mosna reveals different types of cellular interactions predictive of response to immunotherapies in cancer.

Alexis Coullomb[1,2] & Vera Pancaldi[1]

[1]Department of Bioinformatics, Centre de Recherches en Cancérologie de Toulouse, INSERM, 31100, France
[2]Center for Biological Physics and Department of Physics, Arizona State University, Tempe, AZ 85283

Abstract

Single-cell spatially resolved proteomic or transcriptomic methods offer the opportunity to discover cell types interactions of biological or clinical importance. To extract relevant information from these data, we present *mosna*, a Python package to analyze spatially resolved experiments and discover patterns of cellular spatial organization. It includes the detection of preferential interactions between specific cell types and the discovery of cellular niches. We exemplify the proposed analysis pipeline on spatially resolved proteomic data from cancer patient samples annotated with clinical response to immunotherapy, and we show that *mosna* can identify a number of features describing cellular composition and spatial distribution that can provide biological hypotheses regarding factors that affect response to therapies.

Introduction

The spatial organization of cell types and their interactions play a major role in the development and function of organs and organisms, and in the progression of diseases such as solid tumors. Recent developments in single cell RNA sequencing technologies have enabled the discovery of a rich variety of cell types and states in various animal models or diseases, as well as the uncovering of new signaling pathways of clinical importance. However, the tissue dissociation step required for single-cell sequencing eliminates the information regarding spatial context within solid tissues, thus impairing the discovery of interactions between cell types or states potentially relevant to biological or clinical questions. More recent developments in spatially-resolved omics methods allow quantifying with single-cell resolution the abundance of several tens of proteins[1–3] and up to several thousands of transcripts[4,5] while preserving the spatial integrity of samples. Analyzing the complex datasets produced by these methods can be a daunting task, and extracting biologically or clinically relevant information remains a challenge. Here we present the Multi-Omics Spatial Networks Analysis library (*mosna*), which allows analyzing any type of spatially-resolved omic experiment to extract relevant cellular interactions or tissue spatial organization features relative to experimental or clinical groups. This library includes the computation of assortativity and mixing matrices, which quantify preferential interactions between cell subsets in terms of statistical estimates across the whole sample. *mosna* also implements the Neighbors Aggregation Statistics (*NAS*) method to discover local cellular

niches by aggregating various features, such as cell types or raw markers data. Finally, we implemented methods to simplify the analyses and test the relevance of features to the experimental or clinical groups in question, in particular identifying biomarkers of response to immune check-point blockers. Finally, we exploit these features and demonstrate the development of machine learning models to predict biological evolution or disease progression. *mosna* is made publicly available to the community, together with relevant documentation (https://github.com/AlexCoul/mosna).

Results and discussion

Spatial network reconstruction and predictions from cell types counts

To discover potentially clinically relevant cellular interactions, *mosna* leverages our previously published library *tysserand*[6] to represent tissues as spatial networks. In these networks, nodes are cells and edges (or *links*) between cells represent their physical interactions, either proximity or direct contact. This representation is compatible with any type of spatially resolved omic method, including image-based proteomics (MIBI-TOF[2], CyCif[1], CODEX[3]), image-based transcriptomics (MERFISH[4], SeqFISH[5], osmFISH[7]) or sequencing-based spatial transcriptomics (VISIUM 10x[8], Slide-seq[9]). Using *tysserand*, both 2D and 3D networks can be reconstructed with high speed and accuracy, based on segmented images or cells' coordinates, with multiple reconstruction methods, including an optimized Delaunay triangulation method for tissue network reconstruction. This network representation is agnostic of the type of measurement, and nodes in these networks can hold arbitrary attributes, such as cell position, raw and processed transcriptomics or proteomics data, cell morphology, etc…

As *mosna* intends to provide insight on the spatial organization of tissues and cellular interactions within them, an analysis pipeline of increasing refinement is proposed, in order to test simpler hypotheses first and use more sophisticated tools next. For each level of complexity, clinical relevance or predictive power of newly computed variables are estimated with descriptive statistics (Mann-Whitney rank test, Benjamini/Hochberg corrected…), plots, and machine learning predictive models, as described below. We exemplify this analysis pipeline using a spatially resolved proteomic data of Cutaneous T-Cell Lymphoma (CTCL), generated with the CODEX method on 70 samples from 14 patients treated with anti-PD-1 immunotherapy, where 7 patients responded, and 7 are non-responders[3].

