
Master Thesis

Cellular neighbourhood analysis in NSCLC patients using imaging mass cytometry data

by

Johannes H. Roelink

(jrk860)

Supervisors: F. Van Maldegem
E. Bosdriesz

August 29, 2024

Master Bioinformatics and Systems Biology

Cellular neighbourhood analysis in NSCLC patients using imaging mass cytometry data

Johannes H. Roelink

Vrije Universiteit, Universiteit van Amsterdam

Amsterdam, The Netherlands

j.h2.roelink@student.vu.nl

ABSTRACT

NSCLC remains one of the leading causes of death by cancer. While ICI has shown some promising results, there is not always a durable response. This resistance could be caused by a remodelling of the TME by the tumor to affect a more hospitable environment. This remodeling of the TME can be manifested by a change in the cellular communities in the tumor, while the densities of the cells in these communities remain relatively stable. Methods like IMC which can identify cell phenotypes in their spatial location and can thus capture spatial patterns. In this thesis, we clustered the cells in IMC images into cellular communities and analysed the effects of these communities on the patients. We also attempted to gain insight in the causes of these effects of a select few communities by looking at the cellular composition of the communities and the local enrichment of cell types in each other's neighbourhood. Two communities that consisted of a similar cell type fractions, showcased opposite effects on the survival of the patients. The communities differed in the number of B cells and a variety of myeloid cells that they contained. The enrichment analysis supported the findings of Sorin et al. that B cells by themselves are an indicator of a good prognosis while T regulatory cells in close contact with B cells seem to negate this effect.

DEFINITIONS AND ABBREVIATIONS

Cellular neighbourhood: area around a single cell

Cellular community/cluster: group of cells that create a pattern that occurs across the image as well as across patients

ICI: Immune checkpoint inhibitors

IMC: Imaging mass cytometry

TME: Tumour micro-environment

irAE: Immune related adverse effects

1 INTRODUCTION

1.1 Non-small cell lung cancer

Lung cancer is the leading cause of cancer deaths worldwide. The most common type, non-small cell lung cancer (NSCLC), accounts for around 85% of all cases. While a variety of treatment options exist, the overall survival remains very poor [1].

1.2 Immune checkpoint inhibitors

Effective cancer treatment often requires cooperation between the therapy and the immune system. However, cancer has a variety of methods to pacify the immune system. A new class of drugs called immune checkpoint inhibitors (ICI) have shown promising results [2]. These drugs attempt to block the proteins used by cancer to

evasive the immune system. They have the advantage over conventional platinum-based techniques with respect to chemotoxicity, but can cause immune-related adverse effects (irAE). These side-effects are often minor and can resolve by themselves or with the help of corticosteroids [3]. However, in a substantial portion of patients treated with ICI's the irAEs can be severe and even life-threatening.

While ICI have the ability to reactivate the immune system, a significant portion of patients is innately resistant to ICI or acquires resistance during the treatment [4]. The reasons for this are not fully understood, due to a large amount of heterogeneity between tumours [5]. A possible explanation is the remodelling of the tumour micro-environment (TME) by the tumour, to create a more beneficial environment for itself.

1.3 Tumour micro-environment

The TME plays a crucial role in immune response and disease progression. Studies in breast cancer have shown that the spatial structures within the TME can be influenced by the tumour to its benefit [6][7].

Therefore, simply examining single cell type densities, with techniques like flow cytometry is not sufficient to explain changes in immune response to tumours [8]. It is essential to consider the spatial context of these cells.

1.4 Imaging mass cytometry

A technique that allows for the investigation of the TME is Imaging Mass Cytometry (IMC). IMC generates detailed images of the TME, preserving the spatial structure of the cells and allowing for the identification of a wide variety of cell phenotypes based on the selected antibody panel. These antibodies, conjugated with unique rare metal isotopes, are chosen based on their binding to known cell phenotypes.

By staining and tessellating the tissue with these conjugated antibodies we can measure the amount of each specific rare metal isotope in an area, and convert the tissue in a multilayered image, where every channel corresponds to a rare metal isotope. Using image processing algorithms or machine learning, these images can then be segmented and annotated to identify the locations of the cells and their phenotypes in the image.

1.5 Cellular community identification

Multiple methods for defining and determining the cellular communities exist and for these methods choices have to be made on how to define the neighbourhood of a cell. Shurch et al. debuted a method of looking at the 10 nearest cells of an index cell and not placing a limit on the distance a cell could be to be considered a neighbour. They then clustered the cells into 9 communities based

