* **5,** 169–181 (2013).
63. Tay, R. E., Richardson, E. K. & Toh, H. C. Revisiting the role of CD4+ T cells in cancer immunotherapy-new insights into old paradigms. *Cancer Gene Ther* (2020). doi:10.1038/s41417-020-0183-x
64. Gao, G., Wang, Z., Qu, X. & Zhang, Z. Prognostic value of tumor-infiltrating lymphocytes in patients with triple-negative breast cancer: a systematic review and meta-analysis. *BMC Cancer* **20,** 179 (2020).
65. Tsiatas, M. *et al.* Evaluation of the prognostic value of CD3, CD8, and FOXP3 mRNA expression in early-stage breast cancer patients treated with anthracycline-based adjuvant chemotherapy. *Cancer Med* **7,** 5066–5082 (2018).
66. Rathore, A. S. *et al.* Presence of CD3+ tumor infiltrating lymphocytes is significantly associated with good prognosis in infiltrating ductal carcinoma of breast. *Indian J. Cancer* **50,** 239–244 (2013).
67. Vamathevan, J. *et al.* Applications of machine learning in drug discovery and development. *Nat. Rev. Drug Discov.* **18,** 463–477 (2019).

Acknowledgements

This work was funded by the Stanford Departments of Pathology and Biomedical Data Science through a Stanford Clinical Data Science Fellowship to R.Y.

Author Information

Affiliations

Department of Biomedical Data Science, Stanford University, Stanford, CA 94305, USA
Aalok Patwa, Rikiya Yamashita, & Daniel L. Rubin.

Center for Artificial Intelligence in Medicine and Imaging, Stanford University, Stanford, CA 94305, USA
Rikiya Yamashita, Jin Long & Daniel L. Rubin.

Department of Pathology, Stanford University, Stanford, CA 94305, USA
Michael Angelo

Department of Molecular Cell Biology, Weizmann Institute of Science, Rehovot, Israel
Leeat Keren

Archbishop Mitty High School, San Jose, CA 95129, USA

Aalok Patwa

Contributions

A.P. conducted the analysis and wrote the manuscript. R.Y. contributed to the design of the analysis substantially, providing constant feedback and interpretation of results. L.K. provided data, suggested analyses and evaluated results. J.L. suggested and reviewed analysis and results. M.A. provided data and interpreted results. D.R supervised the project. All authors participated in the critical revision and approval of the manuscript.

Corresponding Author

Direct correspondence to Daniel Rubin: dlrubin@stanford.edu.

Ethics declarations

The authors declare no competing interests.

Supplementary Materials

Functional Proteins	Beta Catenin, CD138, CD45RO, CD63, FoxP3, H3K27me3, H3K9ac, HLA-DR, HLA Class 1, IDO, Keratin17, Keratin6, Ki67, Lag3, p53, PD-L1, PD-1, Phospho-S6
Lineage Proteins	CD11b, CD11c, CD16, CD20, CD209, CD3, CD31, CD4, CD45, CD56, CD68, CD8, dsDNA, EGFR, MPO, Pan-Keratin, SMA, Vimentin

Table S1. Delineation between functional and lineage proteins. Lineage proteins are those that are solely used for the identification of cell type, whereas functional proteins provide information related to proliferation or metabolic activity[22].

Protein	Coefficient	Hazard Ratio	Coefficient P	BH-Corrected P
CD45RO	-0.019	0.981	0.051	0.311
PD1	-0.112	0.894	0.063	0.311
CD138	0.018	1.018	0.069	0.311
IDO	-0.067	0.935	0.084	0.311
HLA_Class_1	-0.053	0.948	0.104	0.311
HLA-DR	-0.013	0.987	0.12	0.311
Lag3	-0.232	0.793	0.121	0.311
p53	-0.023	0.978	0.212	0.477
H3K9ac	0.086	1.09	0.331	0.661
H3K27me3	0.724	2.063	0.367	0.661
Keratin17	0.009	1.009	0.432	0.684
Beta catenin	-0.008	0.992	0.476	0.684
phospho-S6	0.012	1.012	0.494	0.684
Keratin6	0.005	1.005	0.651	0.837
PD-L1	-0.004	0.996	0.715	0.858
CD63	0.003	1.003	0.791	0.890
Ki67	-0.007	0.993	0.846	0.892
FoxP3	-0.027	0.973	0.892	0.892

Table S2a. Protein expression Cox regression results for recurrence. This table shows the results from performing univariate Cox regression with each protein's expression as predictors and recurrence outcome as the response. There is no association between protein expression and recurrence in the cohort.