The most immediately available variables that come to mind to characterize samples or understand disease progression are often cell type proportions. Thus, *mosna* facilitates testing whether or not these variables are predictive of patients' status by computing p-values, corrected for false discovery rate with the Benjamini/Hochberg method by default. In the CODEX CTCL cohort, none of the cell type proportions had a statistically significant difference of distribution between responders and non-responders (**Figure SI 1a**), as noticed by the original authors of the paper describing this dataset, which results in a poor bi-clustering of patients and cell type proportions (**Figure SI 1b**). However, even if variables taken alone do not have a high predictive power, a combination of these variables could be predictive of patients' groups. Thus, *mosna* makes it straightforward to train a machine learning model, which is an elastic-net penalized logistic regression (LR) model trained with cross-validation, evaluate its performance, and assess the importance of variables in this

model. After training on cell types proportions, the LR model performance was assessed with the Area Under the Receiver Operating Characteristic Curve (ROC AUC). The performance was poor, as the ROC AUC was 0.5 (**Figure SI 1c**), which is equivalent to a random classifier, thus cell type proportions do not carry information to predict response to immunotherapy in this case.

We then hypothesized that an imbalance in the abundances of specific cell types could explain clinical outcomes. To explore this lead, *mosna* can automatically generate composed variables by computing ratios of cell type proportions, or if needed higher order ratios, i.e. ratios of ratios of cell types, and so on. The first order of composed variables, the ratios of cell type proportions, produces 3 variables that are statistically different between response groups without applying FDR correction: the ratios lymphatics / macrophages (M1=M2), the lymphatics / vasculature and lymphatics / macrophages (M2>M1) (**Figure SI 1d**). As many variables are produced during this process, after FDR correction all composed variables are not statistically different between groups. This lack of predictive power of these combined features is confirmed by a poor bi-clustering of patients and ratios of cell type proportions (**Figure SI 1e**), and by training Machine Learning models with a ROC AUC of 0.25, which is worse than random predictions (**Figure SI 1f**).

Next, after computing the second order of composed variables, meaning the ratios of ratios of cell types, we obtained 1054 composed variables that are statistically significant without FDR correction but, again, as the number of produced variables is even higher than for the first order variables, none of the composed variables are statistically significant after FDR correction. Nonetheless, when using composed variables with p-values below 0.05 and below 0.005 without FDR correction, bi-clustering of patients and composed variables showed a very good blind clustering of patients per response group (**Figure 1a,b**).

Finally, after training a LR model on composed variables, the model had moderately high classification performance (ROC AUC = 0.75) of patients into response groups when trained on variables with p-values below 0.05 (**Figure 1c**), and had perfect performance (ROC AUC = 1) when trained on variables with p-values below 0.005 (**Figure 1d**). This is probably due to the low number of observations (patients) compared to the higher number of variables to be selected in the first case, whereas training of the model is facilitated in the second case by the selection of the most significant variables.

Of note, the third most important variable in the second case, ((CD4+ T cells / stroma) / (Tregs / lymphatics)), could be related to the *SpatialScore* defined by the authors of the original paper to predict response to therapy, as this score is the ratio of distances between CD4+ T cells and their nearest tumor cell, and CD4+ T cells and their nearest Tregs cells.

Preferential cell type interactions are predictive of response to immunotherapy

As evidenced by Philips et *al.* and other studies[10,11], the mere counts of cell types can be insufficient to predict or explain disease progression or response to therapy, and integration of spatial data may be required. To decipher whether cell types interaction patterns at the scale of whole samples can explain response to therapy, *mosna* can compute the mixing matrix (MM) and the assortativity coefficient (AC) of their corresponding spatial networks.

Assortativity is a measure of preferential interactions between nodes that share the same attributes, like cell types or marker positiveness[12,13]. The AC is computed from the mixing matrix (MM), where each element $e_{i,j}$ is the fraction of edges between nodes with attribute i

and nodes with attribute j (**Figure 2a,b**). The MM can thus inform us more precisely on how a specific type of cell tends to be preferentially close to another type of cell, or to the same type in the diagonal of the matrix. The AC and MM of networks are traditionally computed on exclusive attributes, for instance cell types. To our knowledge, *mosna* is the only library able to compute AC and MM on non exclusive attributes, like positiveness to markers. This is particularly interesting to study preferential interactions between cells positive for a pair of markers, when they are also positive for other markers, as is often the case in spatial omics data. In order to account for differences in cell type proportions, we compute the z-scored MM and the z-scored AC after random permutations of cell attributes in the network.

We used *mosna* to compute the z-scored MM on the 70 samples of the CODEX CTCL cohort, whose elements correspond to the preferential interactions between all pairs of cell types. Five pairs of cell types interactions were found significatively different between response groups even after FDR correction, that are neutrophils with themselves, CD8+ T cells with B cells, Langerhans cells with themselves, tumor cells with lymphatics cells, and vasculature cells with themselves. The bi-clustering on all individual samples did not produce homogeneous clusters of responders and non-responders, but an LR model trained on all z-scored interactions between cell types shows perfect performance (**Figure 2d**). Taking only the 4 variables remaining significant after FDR correction also produces a LR model with perfect performance (**Figure 2e**).