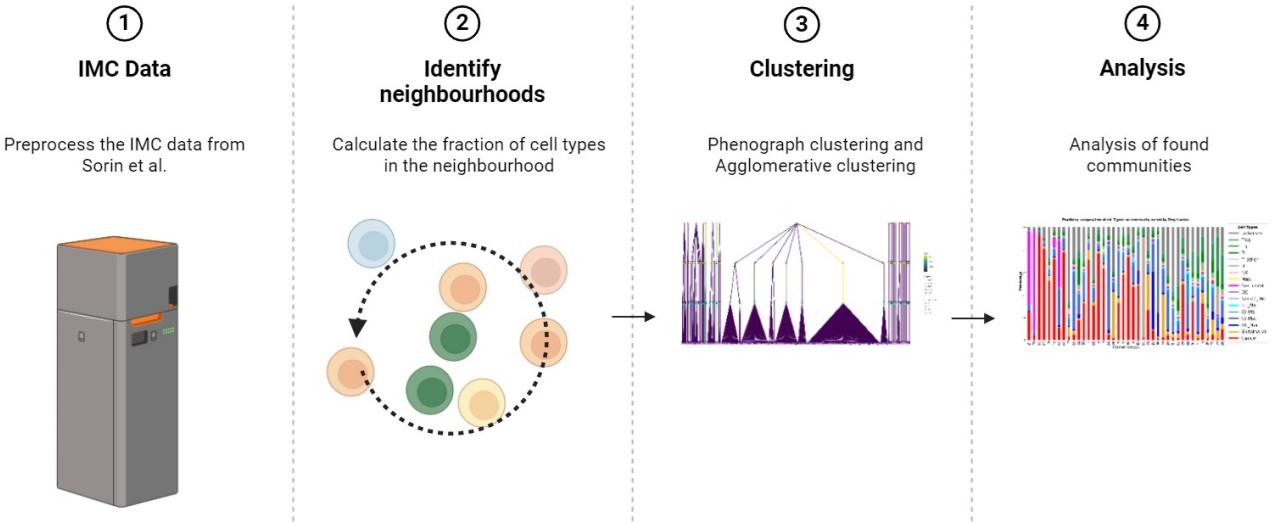


Figure 1: General outline of the workflow.

on these nearest cell proportions [9]. Sorin et al. made use of this method, as well as Enfield et al. [10][11]. The latter employed this method to determine groups of cell types that were enriched in each other's neighbourhood [11].

Other studies defined the neighbourhood as cells found within fixed distance from the index cell and thus not limiting the amount of neighbours a cell could theoretically have [12][13][5]. This results in the dilemma of determining how far a cell can be from the index cell to still be considered a neighbour. Blise et al. base their neighbourhood definition on the reach of cytokine communication, which was researched by others [12][14][15]. They used 60 micrometers from cell center to cell center as a cutoff.

Maus et al. calculated their neighbourhoods from a tumour cell point of view and they required that there was at least one non-index cell in the neighbourhood of the index cell [5]. They determined that at a cut-off of 25 micrometers, over 90% of the tumour cells fit this requirement. To account for differing cell sizes they performed their analysis with 50 micrometers in parallel.

Cole et al. utilized two different techniques on two different datasets. On their murine data, they defined a neighbourhood as every cell within 15 microns of the index cell, measuring from membrane to membrane. They also employed the method developed by Shürch et al. discussed earlier [13][9].

Cole et al. found an important role for T regulatory cells in deactivating the immune system to the benefit of the tumour [13]. For this reason, the analysis in this thesis will place an emphasis on the effect of communities with a high T regulatory cell content.

1.6 Research question

This thesis will focus on the dataset published by Sorin et al. [10]. They performed IMC staining on 416 tumour cores from lung adenocarcinoma (LUAD) patients to generate a dataset containing around 1.6 million cells. These spatially resolved cells are annotated with their cell type using an antibody panel aimed at identifying a wide variety of cells. Furthermore, they provide a variety of clinical information about the patients and the tumours.

This dataset will be utilized in this thesis to identify cellular communities within the TME. Furthermore, the effect of these communities on the outcome of the patients will be investigated. A basic outline of the workflow of this thesis is provided in Figure 1.

2 METHODS

2.1 Cellular community identification

The cellular communities are based on the neighbourhood composition of the cells. A cell was considered a neighbour of another cells if their centers were within 25 pixels (25 microns) of each other. The fraction for every cell type was then calculated per cell. Using this information, the cells were clustered using the Phenograph algorithm with $k=250$ [16]. This resulted in 278 clusters. The centroids of every cluster were calculated and these centroids were clustered again using agglomerative clustering [17]. This was done for every number of clusters between 1 and a 100. For every number of clusters, the information loss, expressed as the loss of variance compared to the original 278 centroids, was calculated.

Three numbers of clusters were identified with relatively more variance retention per number of clusters, as shown by dips in the

