Universiteit Leiden
The Netherlands

Visual analytics for spatially resolved omics data at single cell resolution: methods & applications

Somarakis, A.

Citation

Somarakis, A. (2022, January 20). *Visual analytics for spatially resolved omics data at single cell resolution: methods & applications*. Retrieved from https://hdl.handle.net/1887/3250550

Version:	Publisher's Version
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/3250550

Note: To cite this publication please use the final published version (if applicable).

Visual Analytics for Spatially Resolved omics data at Single Cell Resolution: Methods and Applications

Antonios Somarakis

Printed by: Gildeprint Drukkerijen

Cover design: Elena Chronopoulou

ISBN: 978-94-6419-393-0
© 2022 Antonios Somarakis, Leiden, the Netherlands Visual Analytics for spatially resolved omics data at single cell resolution: Methods and Applications

No part of this thesis may be reproduced or transmitted in any form or by any means, without the prior permission of the author.

The research described in this thesis was supported by the Leiden University Data Science Research Programme.

Visual Analytics for Spatially Resolved omics data at Single Cell Resolution: Methods and Applications

Proefschrift

ter verkrijging van de graad van
doctor aan de Universiteit Leiden,
op gezag van de rector magnificus prof. dr. ir. H. Bijl,
volgens besluit van het college voor promoties
te verdedigen op
donderdag 20 januari 2022 klokke 11:15 uur

door

Antonios Somarakis

geboren te Iraklion, Crete
in 1992

Promotores:
Prof. Dr. ir. B.P.F. Lelieveldt
Prof. Dr. ir. F.J. Verbeek
Copromotor:
Dr. T. Höllt

Leden promotiecommissie:

Prof. Dr. J.V.M.G. Bovee
Prof. Dr. A. Vilanova (Eindhoven University of Technology)
Prof. Dr. ir. M.J.T. Reinders (Technische Universiteit Delft)
Prof. Dr. K.J. Batenburg

The work presented in this thesis was carried out at the division of Image Processing (LKEB), department of Radiology at the Leiden University Medical Center.

Contents

1 Introduction **1**
1.1 Spatial omics Data 4
1.2 Preprocessing of Multiplexed Images 5
1.3 Visual Analytics 6
1.4 Contribution & Outline 9
References 12

2 Preprocessing of Highly Multiplexed Cellular Images **17**
2.1 Normalization strategies for Imaging Mass Cytometry data 18
2.2 Segmentation of microglia cells 37
References 40

3 Micro-Environment Exploration **45**
3.1 Abstract 46
3.2 Introduction 46
3.3 Related work 47
3.4 Background 49
3.5 ImaCytE 49
3.6 Case study 61
3.7 Conclusion and Future Work 67
References 68

4 Cohort Comparison **73**
4.1 Abstract 74
4.2 Introduction 74
4.3 Related Work 75
4.4 Abstraction 76
4.5 Workflow 79
4.6 Validation 86
4.7 Discussion and Conclusion 95
References 97

5 Analysis of multiplexed cellular images in Alzheimer's study **103**
5.1 Abstract 104
5.2 Introduction 104
5.3 Methods 105

5.4 Results 108
5.5 Discussion 117
5.6 Conclusion 120
5.7 Acknowledgements 120
5.8 Supplementary Information 121
References 125

6 Discussion 131
6.1 Discussion 133
References 135

Nederlandse Samenvatting 137

List of Publications 139

Curriculum Vitæ 141

Acknowledgements 143

1

Introduction

A disease is defined as “any harmful deviation from the normal structural or functional state of an organism”[1]. Organisms, though, are “republics of living elementary units ”[2].Therefore in 1839, Schwann stated that cells are the elementary parts of all tissue parts [2] defining them as “the fundamental units of living organisms” [3], thereby reducing the study of healthy and diseased tissue to the understanding of cellular functionality in the tissue.

To study cellular function, information regarding cellular properties is essential, including their molecular characteristics and position in the tissue. Hence, microscopic and molecular imaging systems are utilized to acquire information about cellular properties (e.g. protein abundance, RNA and DNA sequence) and simultaneously locate the position of the cell in the tissue. Until recently, it was either impossible to measure whole gene sequences or more than a couple of proteins at single-cell level [4]. Current ’spatial omics’ imaging modalities can measure tens of proteins [4, 5] to the whole human transcriptome [6, 7] at subcellular resolution. The detailed determination of the cellular properties at single cell resolution empowered researchers to precisely characterise the type of the cells in tissues and explore the spatial patterns they form. Both aspects are important towards the understanding of a cell’s role in tissue functionality, and to understand what goes wrong in diseased tissue.