Finally, after aggregating the preferential interactions in samples per patient, the bi-clustering did not partition patients into homogeneous trees of responders and non-responders (**Figure SI 2e**), although at a first glance patients appeared clustered into nearly perfect groups, which is noticeable with the lower performance of the LR model, with a ROC-AUC of 0.75. This is probably due to a lower number of observations after aggregating samples per patient, compared to the number of variables the optimization procedure has to select during training. When training the LR model on variables with a p-value under 0.05, the LR model showed again perfect performance (**Figure SI 2g**). Of note, the computation of the z-scored MM element is parameter-free, which makes it easy to use in research teams and facilitates comparisons across studies.

The Neighbors Aggregation Statistics approach defines niches predictive of response to therapy

The assortativity and mixing matrix are measures of preferential interactions between cell types at the scale of the whole network. In order to discover local communities, also called *neighborhoods* or *niches*[14,15], made of specific cell type proportions or marker levels, *mosna* also implements the the Neighbors Aggregation Statistics (NAS) method, which we originally developed to study spatial patterns in cell composition in the mouse cortex[16]. For each node, we first aggregate its measured variables with its first neighbors' (**Figure 3a**), and we compute different statistics for each of these variables, such as the mean or median, to identify a central tendency, or the standard deviation or interdecile range, to quantify the variation of the variable within this neighborhood (**Figure 3b**). By repeating this process for all nodes, we obtain a *NAS* table with N^C rows and $N^V x N^S$ columns, where N^C is the number of cells, N^V the number of variables, and N^S the number of statistics computed on aggregated variables. From this table, generally bigger than the original dataset, we perform dimensionality reduction and clustering (**Figure 3c**). The resulting clusters are niches, defined either by proportions of cell types or by statistics on marker levels depending on the

analysis choices. The niches can be visualized on the dimensionality reduced projection of the NAS data, and on the samples' spatial networks (**Figure 3d**).
Using the NAS method on the CODEX CTCL dataset, we could identify 15 niches defined from the spatial configuration of protein markers data. With *mosna*, one can easily look at the distribution of cell types across niches, normalized either by absolute proportions across samples, or by cell type or by niche (**Figure 3e**). Some niches are composed mostly of a single cell type, whereas other niches present several cell types interacting locally. From the proportion of cells in each niche per patient, *mosna* could train a LR model predicting response to immunotherapy with perfect performance (ROC AUC = 1) (**Figure 3g**). Interestingly, the variables contributing to the model were not related to niches made of tumor cells with immune cells, but consisted of niches made of a single cell type, either only vascular cells, only tumor cells, or only epithelial cells. If the presence of immune cells is known to be a major factor in the survival of patients with standard therapies, here the task was to predict the *response* of patients to immunotherapy. We should notice that at the beginning of the analysis pipeline, proportions of cell types, including vascular, tumor and epithelial cells, were not predictive of response. The fact that niches made mostly of these cell types are predictive of response to ICI suggests that local areas appearing to exclude other cell types are prognostic markers of a favorable response to immunotherapy.

Looking at the detailed composition of some of the most interesting niches, namely niches 4, 5 and 11, we noticed that they are made mostly of endothelial cells. In total they gather 7452 endothelial cells, 4 Langherans cells, 1 stromal cell and 1 cancer cell. However, these endothelial cells have been clustered into 3 different niches, and only the niche 5 contributes to the model to predict response to therapy. To explain what features contributed to this clustering, and why only this specific niche contributes to the predictive model, we performed differential variable analysis, comparing either the NAS variables, or the markers data, between pairs of niches. Several markers appeared as having statistically different distributions between niches, with endothelial cells of niche 5 having less FOXP3 and CD45RO, but more CD68 than endothelial cells in niche 4. Interestingly, performing the differential variable analysis on the NAS variables shows that the variability in marker levels can differ between niches, with for example cells in niche 5 presenting a more homogeneous level of expression of CD34 and CD11c than cells in niche 11.

Conclusion

mosna leverages spatial network representation of tissue samples to extract biologically or clinically relevant features. It provides a set of methods to ease the analysis of complex datasets, with a pipeline starting from evaluating the predictive power of standard measures, such as cell types proportions, and computing more elaborated features such as preferential cell type interactions through the z-scored assortativity, and the discovery of local niches with the Neighbors Aggregation Statistics methods. The approach is compatible with all types of spatially-resolved omics experiments, in 2D and 3D, and can be applied on annotated cell types, raw markers data, or any pre-computed attribute per cell, such as morphology or any other phenotype. Spatial networks are commonly reconstructed using the centroids of detected cells, but thanks to *tysserand, mosna* can analyze networks reconstructed with segmentation masks, where edges represent observable direct contacts between cells, even at a distance from their centroid. Furthermore, the discovery of niches is

