Learning tissue representation by identification of persistent local patterns in spatial omics data

Jovan Tanevski[1,2,*], Loan Vulliard[3], Felix Hartmann[3,4], and Julio Saez-Rodriguez[1,*]

[1]Institute for Computational Biomedicine, Heidelberg University and Heidelberg University Hospital, Heidelberg, Germany
[2]Department of Knowledge Technologies, Jožef Stefan Institute, Ljubljana, Slovenia
[3]Systems Immunology and Single-Cell Biology, German Cancer Research Center (DKFZ), Heidelberg, Germany
[4]German Cancer Consortium (DKTK), Heidelberg, Germany

[*]Correspondence: *{jovan.tanevski, pub.saez}@uni-heidelberg.de*

Abstract. Spatial omics data provide rich molecular and structural information about tissues, enabling novel insights into the structure-function relationship. In particular, it facilitates the analysis of the local heterogeneity of tissues and holds promise to improve patient stratification by association of finer-grained representations with clinically relevant features. Here, we introduce Kasumi, a method for the identification of spatially localized neighborhoods of intra- and intercellular relationships, persistent across samples and conditions. We learn compressed explainable representations while preserving relevant biological signals that are readily deployable for data exploration and hypothesis generation, facilitating translational tasks. We address tasks of patient stratification for disease progression and response to treatment in cancer on data coming from different spatial antibody-based multiplexed proteomics platforms. Kasumi outperforms related neighborhood analysis approaches and offers explanations at the level of cell types or directly from the measurements, of the spatial coordination and multivariate relationships underlying observed disease progression and response to treatment. We show that persistent local patterns form spatially contiguous regions of different sizes. However, the abundance of the persistent local patterns is not associated with their relative importance in downstream tasks. We show that non-abundant, localized structural and functional relationships in the tissue are strongly associated with unfavorable outcomes in disease progression and response to treatment.

1 Introduction

The connection between the microanatomy of a tissue and its function is well recognized. Histological staining, immunohistochemical and immunofluorescence panels measuring a small number of molecules are routinely used for diagnosis, tracking disease progression,

assigning patients to treatable groups, and more. The advancement of spatially resolved omics technologies offers an unprecedented insight into the structure of tissues and the molecular state of their constituents[1,2], bridging biological understanding of disease to the clinical pathology practice[3,4]. This is made possible by an increased richness of the measured data, the number and diversity of distinct molecules, and the spatial resolution of tissue observations. The availability of such data prompts a revisit of the understanding of the relationship between structure and function in tissues[5,6]. Subsequently, it motivates the development of computational approaches aimed at facilitating the discovery of novel biological insights from complex and spatially resolved data[7–9].

The computational discovery of knowledge[10] relies on data representations, i.e. transformation of the data to a new set of variables to facilitate exploration and association to biologically and clinically relevant observations. A good representation not only facilitates the efficient and effective application of computational methods but is also explainable. That is, the models built on such representation are communicable to domain experts, experimentally verifiable, and practically deployable. In the frame of analysis of spatial omics data, there are two commonly used approaches to data representation: clustering (phenotyping) and neighborhood (niche) analysis.

Clustering assigns a label to measured spatial profiles, usually corresponding to a cell (sub)type or a functional state. By spatial profile, we here refer to the abundances of measured molecules at a specific location. This location can be captured with different resolutions ranging from tens of micrometers (a small group of cells) to a micrometer (sub-cellular resolution). While the grouping of spatial units, with or without considering spatial locations[11], is performed in a data-driven manner, the discrete labels are assigned based on known cell-type markers, functional markers, or by cross-referencing existing atlases. Consequently, the output of the clustering becomes explainable and communicable. Clustering is a first step towards representing the tissue structure. The assumption behind clustering is that the majority of the relevant information about the spatial unit is captured by its profile. Therefore, the tissue can be represented by the distribution of constituent clusters.

Neighborhood (niche) identification and analysis, although related to clustering, goes one step further. Often starting from the results of the clustering, it aims at capturing a higher-order representation of the tissue by identifying conserved patterns of interactions between clusters. Recent methods for neighborhood analysis adopt different approaches. Some, such as histoCAT[12], Giotto[13], and Cellular Spatial Enrichment Analysis (CSEA)[14,15], focus on the immediate neighborhood and number of interactions between pairs of cell types or functional states. The pairs of interacting cell types are identified by calculating the significance of co-occurrence within the immediate neighborhood by comparing to a null distribution of interactions derived from cell location permutations. Another group, consisting of methods such as iNiche[16], Spatial-LDA[17], and Coordinated Cellular Neighborhoods (CCN)[18,19], identify neighborhood motifs by first representing each cell by the cell-type composition within its neighborhood and then clustering this representation again to infer higher-order motif-oriented identities. Finally, a third group of methods, such as SPACE-GM[20] and STELLAR[21], aim at supervised end-to-end GNN (graph neural network) -based learning of cell representations based on their cellular neighborhoods as defined by a cell graph[22] of the sample.

Most of the previously mentioned methods are constrained by their reliance on the initial clustering of the data as a baseline representation for further neighborhood analysis. The neighborhoods in these cases are often defined based on colocalization or correlation. In the presence of multivariate or non-linear structural patterns[23], these approaches result in the identification of redundant or spurious neighborhoods which are *post hoc* manually relabeled or joined, before being used for downstream analyses. However, neighborhoods defined by colocalization of cell types inherit the explainability of the baseline representation. The simplicity of the underlying assumptions means that they can be applied to data consisting of a small number of samples or even a single sample. On the other end, GNN approaches require a large number of labeled samples for (supervised) training to reliably produce their embedded representation, which for a given tissue and a specific condition might not be available. While GNN approaches can potentially capture complex relationships from the available data, the produced embeddings are not explainable. Starting from the spatial profiles or initial clustering, this group of approaches relies on either additional approximate interpretation or a combination of the learned embeddings with methods from the other two groups. The learned embeddings are considered as spatial profiles and therefore a baseline representation for further neighborhood analysis.

Here, we present *Kasumi*, an approach to learning a representation of tissue structures that overcomes the limitations of methods for neighborhood analysis by: (i) accounting for spatially localized patterns of multivariate, non-linear, and robust intra- and intercellular relationships; (ii) offering the flexibility to analyze neighborhoods without relying on cell-type labels; and (iii) providing an unsupervised and explainable representation of samples as a composition of neighborhoods persistent across samples.

Kasumi aims to expand the current understanding of tissue structure and its relation to condition and function by extracting knowledge available in high dimensional and spatially resolved omics data. To this end, we define our approach around the explainable multi-view framework for dissecting spatial relationships from highly multiplexed data (MISTy)[24].

Our approach to defining neighborhoods based on persistent multivariate and non-linear patterns allows for more flexibility to cross scales of organization[25]. Instead of forming neighborhoods by grouping spatial units with similar profiles or similar local compositions, we define neighborhoods as tissue patches that share robust relationships at different scales, i.e., different spatial contexts across all available samples. In other words, instead of considering the cell as an independent unit of organization of the tissue, we consider the consistent intracellular and intercellular relationships as captured by a non-linear and multivariate predictive model as a tissue representation.

Different biological processes operate at different scales and their spatial regulation can be overlooked when searching for patterns at an inappropriate resolution[26]. Kasumi offers flexibility in the definition of the type and scale of the relationships which can capture the tissue composition (organization of cell types) or function (relationship among markers of molecular processes). Furthermore, it doesn't rely on initial clustering, but can rather use observations at the level of spatial profiles or any other available data representation.

Kasumi is scalable with regard to resolution and number of samples, and it can be used on a single sample or hundreds of samples. It is easily deployable and freely available as an open-source R-package. The representation learned by Kasumi is explainable and it doesn't require the use of complex models for biologically relevant or translational downstream tasks. We demonstrate this on different cancer-related data by addressing tasks of patient stratification related to expected disease progression and response to treatment.

2 Results

2.1 Kasumi identifies persistent local patterns

Kasumi is an unsupervised multi-view modeling approach to learning representations of tissues based on localized relationship patterns in spatial omics data (Figure 1).