The cell type’s heterogeneity provides information to the researchers about the complex biology in tissue , but only the exact determination of the spatial interactions completes the picture. For example, the presence of immune and cancer cells in a tissue sample exhibits an immune response to an existing tumour. Yet, the position of immune cells in relation to the tumour defines the current status of tumour as immune-excluded or immune-inhaled: this distinction is essential in treatment selection and outcome prediction in cancer. Especially now, with the current imaging modalities that offer much more information, the determination of the exact type of the immune and cancer cell interactions enables a deeper understanding of why tumours with similar status from different patients react differently to treatment. The analysis of the spatial cell type configuration in tissue is flanked by additional problems. The comparison of tissue samples originating from multiple subjects is prone to batch effects, as the different tissue preservation and acquisition protocols vary the range of acquired data among tissue samples. Also, the difficulty of boundary definition for complex cellular structures often deteriorates the quality of any finding and demands its “in situ” validation.

Most currently available methods [8, 9] are focused on the analysis of the data utilizing statistical hypothesis testing, requesting from the experts to choose in advance the specific patterns they want to test. This choice is further impeded by the novelty of the data, as the experts do not have a clear view of the main characteristics of the data and the patterns they can discover.A different approach which can incorporate gradually, in a data-driven way, the expert’s prior knowledge and cognitive skills would be more efficient. Visualizing the data, the expert can perceive its main characteristics in a more intelligible manner. In particular, Visual Analytics [10] combining visualization with data analysis techniques breaks down the identification of significant patterns to smaller tasks, where the expert can mobilize prior expertise and cognition to make more informative decisions throughout the identification process. Moreover, the expert can interactively, through the linked visual interfaces, not only explore in a stepwise manner the data, but also easier locate any finding

in the tissue and validate its quality and significance.

1.1. SPATIAL OMICS DATA

Even though the need for systematic exploration of cell contexts has been consolidated for years now [11, 12], the acquisition of images possible to capture in detail the molecular cellular required for high-resolution cell classification was not feasible. Only recently researchers managed to simultaneously identify multiple molecular properties, including DNA sequence[13], RNA transcripts[14, 15] or proteins[4, 16] alongside their spatial locations at cellular level. For example, recent state of the art acquisition methods [17] could measure simultaneously up to 10 proteins in each cell. With the methods presented in this thesis, expression information for more than 40 proteins is measured, enabling researchers to answer in much greater detail vital biological question, such as the identification of the constituent cell types in a tissue sample.

Each biological property is measured by a specific method, providing the expert the option to choose the one that is most suitable to answer a specific biological questions. Except for the biological relevance, the choice of the method is also influenced by the cost, the throughput speed and the characteristics of the tissue samples (e.g. part of the tissue, preservation method). In this thesis, applications were mostly focused on studies related to the immune system with Imaging Mass Cytometry [4]and multispectral immunofluorescence imaging [18], two relatively novel methods to image protein expression.

Imaging Mass Cytometry Data. Imaging Mass Cytometry data acquisition method [4] is a combination of mass cytometry [19], immunocytochemistry [20] (ICC) and immunohistochemistry [21] (IHC) techniques. More specifically, the acquisition starts with the labeling of tissue samples with antibodies conjugated to heavy metals following established IHC and ICC protocols. The selection of antibodies is crucial, as each of them binds to a specific tissue protein. Then, a laser ablates spot by spot every $1\mu m^2$ of the tissue. The ablated material is directed to a Mass Cytometer, where the amount of metals, and by extension of proteins, are measured for each spot. Each spot represents a pixel in the derived image. Up to now, more than 40 proteins can be measured simultaneously with Imaging Mass Cytometry.

Multispectral immunofluorescence imaging Even though Imaging Mass Cytometry offers unprecedented detail in the measurement of protein abundance, its throughput is relatively low, limiting in practice the amount and size of tissue samples that can be scanned. On the other hand, a typical cohort study contains some hundreds to thousands of tissue samples. The Vectra imaging system based on multispectral immunofluorescence (IF) technology [18] can capture up to 7 proteins simultaneously, enabling researchers to conduct high throughput studies to the detriment of the amount of the measured proteins. Also, Vectra derived images offer higher resolution, depicting in every pixel a region of $0.25\mu m^2$ of the tissue sample.