performed jointly on all samples, which allows to find rarer or smaller neighborhoods and spatial patterns that occur across images, and which makes the interpretation of niches between samples more practical. The definition of niches across samples is a requirement to train predictive models, whether it is to estimate the survival of patients or their response to therapy. We have shown here that *mosna* can find clinically relevant niches, and that we can explore the differences in cell types composition, NAS variables or raw markers between these niches. The interpretation of these variables is a crucial task to develop more efficient patient stratification algorithms for personalized medicine, and to discover targetable biological pathways and orient the development of novel therapies. We believe that *mosna* will be a valuable tool for biological and clinical advancements exploiting the wealth of spatially resolved data that is increasingly available in both research and clinical settings.

Methods

Cellular network reconstruction

Spatial tissue networks are reconstructed with the library *tysserand*. For the CODEX CTCL dataset available at https://www.nature.com/articles/s41467-021-26974-6#Sec33, cells' centroid coordinates were available, networks were thus reconstructed with the Delaunay triangulation method after choosing by hand an edge trimming distance threshold of 200 pixels to have sufficient connectivity for all networks, but avoid connecting cells across what appeared as structural gaps, such as vasculature features. For each sample pre-treatment, a table of edges was produced, relating pairs of nodes ids, where nodes are cells and their id is given in the available data. All these edges tables were merged into a single one for access convenience for further analysis, together with their corresponding sample id to relate network data to patients.

Mixing matrix construction

The mixing matrix is built by computing for each element $e_{i,j}$ the proportion of edges between nodes that have the attribute i and those that have the attribute j. We have used undirected networks, but depending on the biological question, directed networks can be used to represent asymmetric interaction, *mosna* also integrates the computation of the mixing matrix for directed networks. The assortativity is a more general measure informing whether nodes tends to interact preferentially with other nodes that have the same attribute. It is computed from the mixing matrix following [13].

Randomization and z-score calculation

The relative proportion of cells with given attributes in the network has an effect on the apparent AC, e.g. if we have 99% cancer cells, the network will appear very assortative because most of the edges are between cells of the same type. In order to correct for this bias we perform network attribute randomization by shuffling the assignment of values of each attribute to the cells. This method preserves the links between cells and the number of cells positive for each attribute. For each sample we compute the MM and the AC for tens of randomized versions of the original network, and we use the mean and standard deviation of the MM and AC to compute their z-score. When several cell type proportions are much lower

than the other ones, interactions between these rare cell types occur more rarely during the randomization events, thus the computed z-score can be a non-finite value. To mitigate this, we randomized networks 500 times for the CODEX CTCL dataset.

Discovery of niches

To discover local communities, also called neighborhoods or niches, made of specific cell type proportions or marker levels, we developed the *Neighbors Aggregation Statistics* (NAS) method. In this method, for each node we search for its first order neighbors, or for its higher order neighbors (neighbors of neighbors, etc…). Attributes of this node and its neighbors are aggregated in a table, and statistics are computed on this table, for example counts or proportions if we aggregate cell types, or mean, standard deviation, median, mean absolute deviation, etc… if we aggregate markers data (mRNA counts, protein levels, morphological features or else). We repeat the process for each node, and we gather these statistics on aggregated variables in a single table, on which we perform dimensionality reduction, for example with UMAP. Then, clustering is performed, using for example the Leiden library, which allows to select a clustering resolution, and the defined clusters correspond to cellular niches. Clusters can then be visualized on spatial networks, and proportions of cells in inches per patient can be used to train machine learning models to predict response to therapy. We should note here that it is strongly advised to choose an appropriate data transformation if the NAS method is used on marker levels, such as log+1 or CLR transformations for gene expression and protein content data respectively.

Machine learning model

For now, *mosna* makes it easy to train a logistic regression classifier, penalized with ElasticNet, which is a combination of L1 and L2 regularization. The training is performed with 5 fold cross validation per default, and the best ratio of L1 and L2 regularization is optimized by grid search. In rare cases training can fail depending on how data is split. In this case training occurs with more training folds, up to 10 by default. Models' performance is assessed with the ROC-AUC, and plots of models' coefficients indicate the importance of variables and whether they are positively or negatively associated with response to therapy. In the near future other machine learning models will be wrapped for easy use within *mosna*.