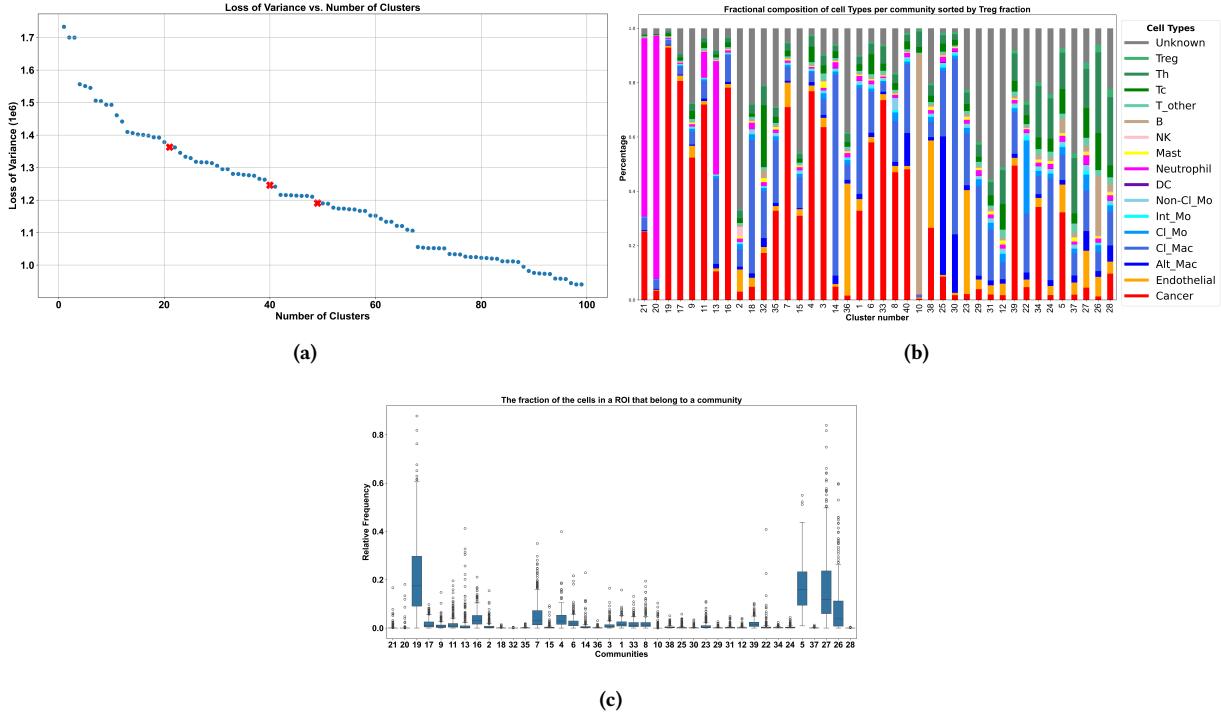


Figure 2: (a) The loss of variance plotted against the number of clusters the agglomerative algorithm merged to. The red crosses at 21, 40 and 49 clusters indicate the cluster that were analysed further. (b) Stacked barplot of the fractional composition per cell type of every cluster in NB₄₀ (c) Boxplot of the percentage of the cells in every ROI that belong to a specific cluster.

information loss landscape. These points were at 21, 40 and 49 clusters, which thus resulted in three groups of cellular neighbourhoods, henceforth called NB₂₁, NB₄₀ and NB₄₉.

The main interest of this analysis was the cellular neighbourhoods with high T regulatory cell content, based on Cole et al. [13]. NB₂₁, NB₄₀ and NB₄₉ were analysed based on their retention of communities with a high T regulatory cell content. We selected the five communities with the highest T regulatory cell fraction in the original 278 clusters and tracked their merging and/or fracturing at the NB₂₁, NB₄₀ and NB₄₉ levels. Some of these clusters merged into a new cluster between 21 and 40 clusters. For this reason NB₄₀, was considered the group with the best trade-off between number of clusters and loss of information and thus used for downstream analysis.

2.2 Community prevalence and overall survival

The images obtained from the patients were divided into three quantiles for every community. These quantiles were calculated based on the Z-score of the fraction of cells that belonged to a community in the image. The overall survival of the patients in these three quantiles was plotted in Kaplan-Meijer curves. A logrank test was performed between the high and low quantiles to obtain the significance of the difference in survival between the quantiles [18].

2.3 Singular cell type enrichment

A linear regression model was trained on the data containing the Z-score of the fractions of every cell type in the neighbourhood of every cell. The weights the model assigns to every feature were ranked. With the exception of the 'Unknown' feature, the five features with the largest weights and their corresponding cell types were used for the neighbourhood enrichment analysis. These cell types were: B cells, cytotoxic T cells, T helper cells, Endothelial cells and classical macrophages. T regulatory cells were also included.

2.4 Neighbourhood enrichment

For every cell, the neighbours, in a 25 micron radius, were obtained and the number of interactions between the cells types identified by the feature importance analysis were calculated per patient and per community. This was the baseline. For the permutation, the annotations of the neighbouring cells were shuffled and the number of cell type-cell type interactions was again calculated. This process was repeated 300 times to obtain a distribution of the number of interactions between cell types. The results of these permutations were averaged and compared to the baseline. The p-values were calculated from the number of permutations where the neighbours counts was higher and lower than the baseline. Furthermore, the log fold change (log FC2) was calculated between the mean of the counts of the permutations and the baseline. ROIs that had a significant p-value ($p < 0.01$) and had a absolute log FC2 of 1 or higher were counted as a 'hit' or a significant difference in

the enrichment compared to the rest of the community. Hits were classified as negative or positive depending on the sign of the log FC2.

3 RESULTS

3.1 Cluster merging

The Phenograph clustering of the cell type fractions per cell resulted in 278 clusters [16]. This number was too high to capture spatial structures in the cell image, because of the low number of average cells per cluster. To merge clusters and obtain more high level clusters, agglomerative clustering was used. Based on the loss of variance in the agglomerated clusters compared to the original 278 clusters, three values were chosen. These values represented the number of clusters to which was agglomerated, with a relatively high retention of variance. These values were 21, 40 and 49. The loss of variance landscape and the chosen values can be seen in Figure ??.

To decide which of these 3 values was most suited for downstream analysis, we selected the 5 clusters with the highest fraction of T regulatory cells in the group of 40 clusters, based on their reported significance in the TME [13]. It was then observed that these clusters remained stable when merging to 40 clusters, but started merging themselves when going to 21 clusters. The clusters that merged between NB_{40} and NB_{49} almost always consisted of more than 30% of cancer cells and often over 50%. Only clusters 16 and 43 contained less than 10% cancer cells, but over half of their cells belonged to the unknown cell type class, which likely explains their merge. For this reason 40 was chosen as the amount of clusters and thus the amount of distinct communities we would track. These communities can be seen in Figure ??.

For every image, the fraction of the cells that belonged to a community was plotted per community. This was done to observe the prevalence of the communities, shown in Figure 2c.

3.2 Community prevalence and overall survival analysis

To analyse the differing relationships between prevalence of a community in a patient and the survival of said patient, Kaplan-Meijer plots were generated. While most communities did not show any significant differences in survival, community 26 and community 27 did. This is shown in Figure 3a and 3b. The high prevalence quantile for community 26 had a significantly higher survival than the low quantile. The logrank test for community 26 had a p-value of $\tilde{0.029}$. This is in contrast to community 27 where the high prevalence quantile had a significantly lower survival than the high prevalence quantile with a p-value of $\tilde{0.020}$.