Kasumi takes as input a dataset of spatially resolved data consisting of one or multiple samples from the same tissue under multiple conditions. The input data is organized in the form of a view composition. In particular, each spatial unit is described by its identity or functional state, forming an intrinsic view (intraview), while additional views capture the relations between each unit's properties in different spatial contexts, or represent complementary aspects of the data. While the view composition can be defined flexibly and can be tailored to existing hypotheses or processes of interest, we focus here on a single-cell resolution and a simple two-view composition. To capture the structure of the tissue, the intraview describes each cell by the identity of the cluster it belongs to. The intraview is complemented by a broader tissue view capturing the distribution of cell identities or marker observations in a radius of interest around each cell (see Methods). The relationships to be estimated given this view composition are therefore patterns of different cell types or patterns of marker expressions within these two spatial contexts.

Within the multi-view framework of MISTy, and therefore Kasumi, these relationships are estimated by constructing a non-parametric and nonlinear machine learning model. In these models, each target variable (cell type or marker abundance) in the intrinsic view is modeled by all other variables (predictors) coming from each view independently. In each model, the estimated importance of each variable as a predictor is a quantification of its potential relationship with the target variable in the context of each view (see Methods). Filtering the set of all relationships to include only those with the highest importance results in a global relationship-oriented representation of the sample. For example, one can select only the predictor-target relationships one standard deviation above the mean importance per target and view. Simply put, a global relationship-oriented representation of a sample is the set of the most robust predictor-target relationships and their estimated importances. Within a single sample, the estimated patterns of relationships between cell types go beyond simple co-localization. They are multivariate, can be non-linear and are determined by the form of the underlying predictive model, while still being explainable through the estimated importance of each predictor-target relationship. For example, as predictive models we use

Random Forests[27] and the importances are estimated as a function of the total reduction of variance of the target as the result of splitting by each of the variables in all constituent trees (see Methods).

Since the above-described modeling occurs at the level of a whole sample, the resulting representation captures a global view of relationships present in the tissue and as a trade-off, doesn't capture the local heterogeneity of tissue sub-structures and their potentially diverse functions. To address this issue, Kasumi learns a representation based on tissue patches sharing relationship patterns that define local neighborhoods.

A neighborhood is a regularly shaped spatially contiguous region (window) of the sample. The sliding window is defined at the level of the view composition, i.e. across all spatial contexts captured by the view composition at the same time. Let m and n be the length and width of a sample in given units of measurements (for example pixels, μm or spots) in two dimensions. The total number of windows z of size s sliding with a fraction of overlap o in both dimensions is then $z = ab$ where $a = \left\lceil \frac{m}{s(1-o)} \right\rceil$ is the number of windows per row and $b = \left\lceil \frac{n}{s(1-o)} \right\rceil$ is the number of windows per column. The k-th window then contains all units with coordinates $x \in [(k \bmod a)s(1-o), (k \bmod a + 1)s(1-o)]$ and $y \in [\lfloor k/a \rfloor s(1-o), (\lfloor k/a \rfloor + 1)s(1-o)]$. The representation of window k then is the ordered set of estimated standardized local importances of the relationships $\{M_{i,j}^{(v,k)}\}$ given local models combining all views v for a target i: $Y_{.,i}^{(k)} = \alpha_I^{(k)} + \alpha_0^{(k)}(F_0^k \circ W)(X, s, o, k, \tilde{Y}) + \sum_v \alpha_v^{(k)}(F_v^k \circ W \circ G_v)(X, s, o, k, \tilde{Y}, T)))$, where $Y_{.,i}$ is the vector of all observations of variable i in the intrinsic view, $\tilde{Y} = Y_{.,\forall q \neq i}$, F_v^k are the trained machine learning models predicting values $Y_{.,i}$ from variables in view v, W is a function that subsets units of view v that belong to window k, G are domain-specific functions that transform the intrinsic view based on spatial context X and functional context T and α are parameters of the meta-model combining predictions from the view-specific models. Each sample is then represented by its compressed and relationship-oriented form of a list of z sets of estimated importances M, comprising the *first* representation of Kasumi.

In general, given data from multiple samples it is difficult to compare directly the abundances and cell-type distributions across samples without further data manipulation and integration. This is mainly due to sampling, batch or technical effects. In contrast, the relationship-based representation is invariant to these effects and can be directly compared. Given the compressed local representation of samples, we can define consistent relationship-oriented clusters of windows as a joint compressed representation of the samples.

Since the importances are estimated per target from the intrinsic view when comparing the importance-based representation within and across samples we are interested in their relative values. To establish a similarity structure, we define a graph of similarities of all windows across all samples. Each window is represented as a node in a weighted undirected graph. The nodes are connected by edges with weights equal to the cosine similarity between the first relationship-oriented representation of the windows. Subsequently, the graph of similarities is clustered by identifying strongly connected communities[28] (see Methods). Labeling each

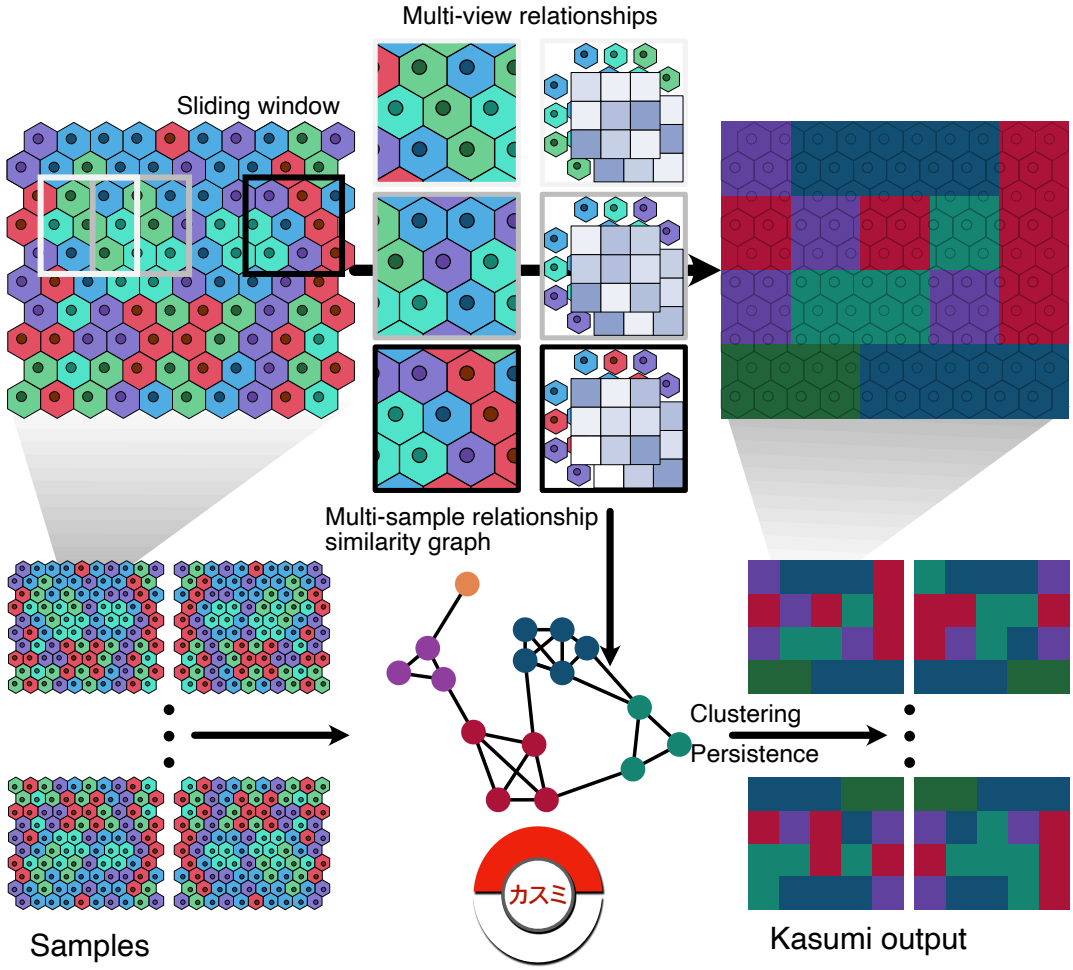


Figure 1: Kasumi takes as input a number of spatially resolved measurements from tissue samples, where each spatial location is assigned a vector of measurements. Each sample is decomposed into tissue patches by sliding a window of fixed size with or without overlap. From each window, Kasumi extracts relationships coming from different spatial contexts by estimating the importance of each measured variable as a predictor of each target variable from a multi-view multivariate non-linear predictive model. The relationships per window across all samples are used to construct a relationship similarity graph. The windows are clustered by graph community detection followed by cluster removal based on a persistence criterion. The output of Kasumi are compressed explainable relationship-based representations of the samples preserving relevant biological signals.

window with the community it belongs to leads to the *second* compressed representation of samples as a vector of discrete window identities based on shared signatures of relationships across samples coming from the same tissue and condition.