References

1. Lin, J.-R. *et al.* Highly multiplexed immunofluorescence imaging of human tissues and tumors using t-CyCIF and conventional optical microscopes. *eLife* **7**, e31657 (2018).
2. 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).
3. Phillips, D. *et al.* Immune cell topography predicts response to PD-1 blockade in cutaneous T cell lymphoma. *Nat. Commun.* **12**, 6726 (2021).
4. Xia, C., Fan, J., Emanuel, G., Hao, J. & Zhuang, X. Spatial transcriptome profiling by MERFISH reveals subcellular RNA compartmentalization and cell cycle-dependent gene expression. *Proc. Natl. Acad. Sci.* **116**, 19490–19499 (2019).
5. Eng, C.-H. L. *et al.* Transcriptome-scale super-resolved imaging in tissues by RNA seqFISH+. *Nature* **568**, 235–239 (2019).
6. Tysserand—fast and accurate reconstruction of spatial networks from bioimages | Bioinformatics | Oxford Academic. https://academic.oup.com/bioinformatics/article/37/21/3989/6313163.
7. Codeluppi, S. *et al.* Spatial organization of the somatosensory cortex revealed by osmFISH. *Nat. Methods* **15**, 932–935 (2018).
8. Ståhl, P. L. *et al.* Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science* **353**, 78–82 (2016).
9. Stickels, R. R. *et al.* Highly sensitive spatial transcriptomics at near-cellular resolution with Slide-seqV2. *Nat. Biotechnol.* **39**, 313–319 (2021).
10. Lundgren, S. *et al.* Quantitative, qualitative and spatial analysis of lymphocyte infiltration in periampullary and pancreatic adenocarcinoma. *Int. J. Cancer* **146**, 3461–3473 (2020).
11. Cao, L. *et al.* M2 macrophage infiltration into tumor islets leads to poor prognosis in non-small-cell lung cancer. *Cancer Manag. Res.* **Volume 11**, 6125–6138 (2019).
12. Newman, M. E. J. Assortative Mixing in Networks. *Phys. Rev. Lett.* **89**, 208701 (2002).
13. Newman, M. E. J. Mixing patterns in networks. *Phys. Rev. E* **67**, 026126 (2003).
14. Stoltzfus, C. R. *et al.* CytoMAP: A Spatial Analysis Toolbox Reveals Features of Myeloid Cell Organization in Lymphoid Tissues. *Cell Rep.* **31**, (2020).
15. Goltsev, Y. *et al.* Deep Profiling of Mouse Splenic Architecture with CODEX Multiplexed Imaging. *Cell* **174**, 968-981.e15 (2018).
16. Lê Cao, K.-A. *et al.* Community-wide hackathons to identify central themes in single-cell multi-omics. *Genome Biol.* **22**, 220 (2021).

Funding

This work was supported by INSERM; Fondation Toulouse Cancer Santé and Pierre Fabre Research Institute as part of the Chair of Bioinformatics in Oncology of the CRCT.
AC acknowledges support from NIH NIMH (1RF1MH128867).