Community 10 also showed a significant difference between the high and low prevalence group of patients, see Figure 3c. The logrank test achieved a p-value of 0.002.

Community 10 consists mostly of B cells, as can be seen in Figure 2b. In this figure the spatial location of the cells is visualized. Cells belonging to community 10 or 26 are colored. B cells are marked with a red cross and T regulatory cells are marked with a blue cross. The B cells in community 26 tend to cluster around the (B) cells belonging to community 10 with very few T regulatory cells intermixed.

3.3 Singular cell type enrichment

This difference in survival between community 26 and 27 was interesting due to the relative similarities between the two communities in cell type fractions as shown in Figure 2b. To explain this difference two different avenues were explored. The first was a potential enrichment of a single cell type in one of the two communities which could drive the difference in survival between community 26 and 27. The second avenue was an enrichment of 2 cell types in each others neighbourhood.

Firstly, a singular cell type enrichment was explored. In Figure 5a, the enrichment of cell types in community 26 and 27 are seen compared to all 40 communities. This shows that both communities were enriched in all identified T cell types, while only community 26 was enriched in B. Community 27 is enriched in monocytes, endothelial cells, dendritic cells, mastoid cells and alternative macrophages in differing amounts. Neither community seemed to be enriched in natural killer cells, neutrophils, classical macrophages, cancer cells and unidentified cells.

3.4 Neighbourhood enrichment

To determine whether an enrichment of two cell types in each other's neighbourhood in community 26 and 27 influences the progression of the patients, an neighbourhood enrichment analysis, similar to the analysis done in [13], was performed. It was unfeasible to analyse every combination of cell types, so a selection was made based on the feature importance ranking from a linear regression. This ranking can be seen in 5b. Based on this ranking, the top five cell types were selected: B cells, classical macrophages, endothelial cells, cytotoxic T cells and T helper cells, as well as T regulatory cells.

The neighbourhood enrichment analysis uses the fold change of the counts of cell type A in the neighbourhood of cell type B using a permutation test. This fold change gets plotted against the significance value of the fold change. To determine whether there is a difference between the two communities, the number of ROIs per community where the fold change exceeds 1 or -1 and has a significance value of <0.05 is counted as positive or negative 'hits'.

Because there were 36 cell type combinations, the following criteria were set to deem a combination as significantly positively enriched in the one of the communities compared to the other: (1) The enriched community has at least 84 positive hits (20% of the ROIs). (2) The enriched community has twice as many positive hits as negative hits. (3) The positive community has twice as many positive hits as the positive hits of the other community. For a community to be negatively enriched the same criteria but inverse must hold true. The volcano plots of the cell type combinations that were enriched are shown in Figure 5c.

One cell type combination, B cell-T helper cell, was both positively enriched for community 27 and negatively enriched for community 26. The following cell type combinations were negatively enriched for community 26: classical macrophage-B cell, endothelial cell-B cell, cytotoxic T cell-B cell, T helper cell-B cell and T regulatory cell-B cell.

If we drop the third criteria: that the positive community has twice as many positive hits as the positive hits of the other community, we obtain a number of cell type combinations that are

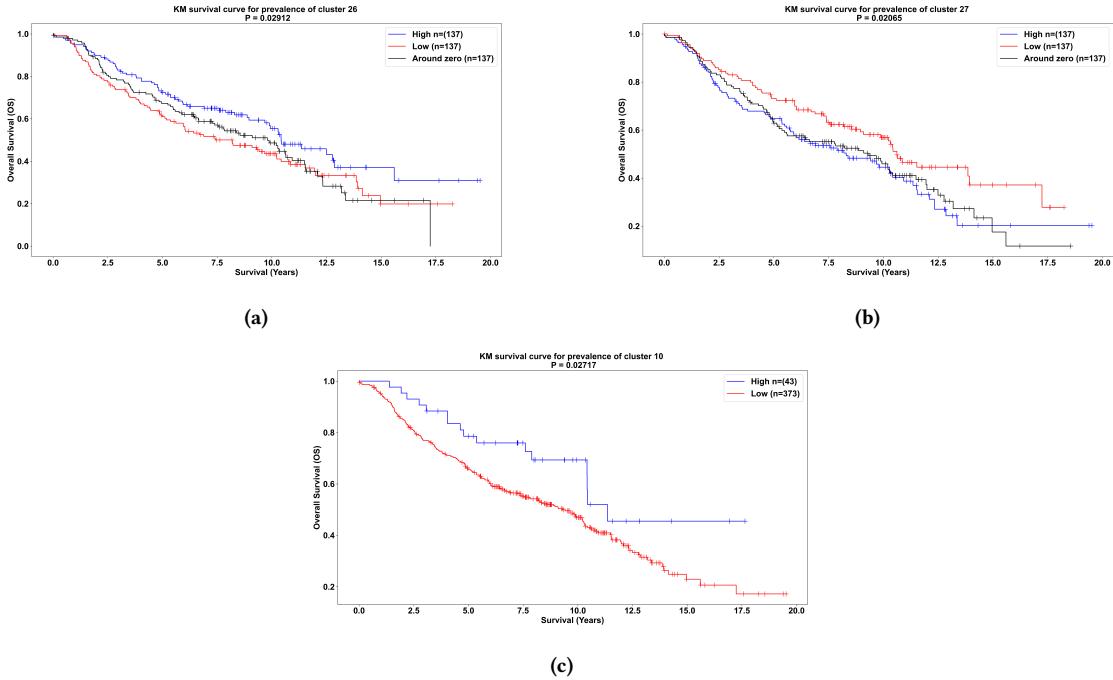


Figure 3: Kaplan-Meier curve of the overall survival of the high, middle and low prevalence group of patients (a) Community 26. (b) Community 27. (c) Community 10. Only a few samples had a z-score above 0 for community 10, so this plot shows the difference between patients with the community (high) and patients without the community (low).

positively (or negatively) enriched in both communities. These combinations are: classical macrophage-classical macrophage, endothelial cell-endothelial cell, cytotoxic T cell-cytotoxic T cell, T helper cell-T helper cell and cytotoxic T cell-T helper cell, T helper cell-endothelial cell. These volcano plots are shown in Figure 5d.