The existence of shared local structural (cell type) and functional (marker) patterns opens up a venue for exploring higher-level organization and empirical models of tissues. To define the building blocks of such models and further reduce the complexity of the representation, we regularized the description by retaining only clusters that match a heuristic persistence criterion. Namely, for a cluster to be considered to represent a meaningful local pattern it has to be persistently present in a number of samples (see Methods). The application of the persistence criterion results in the *third* and final representation of the samples by persistent local patterns.

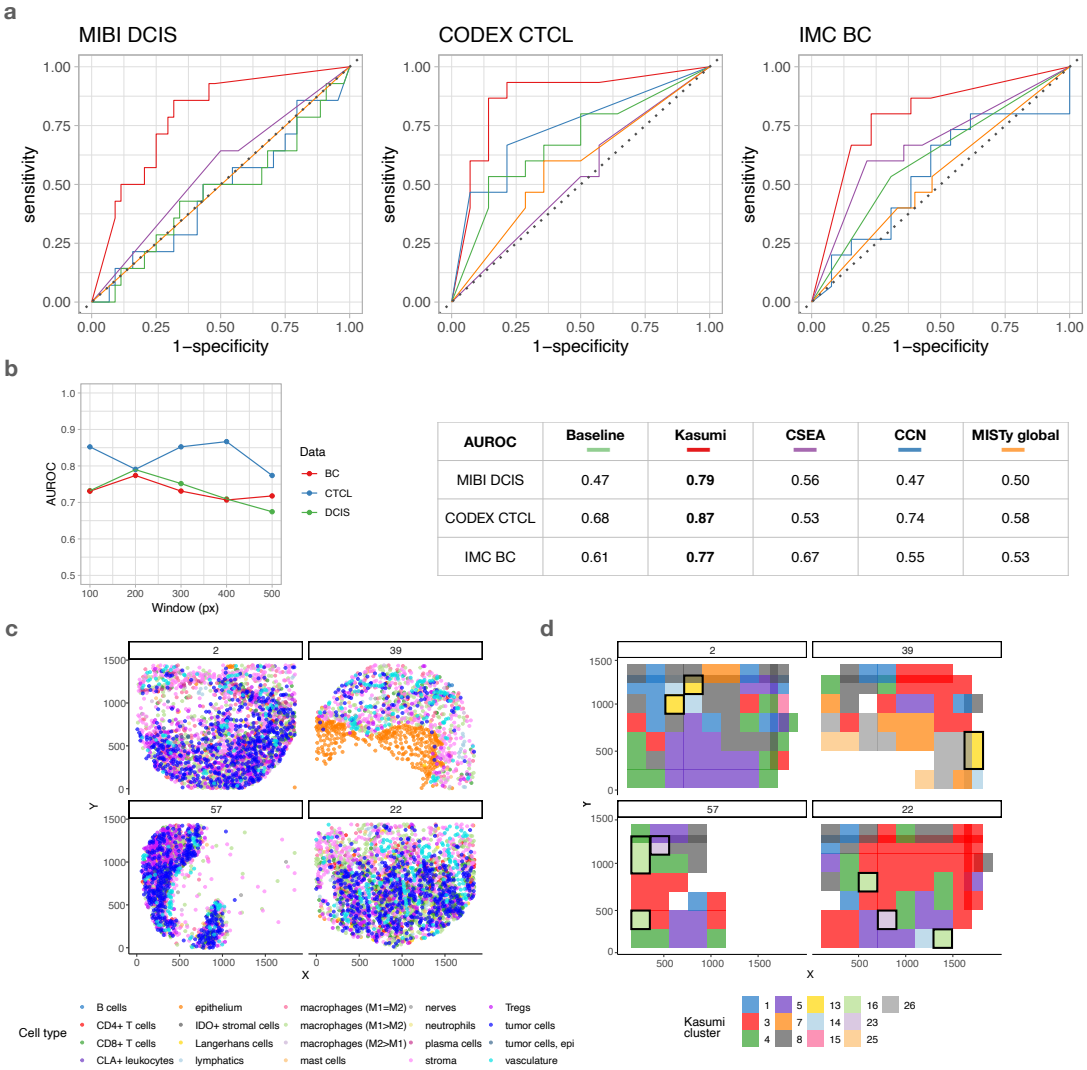
The output of Kasumi consists of the three layers of representation and the summary model statistics for each target, view, and sample. In the following, we used the third layer of representation by persistent local patterns as inputs to the predictive tasks and the other two layers to support the explanation.

2.2 Kasumi preserves relevant biological information

The representation generated by Kasumi is data-driven and unsupervised. It can be used for different downstream tasks. To demonstrate the ability of Kasumi to retain biologically relevant information while offering a simpler, lower-dimensional representation of the tissue, we evaluated it on different clinically relevant predictive tasks.

We considered highly multiplexed and spatially resolved data consisting of samples of ductal carcinoma *in situ* (DCIS)[15] measured by multiplexed ion beam imaging (MIBI)[29], samples of cutaneous T cell lymphoma (CTCL)[19] measured by co-detection by indexing (CODEX)[16] and breast cancer (BC)[30] samples measured by imaging mass cytometry (IMC)[31]. We chose recent studies that shared their data and showcased different types of neighborhood analyses. We focus on explainable approaches that can perform at the level of a single sample and can reliably scale up to a large number of samples, namely the CSEA[15] and CCN[19] methods. Due to the low (although representative) number of samples in the analyzed datasets, GNN approaches were not included in the comparison. As a baseline global representation without taking the spatial context into account, we considered the simple cell-type distributions per sample. To demonstrate the added value of considering local patterns instead of global relationships per sample, we also ran MISTy[24] using the same view composition as Kasumi.

We evaluated the performance of simple logistic regression models on the tasks of predicting progression to breast cancer from the DCIS samples ($n = 58$, 44 non-progressors, 14 progressors), predicting the response to PD-1 inhibition in pre-treatment CTCL samples ($n = 29$, 14 responders, 15 non-responders) and sensitivity to postoperative treatment with hormone therapy in Estrogen and Progesterone receptor positive, HER2 negative BC samples ($n = 30$, 15 sensitive, 15 resistant). The ability to use simple models based on a representation without further transformation suggests that the latter encodes relevant biological information and is easily interpretable.



AUROC	Baseline	Kasumi	CSEA	CCN	MISTy global
MIBI DCIS	0.47	**0.79**	0.56	0.47	0.50
CODEX CTCL	0.68	**0.87**	0.53	0.74	0.58
IMC BC	0.61	**0.77**	0.67	0.55	0.53

Figure 2: *a)* Kasumi outperforms a baseline representation of sample cell-type distribution, a global relationship-based representation, and related local neighborhood analysis approaches on tasks of predicting disease progression and response to treatment. *b)* The window size affects the performance of Kasumi. Larger window sizes decrease the performance, pointing to the importance of local patterns. *c)* Location of cell centroids labeled by cell type (cell state) as represented in the original dataset. Top – two samples of responder patients from the cutaneous T cell lymphoma (CTCL) data; Bottom – two samples from non-responder patients. *d)* Resulting Kasumi representation for the samples shown in *c*. Kasumi offers a spatially-aware, compressed but explainable representation that preserves relevant biological information. Highlighted are the windows of the Kasumi representation, that can be used to distinguish between responder and non-responder samples.