Protein	Coefficient	Hazard Ratio	Coefficient P	BH-Corrected P
Keratin6	0.025	1.025	0.034	0.374
HLA-DR	-0.018	0.982	0.045	0.374
Lag3	-0.336	0.715	0.073	0.374
Keratin17	0.02	1.021	0.096	0.374
HLA_Class_1	-0.055	0.946	0.109	0.374
H3K27me3	-0.073	0.93	0.146	0.374
IDO	-0.049	0.952	0.164	0.374
CD138	0.014	1.014	0.175	0.374
PD1	-0.069	0.934	0.187	0.374
CD45RO	-0.011	0.989	0.241	0.434
p53	-0.014	0.987	0.353	0.578
CD63	0.008	1.008	0.511	0.671
Beta catenin	-0.007	0.993	0.517	0.671
H3K9ac	0.056	1.058	0.522	0.671
PD-L1	-0.006	0.994	0.626	0.751
FoxP3	-0.085	0.918	0.688	0.774
Ki67	0.009	1.009	0.771	0.816
phospho-S6	0.001	1.001	0.928	0.928

Table S2b. Protein expression Cox regression results for survival. This table shows the results from performing univariate Cox regression with each protein's expression as predictors and survival outcome as the response. There is no association between protein expression and survival in the cohort.