4 DISCUSSION

4.1 Conclusion

In conclusion, using a combination of Phenograph and agglomerative clustering, we identified 40 cellular communities in the images of tissues of 416 NSCLC patients. These communities captured a wide variety of spatial structures and cell densities in the TME of these patients. Comparing the survival of the patients with a high and a low prevalence of these communities, guided us to two different communities. Community 26 and 27 were similar based on fractional make up of cell types, but a high occurrence of this community in a ROI was linked to opposite survival outcomes for the patients.

However, the number of B cells differed greatly between the two communities. Around 2% of the cells in community 27 is a B cells, while this percentage is around 25% for community 26. An abundance of B cells have been positively correlated with overall survival in this dataset by Sorin et al. [10].

This is corroborated by community 10, which consists of around 90% B cells. A high prevalence of this community is also linked to a better overall survival. However, relatively few cells were assigned to this community, so conclusions drawn from this are tentative.

The community captures a tight group of B cells with very few other neighbouring cells, which does not occur a lot in the dataset, as can be seen in Figure 2c.

In addition to the link between B cells and survival, Sorin et al. noted that a co-enrichment of B cells and T regulatory cells in each other's neighbourhood negated this survival benefit. This is in contrast to our community 26, which, after community 28, has the highest fraction of T regulatory cells with around 3% of total cells belonging to this phenotype.

The neighbourhood enrichment analysis supports this finding by Sorin et al., however. The cell type combination of T regulatory cell-B cell is negatively enriched in community 26, which indicates that B cells are avoiding T regulatory cells.

This seems to have a positive effect on the survival of the patients with a high content of community 26. The T regulatory cell-B cell combination is not negatively enriched in community 27 and an abundance of cells belonging to community 27 have a negative impact on survival. It would be interesting to look at more images like Figure 4, annotated with community 27 as well, to analyze the location of these differing cell types and the communities they belong to.

The agglomeration of cells from community 10 and it being surrounded by cells belonging to community 26, a community that consists mostly of lymphoid cells, likely indicate that this community corresponds to tertiary lymphoid structures [19]. The presence of these structures have been linked to positive survival [?], just like community 10.