Based on the performance estimated by 10-fold cross-validation summarized by the area under the receiver operating characteristic curve (Figure 2a) the Kasumi representation outperforms the baseline representation of cell-type distribution on all tasks, underscoring the added value of spatial information. Kasumi also clearly outperforms all other representations generated by related approaches (Figure 2a), even with different choices of window sizes (Figure 2b).

The performance of Kasumi across tasks consistently decreases with the window size increase after a peak performance value (Figure 2b), pointing to the importance of capturing the local heterogeneity of tissues. Note also that the optimal choice of window size is dependent on the task and the scale of the important biological processes for that task. An example of the resulting Kasumi compressed representation resulting from a cell-type based sample representation (Figure 2c) is shown in Figure 2d. In general, local representations outperform global representations, and taking the spatial context into account always improves the performance on the selected tasks.

We further illustrate the explainability aspect of the Kasumi representation, providing insights into the neighborhood composition and the local relationship patterns.

2.3 Kasumi offers comprehensive explanation

The representation generated by Kasumi not only performs well in predictive tasks, but it is also explanatory. Given the predictive tasks, we can contextualize the Kasumi representation by assigning each cluster a relevance value for the task at hand by measuring the signed Model Reliance (Figure 3a). The value of the signed Model Reliance (sMR) is an estimate of the relevance of each cluster[32], while the sign assigns the contribution of the cluster to the prediction of one of the two classes (see Methods).

We can explore and explain the cluster-specific models. The Kasumi representation is task-agnostic, i.e., generated in an unsupervised manner, and therefore broadly describes samples. Every cluster can be explained by three aspects. First, we consider the performance of the cluster-specific models in predicting cell identity from different views. The predictive performance of the models acts as a proxy for the amount of information captured from the data. In particular, we express the performance by the variance explained by the model (R^2) for each target and the gain of variance explained when considering the spatial context (i.e. all views). We next estimate the cluster-specific importance of each predictor-target relationship coming from each view. All of the information needed to generate the explanations is already available in the Kasumi output and can be obtained by backtracking the relationship from the third to the first level of representation, i.e. from cluster label to performance and importance relationships.

In all tasks, the simple view compositions we considered consist of two views – an intrinsic view (intraview) and a broader tissue view (paraview). In the case of one-hot encoded cell types, we don't model the intraview as such a prediction is trivial. The gain in variance explained is therefore equal to the total variance explained, and the contribution is limited to the broader tissue context (Figure 3b top). Each cluster is then explained by the estimated importances of the predictor-target relationships (Figure 3b). To remove spurious

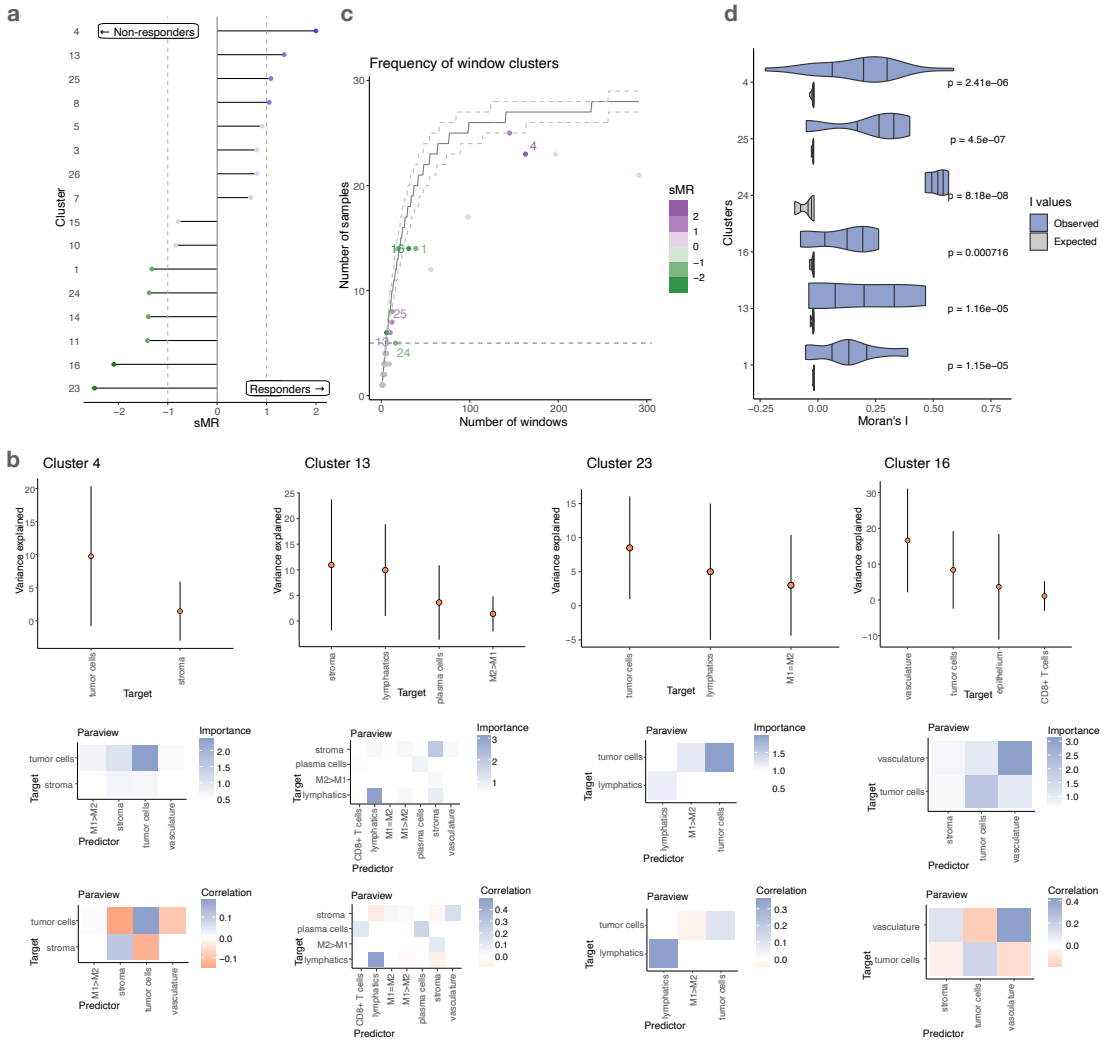


Figure 3: *a)* Signed reliance of a simple logistic regression model on the frequency of individual Kasumi clusters as sample representation for the task of distinguishing responders from non-responders in the cutaneous T cell lymphoma (CTCL) data. *b)* Distribution of variance explained per target cell type (cell state) across all windows from the same cluster (targets with variance explained less than 1% are not shown), estimated average importances of relationships describing/explaining the cluster, and average correlation between the predictor and target of important relationships across all windows coming from all samples. *c)* Number of total observed windows across samples for independent Kasumi clusters. The horizontal dotted line denotes the persistence threshold for the exclusion of clusters. The solid line represents the simulated number of samples for a cluster found in a given number of windows, as modeled by a multivariate hypergeometric distribution. The dashed lines depict the 80% confidence interval of this distribution. Relevant Kasumi clusters are consistently found to be less frequent than expected. *d)* Moran's I values for a sample of relevant clusters. The p-values are Bonferroni-corrected per cluster across all samples having at least two windows matched to the cluster, and Benjamini-Hochberg corrected across clusters.

relationships and minimize the number of false positives in our explanation, we removed targets with low explained variance and estimated importances below one standard deviation above the mean ($M \geq 1$).