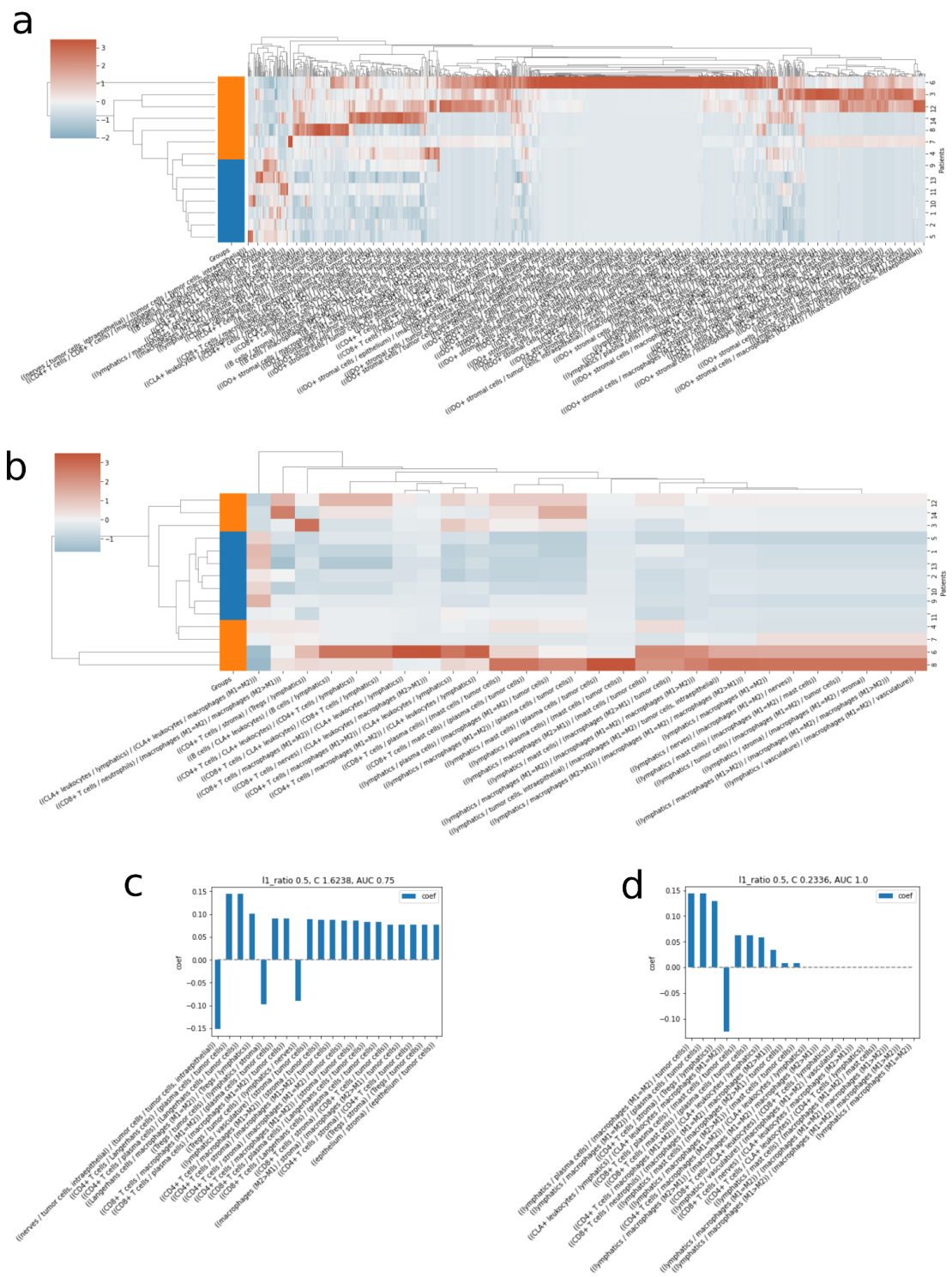

Figure 1: Predictive power of ratios of ratios of cell types proportions, showing (**a**) the bi-clustering of patients on ratios of ratios of cell types proportions with p-values < 0.05 and (**c**) the coefficients and the ROC AUC performance metric of the ElasticNet - penalized logistic regression model trained on these variable. **b**, Bi-clustering of patients on ratios of ratios of cell types proportions with p-values < 0.005. **d**, Coefficients and ROC AUC performance of the ElasticNet - penalized logistic regression model trained on these variables.