As output, both modalities provide a stack of images, where each pixel of the image contains multiple scalar values at subcellular resolution. The development of our pipeline was based on, but is not limited, to the analysis of this type of data. In principle, most parts of our pipeline can be utilized for modalities that capture various properties (e.g. tissue transcriptomics) given that they provide their cell defining characteristics at cellular resolution.

1.2. PREPROCESSING OF MULTIPLEXED IMAGES

The transformation of pixel values to reliable cellular data is an open challenge, as it entails many and complex steps. Data normalization and cell segmentation are two of the most significant.

Data Normalization. Tissue preservation and handling methods influence the signal-to-noise ratio of the acquired images. A differentiation in the tissue fixation protocol or age among two samples can vary the range of their acquired values. Hence, normalization of the data and removal of noise during tissue acquisition are essential for the combination of images in cohorts for downstream analysis. Most of the time, an offset is expressed in the image values as an almost uniform disproportionate high value. A common procedure for the normalization of these data is the manual subtraction of the offset from the pixel values so as the expression pattern to match expert's prior knowledge. However, such an approach is prone to subjective observer errors and extremely time-consuming for large cohorts of images. On the other hand, automatic approaches [22] reduce processing time and inter-expert heterogeneity, but remove important biological variance as well. In Chapter 2.1, we propose a hybrid method to deal with the previously mentioned shortcomings for Imaging Mass Cytometry data.

Cell Segmentation. Cell segmentation is one of the most difficult problems in the analysis of histopathological images. Many methods have been utilized spanning from traditional unsupervised segmentation algorithms [23, 24], over fully-supervised deep learning methods [25], to semi-supervised approaches [8]. Usually, the unsupervised methods start with the nucleus segmentation followed by their dilation to identify the cellular borders. The deep learning methods extract features from multiple dimensions of the multiplexed images providing better accuracy, given that they are trained with a large amount of annotated images. The semi-supervised approaches are based on weakly annotated data and automated machine-learning methods [26], providing the best solution for data derived from recent imaging modalities with minimum available annotated images. However, these segmentation algorithms are developed for circular shaped cellular data (Figure 1.1a). For complex cellular structures, such as microglia cells (Figure 1.1b) where their branches have been detached from the cellular body during the image acquisition procedure, specific algorithms are needed. Current approaches are either focused on the identification of the cell's skeleton [27] or nucleus [28]. In Chapter 2.2, we propose an unsupervised method for the identification of microglia cells from high-dimensional multiplexed images.

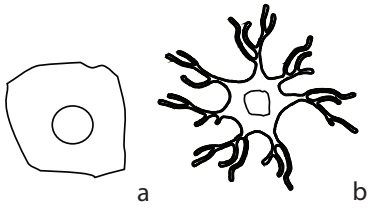

Figure 1.1: ***Cellular structures.*** *(a) A typical structure of a cell and (b) the structure of microglia cell, with the inner circural part (soma) and the many branches (processes).*

1.3. VISUAL ANALYTICS

Given the novelty of the aforementioned spatial cell phenotyping modalities, the experts do not always know in advance which patterns to expect in the data. The cellular composition of captured spatial omics images is often unknown, as the cellular characteristics that differentiate a patient and a healthy individual. When experts are neither aware of these characteristics, such as their main trends or outliers, nor the patterns they want to identify, experts should first perceive their data before starting affirming or rejecting a hypothesis. Exploratory Data Analysis [29] can play a significant role towards this direction.

Quantitative metrics can often provide a decent, but fragmented picture of data's characteristics, especially for the high-dimensional (i.e. multiple measured cellular properties) and spatial (i.e. cellular resolution) nature of such data. Anscombe's Quartet [31] is a typical example of the inability of quantitative metrics inability to provide the full spectrum of data characteristics. It contains four different two-dimensional datasets (Figure 1.2) with almost identical statistics; mean value for x and y, variance for x and y, correlation among x and y, linear regression line and coefficient. Their visual representations (Figure 1.2), though, illustrate four different datasets with discrete inherent characteristics, highlighting the definition given from Pickover and Tewksbury [32] for visualization, as the "*the art and science of making the unseen workings of nature visible*".

Therefore, we integrate visualization techniques in the exploratory analysis of our complex, novel data, as they are described in Chapters 3 and 4 to enable the expert "*to analyze data when they don't know exactly what questions they need to ask in advance*" [33].