The underlying model is multivariate and non-linear and therefore the Kasumi relationships should not be interpreted as linear or in a pairwise predictor-target manner. The interpretation of the results should be at the level of estimated importances. However, to improve the explainability of the results, an approximation of the sign of the relationship between the cell types (co-localization or avoidance) can be estimated by the correlation between the presence of each target cell type and the abundance of the important predictor cell types in the broader tissue structure across all windows belonging to the same Kasumi cluster. Note that low correlations of predictor-target relationships with high estimated importances are indicative of non-linearity or heterogeneity of the underlying relationships. In particular, the sign of the interaction between important relationships might be different in different windows, which in turn can be reflected as a low overall correlation. However, high correlations are indicative of a consistent linear relationship between a predictor and a target.

For example in the best-performing case, analyzing the results of the task pertaining to PD-1 inhibition treatment in CTCL, the model is highly reliant (sMR < -2) on clusters 23 and 16 to determine if the sample belongs to the group of non-responders, and on clusters 4 and 13 (sMR > 1) to responders (Figure 3a also shown on samples in Figure 2b). From these clusters 16 and 4 capture patterns of negative relationships between the tumor, stroma, and vasculature. The difference in patterns between the clusters is the presence of a positive relationship between the tumor and macrophages (M1>M2) in cluster 4 that have been shown to be associated with favorable outcome due to their anti-tumorigenesis role[33]. In cluster 13, associated with responder outcome, we also observed tumor-associated macrophages, in particular M1=M2 and M1>M2, to be positively interacting with the stroma. On the non-responder side a pattern of avoidance of M1>M2 macrophages and is also captured by cluster 23. In summary, patients with low immune activity in the tumor microenvironment are unlikely to benefit from PD-1 inhibition. This exemplifies how the interpretation of the Kasumi models can offer new perspectives on multiplexed imaging datasets and guide hypothesis generation.

The description of the interactions supporting each cluster can be complemented by its abundance and distribution across samples. How often and where a cluster is found informs us about the biological process it may represent. We observe that the frequency at which clusters are found varies greatly, demonstrating the ability of Kasumi to model phenomena at different scales (Figure 3c). For instance, cluster 4 is found in most samples although 1.6 times more common in samples from patients who responded to the immunotherapy. Cluster 16, on the other hand, is associated with non-responders although present in only a small number of windows. The contrast between cluster 4, describing a common tumor-related pattern, and cluster 16, highlights the relevance of both global, cancer-related patterns and local microenvironment structures. Similar observation can be made for clusters 13 and 23. The cluster frequency is independent of the model reliance for these clusters and of whether they are associated with positive or negative outcomes. Clusters 23, 16, and 13 have relatively low frequency but high relevance for the predictive task, while cluster 4 has high

frequency but is also found to be relevant. Clusters that are found in multiple windows per sample (Figure 3c) are also spatially autocorrelated (Figure 3d), i.e. show spatial contiguity. The spatial contiguity of Kasumi clusters is significantly stronger than expected from a random uniform distribution and systematically pointed at processes restricted to local tissue regions.

In addition to offering a more comprehensive explanation, Kasumi reduces the complexity of representation (number of clusters) compared to the number of cell types (Figure 2c vs. d). For simplicity and in order to show the amount of redundancy of the representation, the logistic models were not regularized. Clusters with an estimated absolute value of sMR above 1 are considered to be relevant for the task and can be further ordered by importance. The clusters with an absolute value of sMR below 1 are not relevant to the task or contain redundant information and can be excluded from further interpretation or analysis.

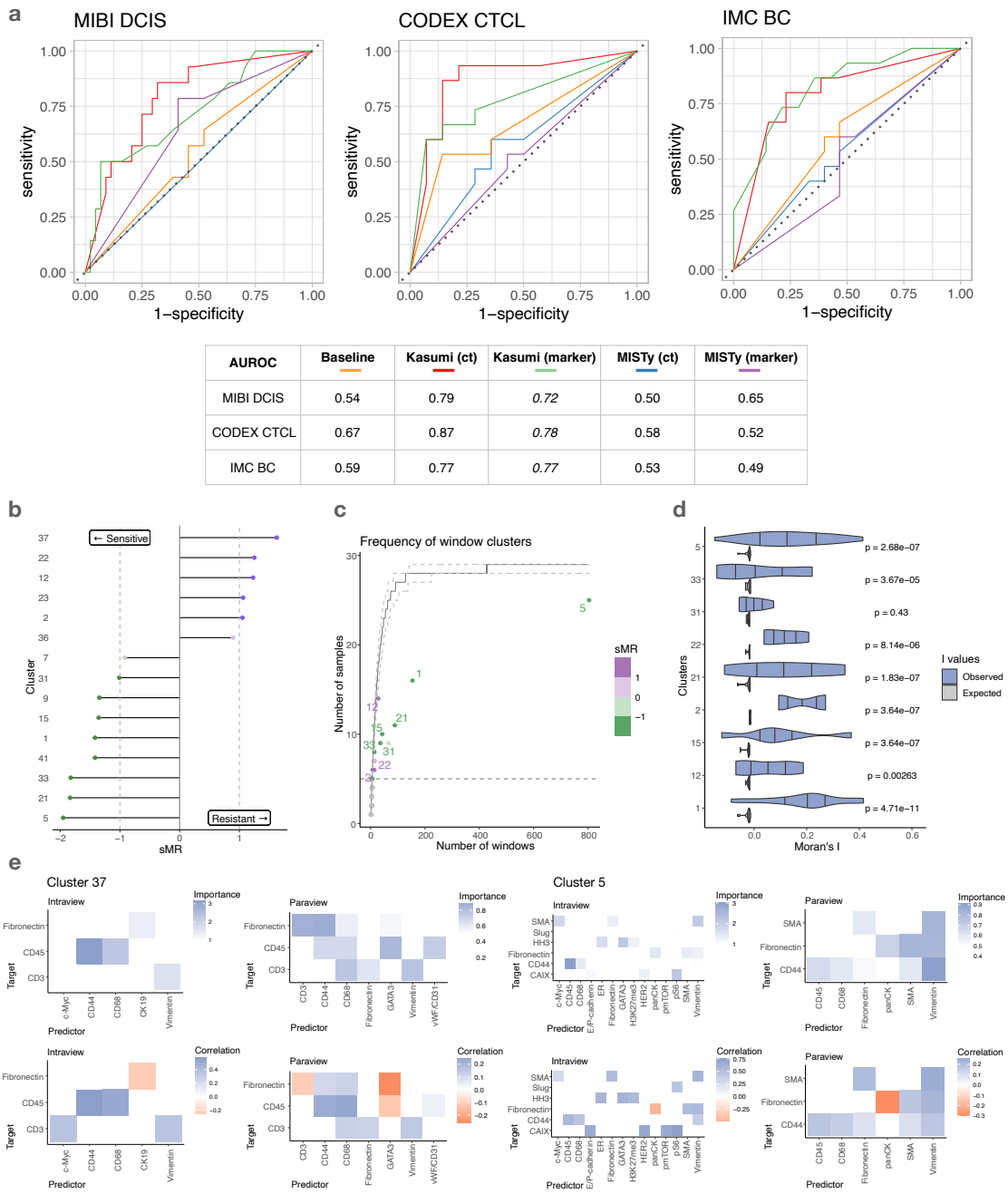
2.4 Kasumi can model quantitative cell features

Many spatial analyses are solely based on cell types inferred from lineage markers, disregarding other information typically available, such as the abundance of functional markers. Kasumi, however, is not limited to cell-type level input and can model all available measurements. The flexible definition of view composition as input allows for cell-type-agnostic analysis directly at the level of the observed spatial profiles. To broaden the definition of the state of each cell to include functional aspects, the intrinsic view can alternatively describe each cell by the mean or median abundance of the observed markers. The broader tissue context for each cell then corresponds to the relative-distance-weighted sum of abundances (see Methods). In this manner, Kasumi facilitates a more consistent relationship-oriented representation across spatial contexts. Instead of estimating relationships between predefined cell types, here each window is also characterized by the intrinsic relationships that are preserved within the window, offering a different perspective and generating alternative hypotheses to explore.