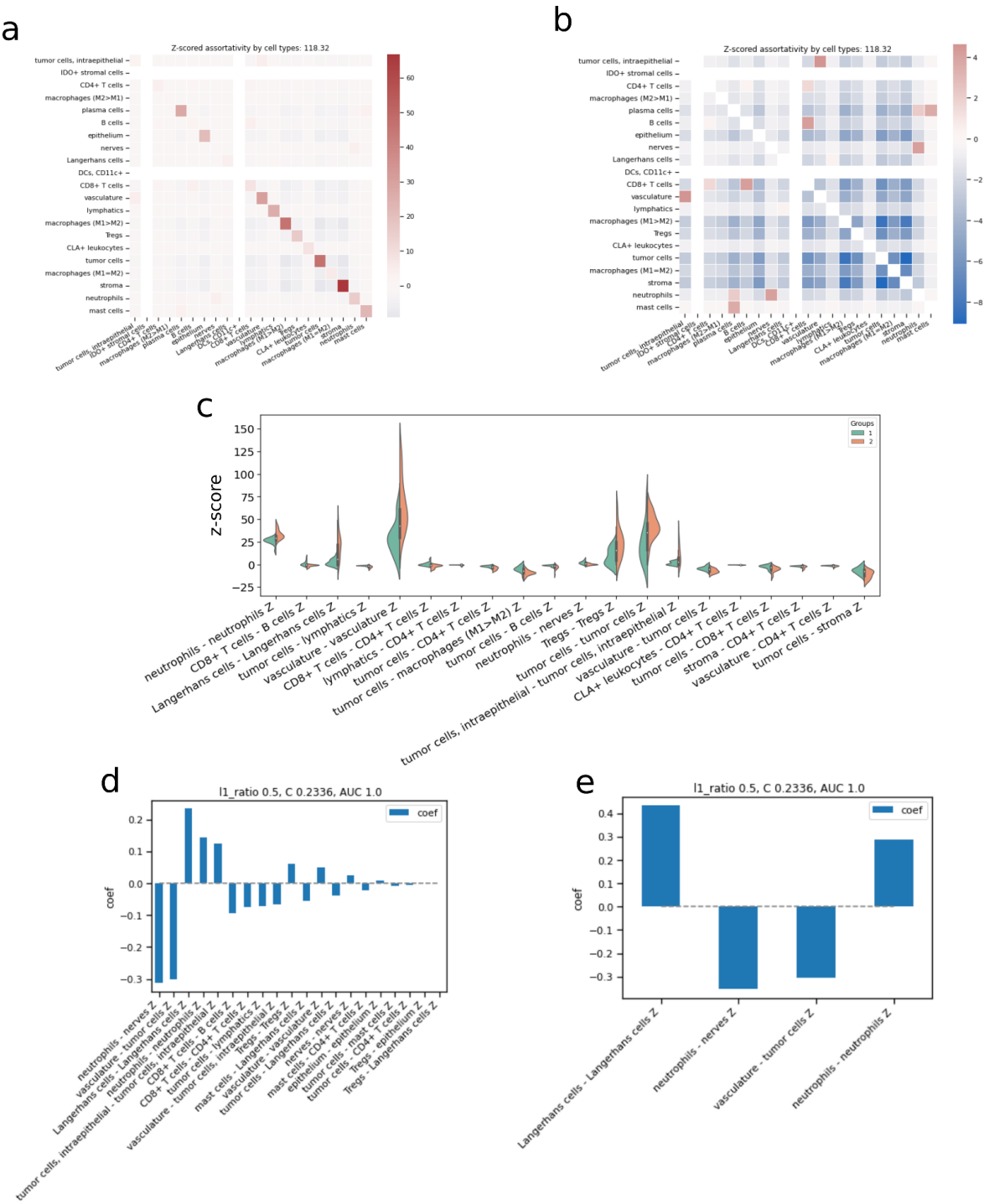

Figure 2: Assortativity measurements of cellular interactions show predictive power of response to therapy. **a**, Z-scored mixing matrix showing preferential interactions between cells of identical or different types. **b**, After removing the diagonal of the z-scored mixing matrix, we notice that most cell types tend to avoid each other (blue elements), but some of them have positive preferential interactions (red elements). **c**, Distributions of z scored interactions between cell types, computed on samples. **d**, Coefficients for the ElasticNet -

penalized logistic regression model trained on the z-scored interactions between cell types. **e**, Coefficients for the ElasticNet - penalized logistic regression model trained on the significant z-scored interactions between cell types after FDR correction.