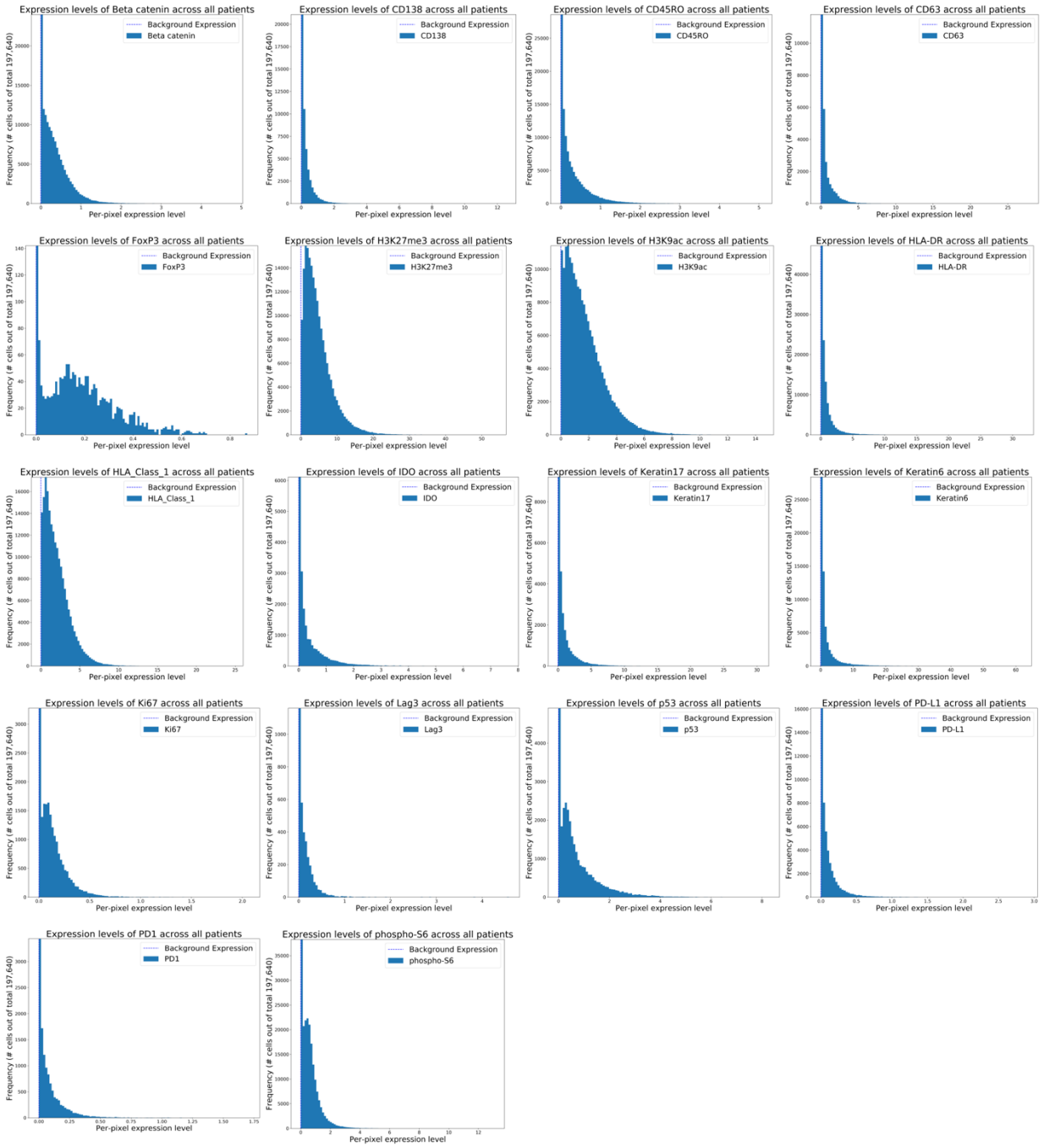


Figure S1. Histograms of expression for all functional proteins. Each of the plots shown in the figure is a histogram of the per-pixel expression levels of that protein across all of the cells in all of the patients. The y-axis has been shrunk to better show the distribution, as in reality, the values are extremely right-skewed, and the frequency of 0 expression dwarfs the rest of the values. However, it is still valuable to see the rest of the distribution, as this is where the difference in cellular phenotype is made. As such, the y-axis upper limit has been made twice the second highest frequency.

Features	Chosen Clusters	Mean Silhouette Score
Protein co-expression	2	0.42
	3	0.18
	4	0.04
	5	0.04
	6	0.02
Functional proteins interactions	2	0.38
	3	0.11
	4	0.05
	5	0.04
	6	-0.23
Immunoregulatory proteins interactions	2	0.47
	3	0.17
	4	0.06
	5	0.04
	6	-0.29

Table S3. Silhouette score analysis on hierarchical clustering. This table shows the mean silhouette score of hierarchical clustering based on the number of chosen clusters for each spatial analysis. In all cases, choosing two clusters led to the highest silhouette score.

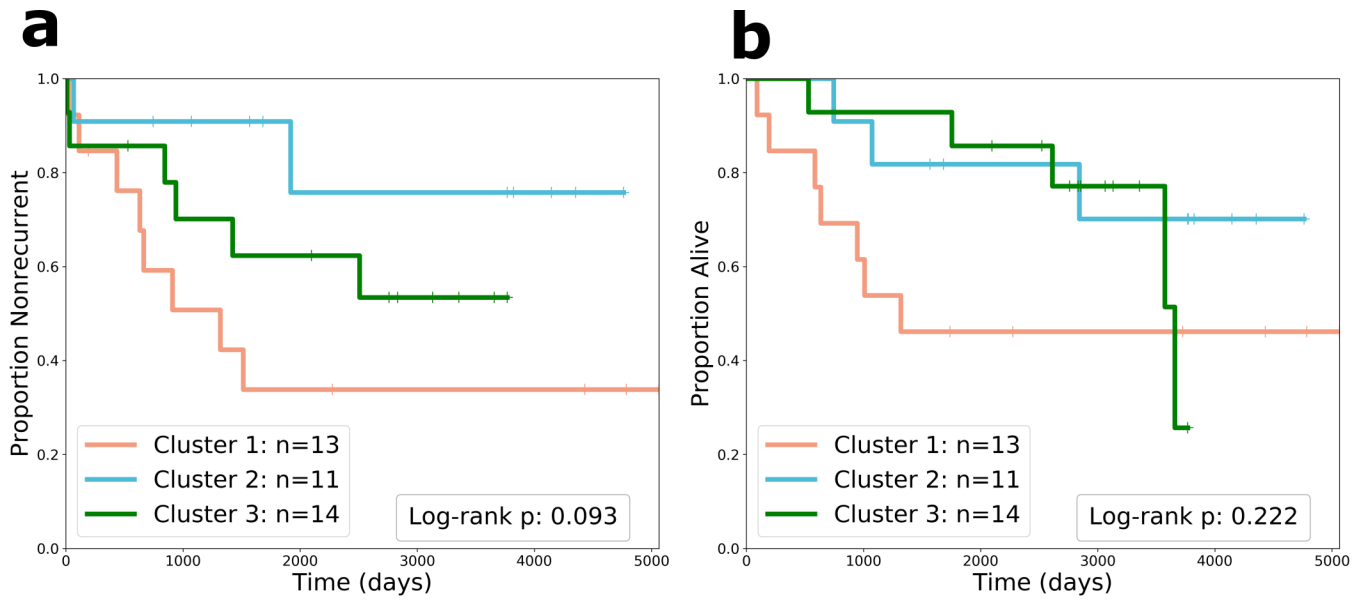


Figure S2. Kaplan-Meier curves comparing three patient clusters chosen from protein co-expression features. **a** Kaplan-Meier curve comparing recurrence across the three patient clusters. The log-rank p-value is shown. **b** Kaplan-Meier curve comparing survival across the three patient clusters. The log-rank p-value is shown.