The performance of the Kasumi representation on the same tasks, given spatial profiles (marker abundances) at input remains consistently high (Figure 4a). As in the cell-type-based analyses, local representations outperform global representations. Here, a global representation was obtained by running MISTy at the level of the whole slide with the same composition as Kasumi. As a baseline global representation with no spatial context, we considered the sum of abundances per marker relative to the total abundance in the sample.

The same task-oriented analyses and explanations can be applied in this context. As an example, Figure 4b shows the sMR values of the Kasumi clusters for the task of predicting sensitivity to treatment with hormonal therapy in breast cancer samples, which we selected as the task with the least performance loss relative to the cell-type-based task.

Interaction patterns within clusters show local (window-scale) cell-type marker but also functional marker relationships that are consistent and shared in up to 25 out of 29 samples (Figure 4c). Here as well, Kasumi is able to highlight local organization at different scales and frequencies. Furthermore, across all Kasumi clusters, the widows assigned the same cluster label show significant spatial autocorrelation (Figure 4d). We observe that some markers, and by extension the biological processes they contribute to, are more influenced by their



AUROC	Baseline	Kasumi (ct)	Kasumi (marker)	MISTy (ct)	MISTy (marker)
MIBI DCIS	0.54	0.79	*0.72*	0.50	0.65
CODEX CTCL	0.67	0.87	*0.78*	0.58	0.52
IMC BC	0.59	0.77	*0.77*	0.53	0.49

Figure 4: *a)* Kasumi performance compared to a baseline representation sample marker abundance distribution, a global relationship-based representation and cell type (cell state) based Kasumi representation on tasks of predicting disease progression and response to treatment *b)* Signed reliance of a simple logistic regression model on the frequency of individual Kasumi clusters as sample representation for the task of distinguishing resistant from sensitive samples in the breast cancer data. *c)* Number of total observed windows across samples for independent Kasumi clusters. The horizontal dotted line denotes the persistence threshold for the exclusion of clusters. The solid line represents the simulated number of samples for a cluster found in a given number of windows, as modeled by a multivariate hypergeometric distribution. The dashed lines depict the 80% confidence interval of this distribution. Relevant Kasumi clusters are consistently found to be less frequent than expected. *d)* Moran's I values for a sample of relevant clusters. The p-values are Bonferroni-corrected per cluster across all samples having at least two windows matched to the cluster, and Benjamini-Hochberg corrected across clusters. *e)* Estimated average importances of relationships describing/explaining the cluster and average correlation between the predictor and target of important marker relationships across all windows coming from all samples in two spatial contexts: intrinsically (within each cell) and in the broader tissue structure.

spatial context than others. This includes both common lineage markers, such as CD3 and CD45, and markers more directly related to cellular state and function, such as CAIX and HER2.

In Figure 4e we explore features underlying Clusters 37 and 5, found to be most relevant for the predictive task. Cluster 37 associated with resistance to therapy was present in a small number of windows and is explained by an immune pattern involving CD44, CD45, and CD68, while a pattern of negative association between CD3, and GATA3 as predictors of Fibronectin is specific to the broader tissue context hinting at a loss of structure and a particular conformation of T lymphocytes hampering the response to hormonal therapy. On the other hand, cluster 5, found in the most samples and the most relevant for prediction of sensitivity to therapy shows a similar immune pattern as in cluster 37 complemented with complemented with a well-coordinated organization of multiple cell types in the tumor environment. Intrinsically we also observed a positive association between CAIX, mTOR and pS6.

3 Discussion

By establishing similarity based on common relationships across different spatial contexts instead of similarities in spatial profiles, Kasumi offers a novel representation that can be used for data exploration and hypothesis generation.

Spatially resolved data can improve patient stratification into treatable groups beyond clinical histology and beyond highly multiplexed data from dissociated tissues. Instead of characterizing samples based on standardized pre-selected clinical features, the Kasumi representation offers an unsupervised but meaningful and explainable association to observable clinical outcomes.

Kasumi representations can easily be adapted to explore tissue-specific common and differential patterns between conditions, even beyond association to clinical features. Furthermore, while we considered view compositions capturing a single spatial context, Kasumi can be deployed with more complex compositions addressing different spatial and functional contexts and different omics technologies[34–38].

On tasks of predicting disease progression and response to treatment, the performance of Kasumi shows the added value of considering spatially resolved data in contrast to baseline cell-type distributions, significantly outperforming related methods. We further show that biologically relevant patterns are localized and that the local heterogeneity of tissues improves association to clinical features not only when considering cell-type level annotated samples, but also directly from measurements.

Kasumi offers an interpretable framework to describe spatial data, highlighting both expected relationships and potentially new findings that can guide experimental follow-ups. We found that Kasumi clusters have more predictive power than simple local descriptions of cell type or marker composition. We achieved the best predictive performance on the prediction of immune checkpoint blockade response in the CTCL dataset. Of note, the most predictive clusters did not directly match the response criteria proposed in the original publication (distance between CD4+ cells and tumor cells, or between CD4+ cells and

Tregs)[19]. These results illustrate how Kasumi allows users to look further than pairwise relations and identify more complex global and local patterns, exploring systematically cellular organization within and across samples. In agreement with the original publication, we do find that both the properties of the tumor and the lymphoid component of its immune micro-environment are needed to understand the response to immune checkpoint blockade. Altogether, our observations underscore the explainability of our approach and the potential of Kasumi for hypothesis generation.

Of particular interest for further work is the exploration of the integration of multiple modalities. Such modalities include not only different omics measurements but also image-based pathomics features. Thanks to the advancement of spatially resolved multi-omics technologies[39–42], we can explore relationships across omics layers from the Kasumi representation to characterize the functional aspects of tissue patterns, that should lead to enhanced mechanistic insights.

Besides extending the layers, as datasets increase in numbers, we will be able to expand the complexity of the models. Higher-order (recursive) neighborhood analysis[43] is a step towards building models in the form of taxonomies of tissue structures. For this purpose, with Kasumi, we can incorporate both structural (cell-type level) and functional (marker level) aspects and derive higher-order relationships and representations. To explore the extent of preserved higher-order relationships concerning different measures of similarities and towards establishing empirical taxonomies of tissue structure, Kasumi can also be combined with the existing related neighborhood analysis approaches.

In summary, Kasumi captures and accounts for multivariate and non-linear persistent spatially localized relationship patterns. It outperforms related approaches for neighborhood analysis, both at the level of cell types or without relying on cell-type labels. Kasumi is practically deployable even with a limited number of samples and can support early-stage translation efforts by targeted data exploration and hypothesis generation.

4 Methods

4.1 Data and preprocessing

The datasets used in this study were obtained from the papers in which they were originally presented.

The DCIS data was obtained from Risom *et al.*[15]. It consisted of 79 samples coming from 70 patients. From those, we selected all 58 samples coming from 58 patients that were associated with either progressor ($n = 14$) or nonprogressor ($n = 44$) status. The samples were segmented and annotated with 17 distinct sublineage labels based on the spatial profiles of the cells. The panel consisted of 38 antibody markers. For our marker abundance level analysis, we considered all of them without any further processing beyond the aggregation at the level of cells already available in the data.

The CTCL data was obtained from Phillips *et al.*[19]. It consisted of 69 samples coming from 14 patients. From those, we selected 29 samples coming from 14 patients taken pre-treatment and were associated with responder ($n = 14$) or nonresponder ($n = 15$) status.

The samples were segmented and annotated with 21 distinct cluster labels associated with cell types, based on the spatial profiles of the cells. The panel consisted of 56 antibody markers. For our marker abundance level analysis, we considered all of them without any further processing beyond the aggregation at the level of cells already available in the data.

The BC data was obtained from Jackson *et al.*[30]. We considered the Basel cohort only, due to the availability of required metadata. It consisted of 376 samples coming from 285 patients. From these, we selected 85 samples coming from the tumor region of 82 patients with ER+PR+HER2- clinical subtype, treated with hormonal therapy and associated with sensitive ($n = 68$) or resistant ($n = 17$) response. To obtain a balanced dataset we further randomly subselected 15 samples from each group. The samples were segmented and annotated with 27 distinct cluster labels. The panel consisted of 33 antibody markers. For our marker abundance level analysis, we considered all of them without any further processing beyond the aggregation at the level of cells already available in the data.