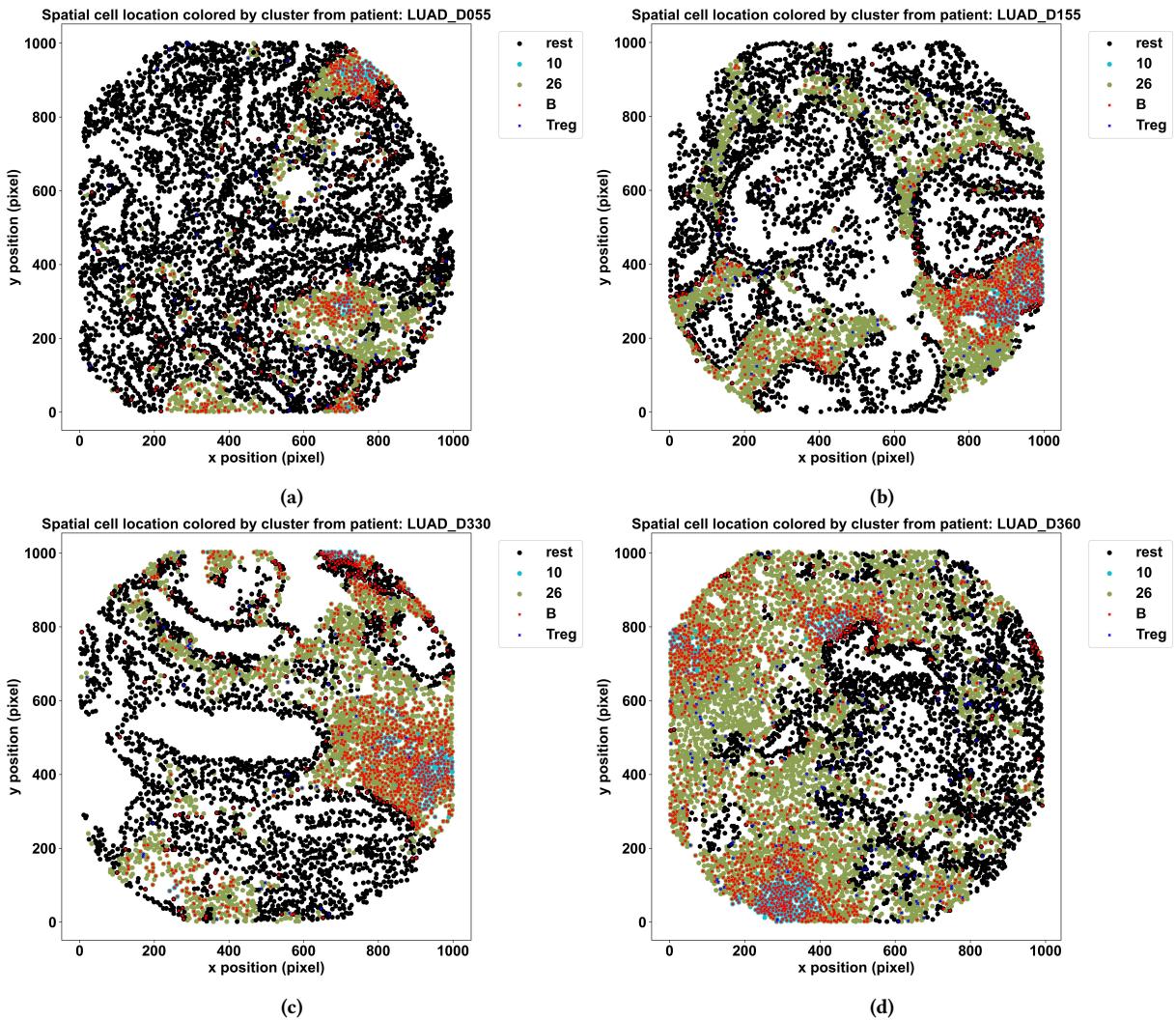


Figure 4: Random selection of ROI images from four patients out of the 20 patients with the most cells belonging to community 10. Cells belonging to community 10 or 26 are colored with green or light blue respectively. The rest of the cells are colored black. B cells are marked with a red cross and T regulatory cells are marked with a blue cross.

Every cell type that was selected for neighbourhood enrichment analysis, classical macrophages, T helper cells, cytotoxic T cells, endothelial cells and T regulatory cells, was negatively enriched with B cells in community 26. It is therefore likely that B cells tend to be enriched with a cell type that was not selected for the analysis. B cells were neither positively nor negatively enriched with themselves.

Another major difference between the communities 26 and 27 is the B cell-T helper cell combination. This combination is positively enriched in community 27 and negatively enriched in community 26, indicating a large difference between the two communities. This, in combination with the previously mentioned avoidance of the analyzed cell types from B cells is perplexing. Literature suggests that structures of lymphoid aggregates are correlated with a good prognosis [19].

NOTE: Not sure how to explain this. Literature I read tends to link avoidance of T helper cells by B cells as a negative prognostic factor but here it is correlated with a good outcome.

There is group of cell type combinations that are positively enriched with itself in both communities consisting of: Classical macrophages, Endothelial cells, cytotoxic T cells and T helper cells. The consistency of this enrichment implies that these structures do not differ much between the two communities. These cell types also have similar abundances, Figure 5a, in the two communities. This likely indicates that these cells play a small role in explaining the differences in survival between the two communities but might still affect the survival in an absolute sense.

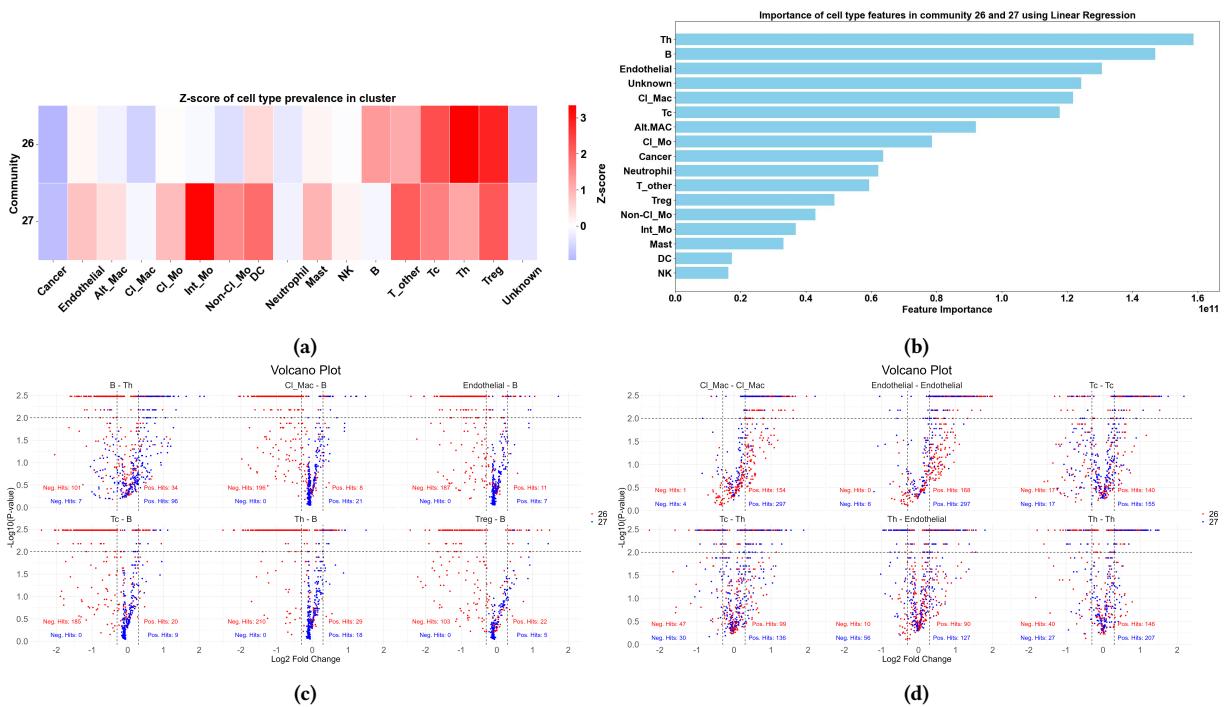


Figure 5: (a) Heatmap of the z-scores of the cell types in community 26 and 27. (b) Ranking of the weights the linear regression on survival assigned to the cell types. The top five cell types were used for neighbourhood enrichment analysis. (c) Volcano plots of the cell type combinations that were positively or negatively enriched. The x-axis shows the log FC2 and the y axis shows the negative log of the p-value. Red dots are ROIs belonging to community 26 and blue dots belong to community 27. (d) Volcano plots of the cell type combinations that were enriched in both communities. The x-axis shows the log FC2 and the y axis shows the negative log of the p-value. Red dots are ROIs belonging to community 26 and blue dots belong to community 27.