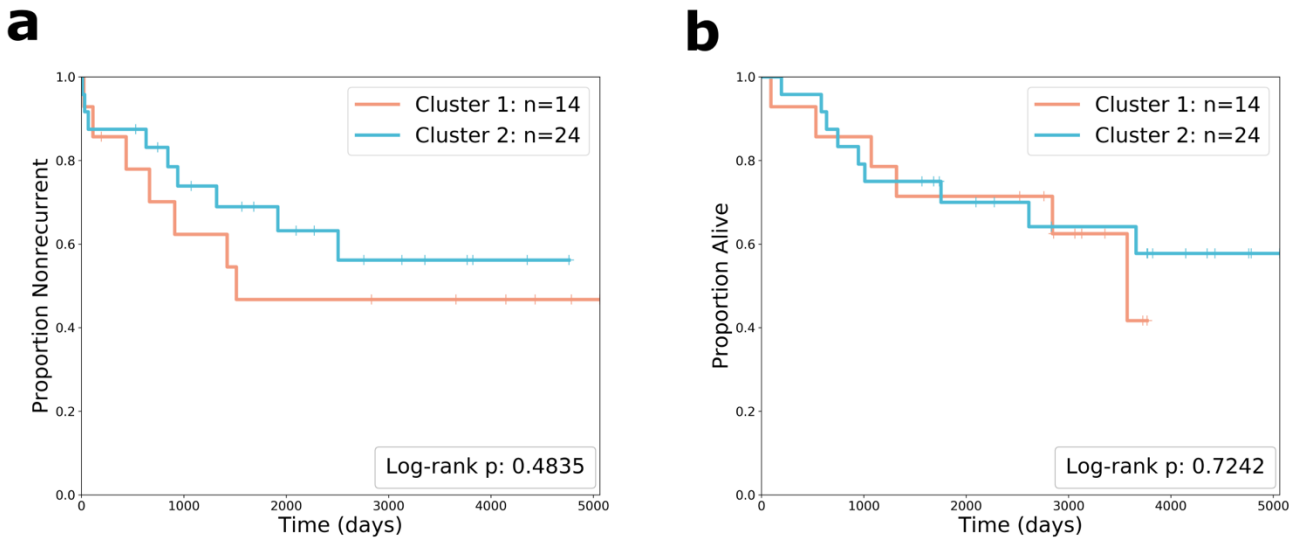


Figure S3. Kaplan-Meier curves of lineage proteins interactions features. **a** Kaplan-Meier curve comparing recurrence across patient clusters formed from lineage proteins interactions features. There is not clear divergence between the clusters. **b** Kaplan-Meier curve comparing overall survival across patient clusters formed from lineage proteins interactions features. There is no clear divergence.