4.2 View construction

For all tasks, we used either the cell-type information available for each segmented cell or the aggregated abundance of markers per cell. The Kasumi and MISTy view composition consisted of an intrinsic view (intraview) where each cell was described by its one-hot encoded cell type or by the complete marker abundance spatial profile.

For the cell-type-based analysis, the intraview was complemented by a broader tissue view (paraview) capturing the cell-type composition of the 10 nearest neighbors surrounding each cell ($l = 10$) as a sum of the one-hot encoded representation of the neighbors. The parameters of the related approaches were set as suggested in the original studies describing the method or its application on one of the datasets used in this study.

For the marker-abundance-based analysis the intraview was complemented by a paraview capturing the distance-weighted marker abundance in the broader tissue structure (radial basis $w = e^{-\frac{d^2}{l^2}}$ with parameter $l = 100$px).

The paraview for each cell c was generated as $G_c = \sum_{j=1}^{n} w(d_c, l) Y_{j,.}$, where Y is the intraview matrix, d_{cj} is the Euclidean distance between cells c and j, l is a parameter of the weight function w.

4.3 Importance estimation

The importance of the estimated predictor-target relationships I_{kj} by Kasumi and MISTy is based on the total reduction of the target (k) variance by the predictor (j) in the underlying target and view specific Random Forest model[27]. It is calculated as the view (v) contribution weighted standardized importance $M_{kj}^{(v)} = \frac{I_{kj}^{(v)} - \bar{I_k^{(v)}}}{\sigma^2_{I_k^{(v}}}(1 - p_k^{(v)})$, where $\bar{I_k^{(v)}}$ is the mean, $\sigma_{I_k^{(v}}$ the standard deviation of the estimated predictor importances for target k and view v, and $p_k^{(v)}$ is the p-value of the contribution of view v for target k in the multi-view model.

4.4 Kasumi importance signatures

The Kasumi importance signature for window k is represented as a vector generated by concatenating the estimated importances for each predictor-target pair from all views. The resulting vector is filtered by setting values to 0 for the predictor-target pair where the gain of variance explained for the target was estimated to be less than 1%. The gain of variance explained per target is a proxy of the amount of information available for that target. If no relevant information was captured by the model then the estimated importances are also not relevant. The signature vector was additionally filtered such that it doesn't contain any values less than zero, i.e., doesn't contain information about interactions with less than the mean value across all importances.

4.5 Predictive tasks and sensitivity analysis

For the predictive task, we summarized all representations in the same way, by the distribution of assigned identities per sample. In the case of the baseline this is the proportion of cell types, for CSEA the significance estimated significance of pairwise interactions for all pairs of cell types, for Kasumi and CCN the proportion of clusters and for MISTy the importance signature. The size of the neighborhood is chosen to match the information given in the original data and method papers.

The relationship between the paraview parameter and the performance of the multi-view models was established in Tanevski *et al.*[24]. The paraview parameters were chosen accordingly but kept same for all targets.

For the Kasumi analysis, we considered windows with a number of cells approximately equal to or larger than the number of variables in the intraview. We established a concave relationship between the selected window size and the performance of the Kasumi representation for the predictive tasks (Figure 2b). Therefore, we selected the window size for each task to correspond to the size associated with the best performance.

The clustering of the first Kasumi representation depends on the choice of two parameters. First the choice of a similarity cutoff for generating a graph representation of the signatures. In the graph representation the the edges between nodes with cosine similarity less than the threshold are removed. The second parameter is the resolution parameter of the Leiden community detection. We further explored the sensitivity of the Kasumi representations to the choice of these two parameters as measured by the performance on the predictive tasks by a grid search and selected the values that lead to the highest performance for each task (Figure 5).

Finally, with regards to the choice of thresholds for the persistence criterion, in these examples, we retain clusters that are present in at least 10% or at least 5 samples. The clusters not fulfilling this criterion are not considered further.

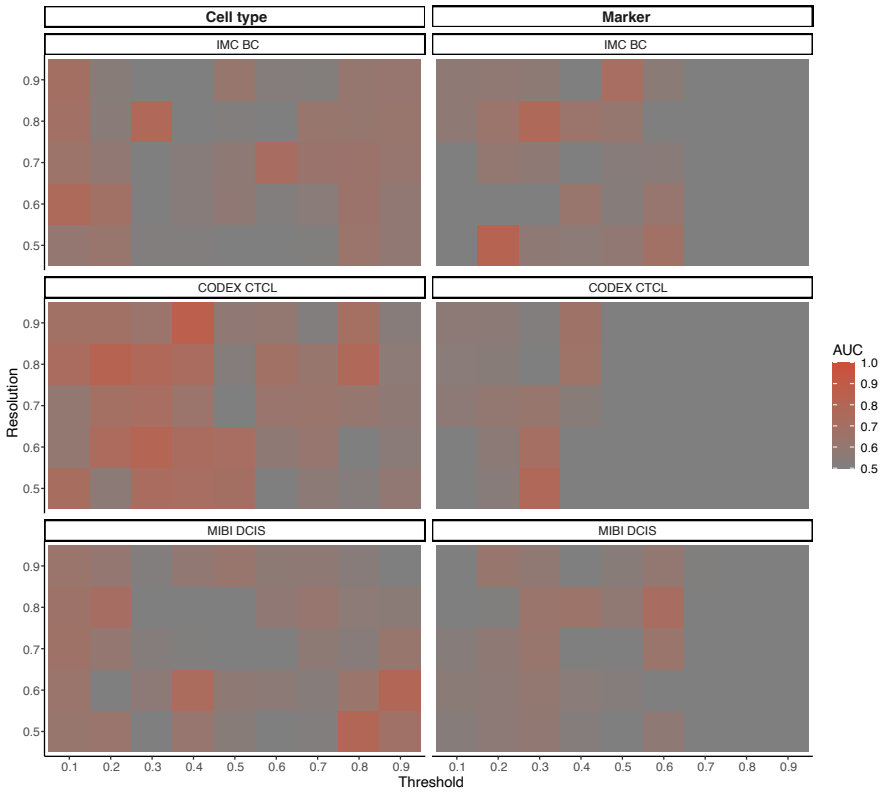


Figure 5: Sensitivity of the performance of Kasumi representation to the clustering parameters for all tasks: the cutoff threshold for the cosine similarity used to remove edges from the graph of windows and the resolution parameter of the Leiden clustering of that graph.

4.6 Signed model reliance

The relevance of each Kasumi cluster for the prediction task is estimated by the signed Model Reliance (sMR). Derived from permutation-based and model-agnostic Model Reliance[32], in the context of a linear (logistic) model sMR is defined as:

$$\mathrm{sMR}(f, D, i) = \mathrm{sgn}(\beta_i) \frac{L(f, D_{[p(i),.]})}{L(f, D)},$$

where β_i is the estimated coefficient for the i-th variable for the linear(logistic) model f given unpermuted data D, $L(f, D) = 1 - \mathrm{AUROC}(f, D)$ is the loss function, $p(i)$ is a permutation of the values of the i-th variable only.

The model reliance[32] estimates the relevance of a variable by the change in the performance of the model under noise conditions. Here, the noise is introduced by permuting the values of one feature at a time and estimating the loss relative to the performance on the unpermuted dataset. The higher the loss the more reliant the model is on that particular feature.

Declarations

Data and code availability

The datasets used in this study have been made freely available as part of the original publications as outlined in the Data and preprocessing section. Kasumi is implemented in the open-source R package mistyR and is freely available from `https://www.github.org/jtanevski/mistyR`. The code used to analyze the data and produce the results is freely available from `https://www.github.org/saezlab/kasumi_bench`.

Author contributions

J.T.: Conceptualization, Methodology, Software, Validation, Formal analysis, Writing - Original Draft. L.V.: Software, Formal analysis, Writing - Review & Editing. F.H.: Validation, Writing - Review & Editing. J.S.R.: Supervision, Resources, Writing - Review & Editing, Funding acquisition.