4.2 Evaluation

While the dataset of Sorin et al. contained a variety of clinical information on the tumours and the patients [10]. They have done a detailed analysis on this information and the correlations between the survival. This information has largely been discarded for the analysis done in this thesis. Conditioning on information like the stage and progression of the tumour and the amount of pack years of the patient could have a large impact on the found communities.

Applying the phenograph clustering algorithm on the neighbour vectors of the cells outputted 278 clusters. Reducing this amount to a smaller group of more archetypical communities was done by looking at the loss of information in the dataset and the retention of T regulatory cell rich communities. This was done with the aim of finding a balance between three conditions: a low information loss, a high retention of T regulatory cell rich communities and a low number of communities. However, the 'elbows' at 21, 40 and 49 in the information loss landscape are not very pronounced, thus it is unclear how good of an optimum these numbers are with respect to the previously mentioned conditions.

The neighbourhood enrichment analysis and the cellular community identification are both based on cellular neighbours found in the 25 micron circle from the center of the index cell. Thus the

enrichment of cell type A in cell type B's neighbourhood in a community indicates that cell type B is drawn towards cell type A in that community compared to other cells but not so much that this enrichment causes the cells to be assigned to a different community. Due to this shared basis, the enrichment analysis is difficult to interpret.

Another problem with the enrichment analysis is the arbitrary choice of analysing two cell types together. The presence of cellular communities suggests that there is an interplay between large and diverse group of cell types and that the interaction between two cell types cannot explain the workings of the whole. However, analysing the combinations of more than three cell types exponentially increases the number of combinations and would require more rigorous statistical safeguards against false positives. These safeguards in turn might also obfuscate any true positive results.

An attempt was made to reproduce the work in this thesis, on other datasets to allow for comparison of the found communities. However, due to differing antibody panels for the IMC and thus differing identified cell phenotypes, it was unfeasible to compare communities across datasets. Future work should coordinate antibody panels to allow for cross-referencing of communities across patient cohorts.