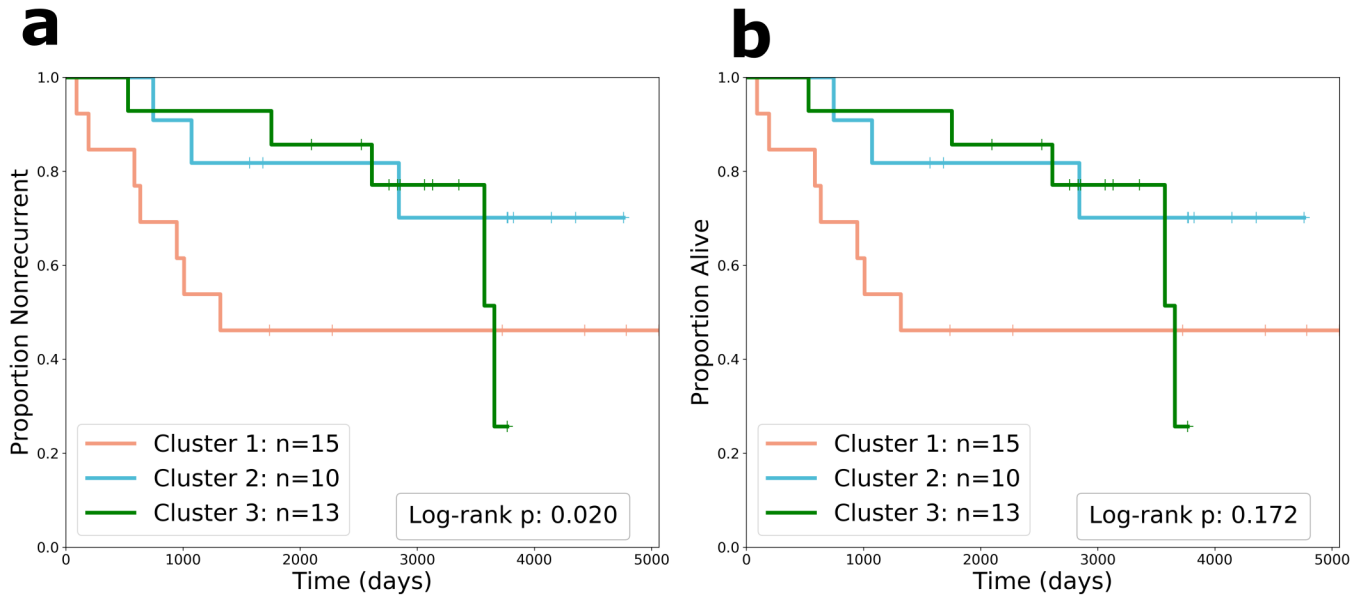
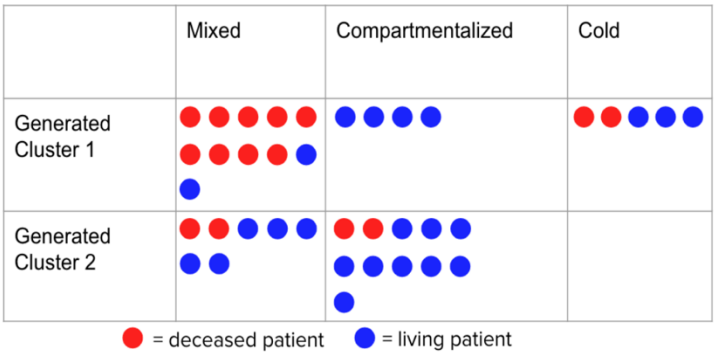


Figure S4. Kaplan-Meier curves comparing three patient clusters chosen immunoregulatory protein interaction features.. **a** Kaplan-Meier curve comparing recurrence across the three patient clusters. The log-rank p-value is shown. **b a** Kaplan-Meier curve comparing survival across the three patient clusters. The log-rank p-value is shown.

Interaction Features	Clinical Outcome	Cluster 1 n	Cluster 2 n	Log-rank p
All functional proteins	OS	20	18	0.5281
	Recurrence	20	18	0.2316
Immunoregulatory proteins	OS	2	36	0.8491
	Recurrence	2	36	0.8784

Table S4. Results of hierarchical clustering using interaction analyses performed by Keren et al[22]. The table was created by selecting two different subsets of features from the interaction matrices defined in Keren et al: all features of functional proteins interactions and features of immunoregulatory proteins. Those features were then used to cluster patients, and the clusters were compared using the log-rank test.



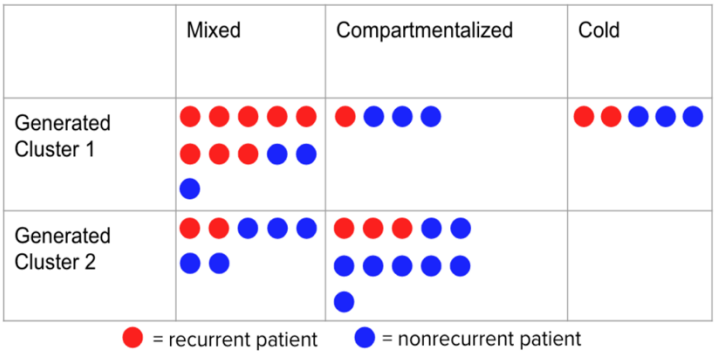


Figure S5. Drawing comparing the patients in clusters formed from interaction features to Keren et al's designation of "mixed", "compartmentalized", and "cold" architectures.