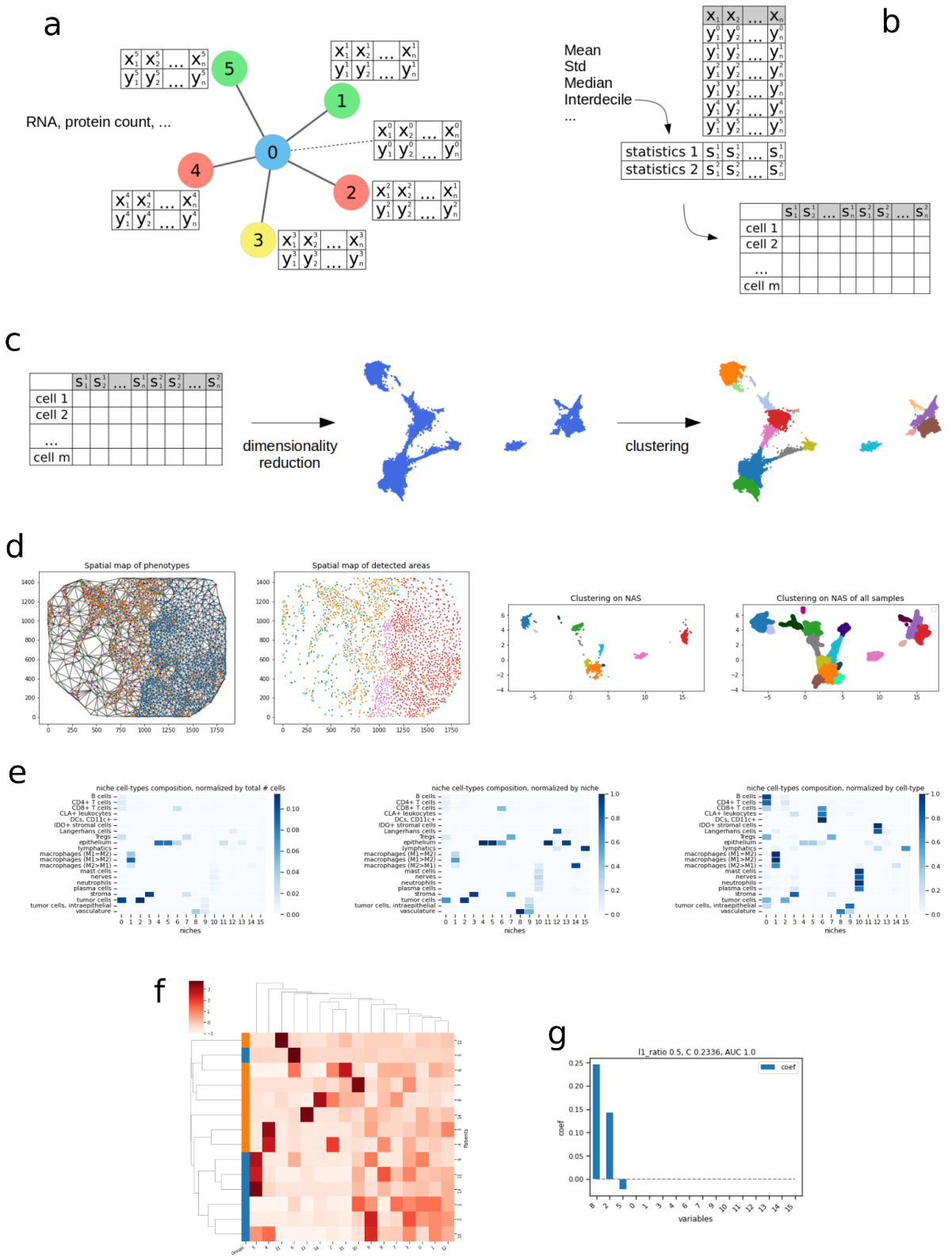

Figure 3: The NAS method defines niches predictive of response to therapy. **a**, In the NAS method, for each node variables from this node and its neighbors are aggregated. **b**, Then multiple statistics can be computed, such as the mean or the standard deviation, for each variable aggregated across neighbors. **c**, From the table of aggregation statistics, we perform dimensionality reduction and clustering to define niches. **d**, We can inspect for each sample the spatial organization of cell types (first image), the attributed niches (second), and all its cells remapped (third) to the dimensionality reduced space of niches (fourth image). **e**, Proportions of cell types in niches, normalized by total number of cells (left), by niche (middle) and by cell type (right). **f**, Bi-clustering of patients and proportions of cells in niches. **g**, Coefficients and performance of the LR model trained on niches.

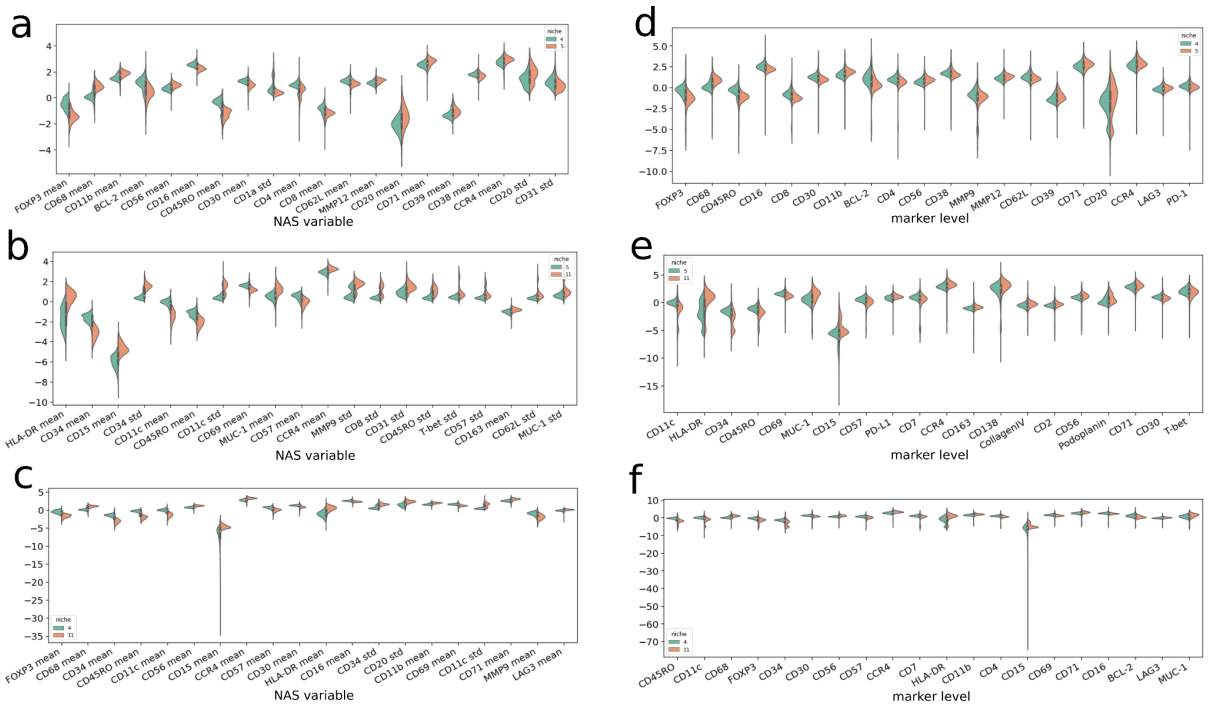

Figure 4: Differential variable analyses on NAS statistics (left column) and on marker levels (right column), between niches 4 and 5 (a & d), between niches 5 and 11 (b & e) and between niches 4 and 11 (c & f).