REFERENCES

- [1] Douglas B. Johnson, Matthew J. Riotti, and Leora Horn. Immune Checkpoint Inhibitors in NSCLC. *Current Treatment Options in Oncology*, 15(4):658–669, December 2014.
- [2] Stéphane Champiat, Ecaterina Ileana, Giuseppe Giaccone, Benjamin Besse, Giannis Mountzios, Alexander Eggermont, and Jean-Charles Soria. Incorporating Immune-Checkpoint Inhibitors into Systemic Therapy of NSCLC. *Journal of Thoracic Oncology*, 9(2):144–153, February 2014.
- [3] A. Kartolo, J. Sattar, V. Sahai, T. Baetz, and J.M. Lakoff. Predictors of immunotherapy-induced immune-related adverse events. *Current Oncology*, 25(5):e403–e410, October 2018.
- [4] Febe van Maldegem, Karishma Valand, Megan Cole, Harshil Patel, Mihaela Angelova, Sareena Rana, Emma Colliver, Katey Enfield, Nourdine Bah, Gavin Kelly, Victoria Siu Kwan Tsang, Edurne Mugarza, Christopher Moore, Philip Hobson, Dina Levi, Miriam Molina-Arcas, Charles Swanton, and Julian Downward. Characterisation of tumour microenvironment remodelling following oncogene inhibition in preclinical studies with imaging mass cytometry. *Nature Communications*, 12(1):5906, October 2021. Publisher: Nature Publishing Group.
- [5] Rachel L.G. Maus, Alexey A. Leontovich, Raymond M. Moore, Laura Becher, Wendy K. Nevala, Thomas J. Flotte, Ruifeng Guo, Jill M. Schimke, Betty A. Dickey, Yiyi Yan, and Svetomir N. Markovic. Resolving the Heterogeneous Tumor-Centric Cellular Neighborhood through Multiplexed, Spatial Paracrine Interactions in the Setting of Immune Checkpoint Blockade. *Cancer Research Communications*, 2(2):78–89, February 2022.
- [6] H. Raza Ali, Hartland W. Jackson, Vito R. T. Zanotelli, Esther Danenberg, Jana R. Fischer, Helen Bardwell, Elena Provenzano, CRUK IMAXT Grand Challenge Team, Oscar M. Rueda, Suet-Feung Chin, Samuel Aparicio, Carlos Caldas, and Bernd Bodenmiller. Imaging mass cytometry and multiplatform genomics define the phenogenomic landscape of breast cancer. *Nature Cancer*, 1(2):163–175, February 2020.
- [7] Esther Danenberg, Helen Bardwell, Vito R. T. Zanotelli, Elena Provenzano, Suet-Feung Chin, Oscar M. Rueda, Andrew Green, Emad Rakha, Samuel Aparicio, Ian O. Ellis, Bernd Bodenmiller, Carlos Caldas, and H. Raza Ali. Breast tumor microenvironment structures are associated with genomic features and clinical outcome. *Nature Genetics*, 54(5):660–669, May 2022.
- [8] Jian-Rong Li and Chao Cheng. Immune cell pair ratio captured by imaging mass cytometry has superior predictive value for prognosis of non-small cell lung cancer than cell fraction and density. *Cancer Communications (London, England)*, March 2024.
- [9] Christian M. Schüürch, Salil S. Bhate, Graham L. Barlow, Darci J. Phillips, Luca Noti, Inti Zlobec, Pauline Chu, Sarah Black, Janos Demeter, David R. McIlwain, Shigemi Kinoshita, Nikolay Samusik, Yury Goltsev, and Garry P. Nolan. Coordinated Cellular Neighborhoods Orchestrate Antitumoral Immunity at the Colorectal Cancer Invasive Front. *Cell*, 182(5):1341–1359.e19, September 2020.
- [10] Mark Sorin, Morteza Rezanejad, Elham Karimi, Benoit Fiset, Lysanne Desharnais, Lucas J. M. Perus, Simon Milette, Miranda W. Yu, Sarah M. Maritan, Samuel Doré, Émilie Pichette, William Enlow, Andréanne Gagné, Yuhong Wei, Michele Orain, Venkata S. K. Manem, Roni Rayes, Peter M. Siegel, Sophie Camilleri-Broët, Pierre Olivier Fiset, Patrice Desmeules, Jonathan D. Spicer, Daniela F. Quail, Philippe Joubert, and Logan A. Walsh. Single-cell spatial landscapes of the lung tumour immune microenvironment. *Nature*, 614(7948):548–554, February 2023. Publisher: Nature Publishing Group.
- [11] Katey S.S. Enfield, Emma Colliver, Claudia Lee, Alastair Magness, David A. Moore, Monica Sivakumar, Kristiana Grigoriadis, Oriol Pich, Takahiro Karasaki, Philip S. Hobson, Dina Levi, Selvaraju Veeriah, Clare Puttick, Emma L. Nye, Mary Green, Krijn K. Dijkstra, Masako Shimato, Ayse U. Akarca, Teresa Marafioti, Roberto Salgado, Allan Hackshaw, TRACERx consortium, Mariam Jamal-Hanjani, Febe Van Maldegem, Nicholas McGranahan, Benjamin Glass, Hanna Pulaski, Eric Walk, James L. Reading, Sergio A. Quezada, Crispin T. Hiley, Julian Downward, Erik Sahai, Charles Swanton, and Mihaela Angelova. Spatial Architecture of Myeloid and T Cells Orchestrates Immune Evasion and Clinical Outcome in Lung Cancer. *Cancer Discovery*, pages OF1–OF30, April 2024.
- [12] Katie E. Blise, Shamilee Sivagnanam, Courtney B. Betts, Konjit Betre, Nell Kirchnerberger, Benjamin J. Tate, Emma E. Furth, Andressa Dias Costa, Jonathan A. Nowak, Brian M. Wolpin, Robert H. Vonderheide, Jeremy Goecks, Lisa M. Coussens, and Katelyn T. Byrne. Machine Learning Links T-cell Function and Spatial Localization to Neoadjuvant Immunotherapy and Clinical Outcome in Pancreatic Cancer. *Cancer Immunology Research*, 12(5):544–558, May 2024.
- [13] Megan Cole, Panayiotis Anastasiou, Claudio Lee, Chris Moore, Edurne Mugarza, Martin Jones, Karishma Valand, Sareena Rana, Emma Colliver, Mihaela Angelova, Katey S. S. Enfield, Alastair Magness, Asher Mollokandov, Gavin Kelly, Tanja D. de Gruijl, Miriam Molina-Arcas, Charles Swanton, Julian Downward, and Febe van Maldegem. Spatial proteomic analysis of a lung cancer model reveals regulatory T cells attenuate KRAS-G12C inhibitor-induced immune responses, April 2024. Pages: 2024.04.11.588725 Section: New Results.
- [14] Alon Oyler-Yaniv, Jennifer Oyler-Yaniv, Benjamin M. Whitlock, Zhiduo Liu, Ronald N. Germain, Morgan Huse, Grégoire Altan-Bonnet, and Oleg Krichevsky. A Tunable Diffusion-Consumption Mechanism of Cytokine Propagation Enables Plasticity in Cell-to-Cell Communication in the Immune System. *Immunity*, 46(4):609–620, April 2017.
- [15] Dorothea Busse, Maurus de la Rosa, Kirstin Hobiger, Kevin Thurley, Michael Flossdorf, Alexander Scheffold, and Thomas Höfer. Competing feedback loops shape IL-2 signaling between helper and regulatory T lymphocytes in cellular microenvironments. *Proceedings of the National Academy of Sciences of the United States of America*, 107(7):3058–3063, February 2010.
- [16] Jacob H. Levine, Erin F. Simonds, Sean C. Bendall, Kara L. Davis, El-ad D. Amir, Michelle D. Tadmor, Oren Litvin, Harris G. Fienberg, Astraea Jager, Eli R. Zunder, Rachel Finck, Amanda L. Gedman, Ina Radtke, James R. Downing, Dana Pe'er, and Garry P. Nolan. Data-Driven Phenotypic Dissection of AML Reveals Progenitor-like Cells that Correlate with Prognosis. *Cell*, 162(1):184–197, July 2015.
- [17] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, Jake Vanderplas, Alexandre Passos, David Cournapeau, Matthieu Brucher, Matthieu Perrot, and Édouard Duchesnay. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12(85):2825–2830, 2011.
- [18] John P Klein, Brent Logan, Mette Harhoff, and Per Kragh Andersen. Analyzing survival curves at a fixed point in time. *Statistics in medicine*, 26(24):4505–4519, October 2007.
- [19] Wolf H. Friedman, Maxime Meylan, Florent Petitprez, Cheng-Ming Sun, Antoine Italiano, and Catherine Sautès-Fridman. B cells and tertiary lymphoid structures as determinants of tumour immune contexture and clinical outcome. *Nature Reviews Clinical Oncology*, 19(7):441–457, July 2022. Publisher: Nature Publishing Group.