Acknowledgements

We would like to thank Ricardo Omar Ramirez Flores, Philipp Sven Lars Schäfer, and Sebastian Lobentanzer for the helpful discussions and Miguel Ibarra for his work on Figure 1. This publication was supported through state funds approved by the State Parliament of Baden-Württemberg for the Innovation Campus Health + Life Science Alliance Heidelberg Mannheim issued to L.V.

Competing interest

J.S.R. reports funding from GSK, Pfizer, and Sanofi and fees/honoraria from Travere Therapeutics, Stadapharm, Astex, Pfizer, Owkin and Grunenthal.

References

1. Moffitt, J. R., Lundberg, E. & Heyn, H. The emerging landscape of spatial profiling technologies. *Nature Reviews Genetics* **23,** 741–759 (2022).
2. Bressan, D., Battistoni, G. & Hannon, G. J. The dawn of spatial omics. *Science* **381** (2023).
3. Irmisch, A. *et al.* The Tumor Profiler Study: integrated, multi-omic, functional tumor profiling for clinical decision support. *Cancer Cell* **39,** 288–293 (2021).
4. Pucciarelli, D. M., Lu, B. Y., Zlobec, I. & DiStasio, M. Beyond the Lab and Into the Hospital: An Outlook on the Clinical Utility of Spatial Omics Technologies. *GEN Biotechnology* **2,** 360–371 (2023).

5. Elhanani, O., Ben-Uri, R. & Keren, L. Spatial profiling technologies illuminate the tumor microenvironment. *Cancer Cell* **41,** 404–420 (2023).

6. Palla, G., Fischer, D. S., Regev, A. & Theis, F. J. Spatial components of molecular tissue biology. *Nature Biotechnology* **40,** 308–318 (2022).

7. Alexandrov, T., Saez-Rodriguez, J. & Saka, S. K. Enablers and challenges of spatial omics, a melting pot of technologies. *Molecular Systems Biology* (2023).

8. Heumos, L. *et al.* Best practices for single-cell analysis across modalities. *Nature Reviews Genetics* **24,** 550–572 (2023).

9. Zeng, Z., Li, Y., Li, Y. & Luo, Y. Statistical and machine learning methods for spatially resolved transcriptomics data analysis. *Genome Biology* **23** (2022).

10. Langley, P. The computational support of scientific discovery. *International Journal of Human-Computer Studies* **53,** 393–410 (2000).

11. Cheng, A., Hu, G. & Li, W. V. Benchmarking cell-type clustering methods for spatially resolved transcriptomics data. *Briefings in Bioinformatics* **24** (2023).

12. Schapiro, D. *et al.* histoCAT: analysis of cell phenotypes and interactions in multiplex image cytometry data. *Nature Methods* **14,** 873–876 (2017).

13. Dries, R. *et al.* Giotto: a toolbox for integrative analysis and visualization of spatial expression data. *Genome Biology* **22** (2021).

14. 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).

15. Risom, T. *et al.* Transition to invasive breast cancer is associated with progressive changes in the structure and composition of tumor stroma. *Cell* **185,** 299–310.e18 (2022).

16. Goltsev, Y. *et al.* Deep Profiling of Mouse Splenic Architecture with CODEX Multiplexed Imaging. *Cell* **174,** 968–981.e15 (2018).

17. Chen, Z., Soifer, I., Hilton, H., Keren, L. & Jojic, V. Modeling Multiplexed Images with *Spatial-LDA* Reveals Novel Tissue Microenvironments. *Journal of Computational Biology* **27,** 1204–1218 (2020).

18. Schürch, C. M. *et al.* Coordinated Cellular Neighborhoods Orchestrate Antitumoral Immunity at the Colorectal Cancer Invasive Front. *Cell* **182,** 1341–1359.e19 (2020).

19. Phillips, D. *et al.* Immune cell topography predicts response to PD-1 blockade in cutaneous T cell lymphoma. *Nature Communications* **12** (2021).

20. Wu, Z. *et al.* Graph deep learning for the characterization of tumour microenvironments from spatial protein profiles in tissue specimens. *Nature Biomedical Engineering* **6,** 1435–1448 (2022).

21. Brbić, M. *et al.* Annotation of spatially resolved single-cell data with STELLAR. *Nature Methods* **19,** 1411–1418 (2022).

22. Yener, B. Cell-graphs: Image-Driven Modeling of Structure-Function Relationship. *Communications of the ACM* **60,** 74–84 (2016).

23. Ramirez Flores, R. O., Schäfer, P. S. L., Küchenhoff, L. & Saez-Rodriguez, J. Complementing cell taxonomies with a multicellular analysis of tissues. *Physiology* (2024).

24. Tanevski, J., Ramirez Flores, R. O., Gabor, A., Schapiro, D. & Saez-Rodriguez, J. Explainable multiview framework for dissecting spatial relationships from highly multiplexed data. *Genome Biology* **23** (2022).

25. Nagle, M. P., Tam, G. S., Maltz, E., Hemminger, Z. & Wollman, R. Bridging scales: From cell biology to physiology using in situ single-cell technologies. *Cell Systems* **12,** 388–400 (2021).

26. El Marrahi, A. *et al.* NIPMAP: niche-phenotype mapping of multiplex histology data by community ecology. *Nature Communications* **14** (2023).

27. Breiman, L. Random Forests. *Machine Learning* **45,** 5–32 (2001).

28. Traag, V. A., Waltman, L. & van Eck, N. J. From Louvain to Leiden: guaranteeing well-connected communities. *Scientific Reports* **9** (2019).

29. Keren, L. *et al.* MIBI-TOF: A multiplexed imaging platform relates cellular phenotypes and tissue structure. *Science Advances* **5** (2019).

30. Jackson, H. W. *et al.* The single-cell pathology landscape of breast cancer. *Nature* **578,** 615–620 (2020).

31. Giesen, C. *et al.* Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry. *Nature Methods* **11,** 417–422 (2014).

32. Fisher, A., Rudin, C. & Dominici, F. All Models are Wrong, but Many are Useful: Learning a Variable's Importance by Studying an Entire Class of Prediction Models Simultaneously. *Journal of Machine Learning Research* **20,** 1–81 (2019).

33. Lin, Y., Xu, J. & Lan, H. Tumor-associated macrophages in tumor metastasis: biological roles and clinical therapeutic applications. *Journal of Hematology & Oncology* **12** (2019).

34. Kuppe, C. *et al.* Spatial multi-omic map of human myocardial infarction. *Nature* **608,** 766–777 (2022).

35. Valdeolivas, A. *et al.* Profiling the heterogeneity of colorectal cancer consensus molecular subtypes using spatial transcriptomics. *npj Precision Oncology* **8** (2024).

36. Masarapu, Y. *et al.* Spatially resolved multiomics on the neuronal effects induced by spaceflight. *Research Square* (2023).

37. Lerma-Martin, C. *et al.* Spatial cell type mapping of multiple sclerosis lesions. *bioRxiv* (2022).

38. Shi, Y. *et al.* Spatiotemporal cell landscape of human embryonic tooth development. *bioRxiv* (2023).

39. Vicari, M. *et al.* Spatial multimodal analysis of transcriptomes and metabolomes in tissues. *Nature Biotechnology* (2023).

40. Vickovic, S. *et al.* SM-Omics is an automated platform for high-throughput spatial multi-omics. *Nature Communications* **13** (2022).

41. Liu, Y. *et al.* High-plex protein and whole transcriptome co-mapping at cellular resolution with spatial CITE-seq. *Nature Biotechnology* (2023).

42. Ben-Chetrit, N. *et al.* Integration of whole transcriptome spatial profiling with protein markers. *Nature Biotechnology* **41,** 788–793 (2023).

43. Bhate, S. S., Barlow, G. L., Schürch, C. M. & Nolan, G. P. Tissue schematics map the specialization of immune tissue motifs and their appropriation by tumors. *Cell Systems* **13,** 109–130.e6 (2022).