Table 1.1: ***Anscombe's quartet*** *includes four datasets with identical statistics with 2 decimal points accuracy.*

	I		II		III		IV	
Observation	*x*	*y*	*x*	*y*	*x*	*y*	*x*	*y*
1	10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
2	8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
3	13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
4	9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
5	11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
6	14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
7	6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
8	4.0	4.26	4.0	3.10	4.0	5.39	19.0	12.50
9	12.0	10.84	12.0	9.13	12.0	8.15	8.0	5.56
10	7.0	4.82	7.0	7.26	7.0	6.42	8.0	7.91
11	5.0	5.68	5.0	4.74	5.0	5.73	8.0	6.89
Summary Statistics								
Mean	9	7.50	9	7.50	9	7.50	9	7.50
Variance	11	4.13	11	4.13	11	4.13	11	4.13
Correlation X, Y	0.82		0.82		0.82		0.82	
Linear Regression	*y = 3.00 + 0.500x*		*y = 3.00 + 0.500x*		*y = 3.00 + 0.500x*		*y = 3.00 + 0.500x*	

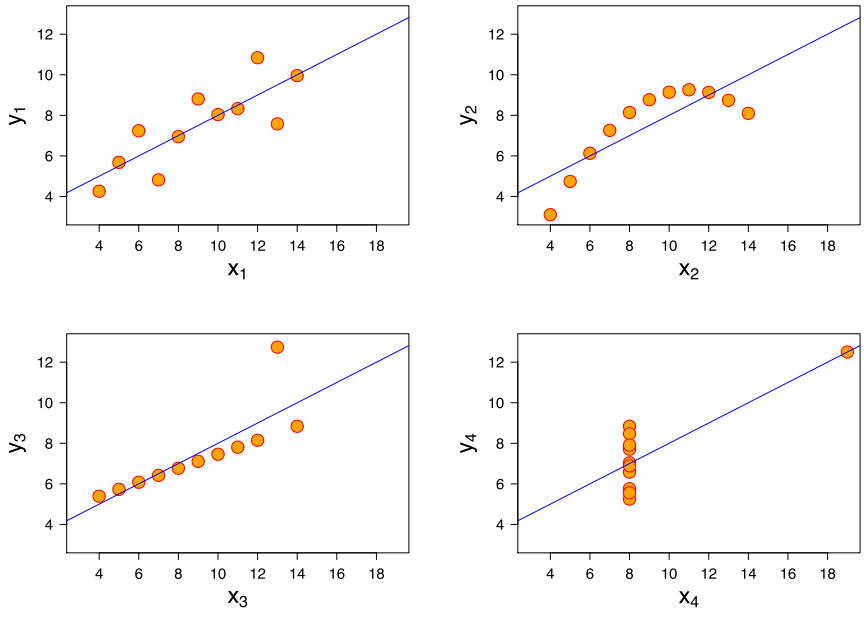

Figure 1.2: ***Anscombe's quartet*** *visual representation in the x-y space shows a significant difference from the above illustarted decriptive statistics. (Image Source: Wikipedia's Lemma [30])*

A significant factor of visualization efficiency is the expert that is performing the specific exploratory task. The identification of the number of clusters and outliers in a two-dimensional scatter plot is performed much more efficiently by a human than a computer. This can exemplify the functioning of visualization; a visual representation is designed to represent data in a way that can facilitate the expert to employ internal cognition and memory usage (i.e. points that are close together form a cluster, points that deviate from the group and are not close to another group are outliers) to gain a deeper understanding of the data (i.e. clusters) [33]. A vital part of the whole visualization process plays the interaction among the expert and computing system (Figure 1.3) throughout this cognitive process. This feedback mechanism is taking place through the interactivity among humans and computers, enabling the expert - on the fly - to explore multiple data attributes.

However, it is impossible for the expert to explore raw cellular properties values or the total amount of spatial cellular combinations merely through interactive visualizations, due to their large amount. To that end, the utilization of automated data analysis techniques alongside interactive visualization systems is deemed necessary. The discipline that studies such systems is called Visual Analytics [10, 34]. In addition to that, there are numerous problems for which research questions can not be defined from the onset of their analysis. Hence, for such ill-defined problems automated solutions fall short Visual Analytics solutions can have substantial contribution by integrating analysis with hypothesis generation [35]. The rationale behind Visual Analytics, as illustrated in Figure 1.3, is based on the combination of the processing power of automated analysis techniques to deal with large amounts of data with the human's ability for analytical reasoning through interactive

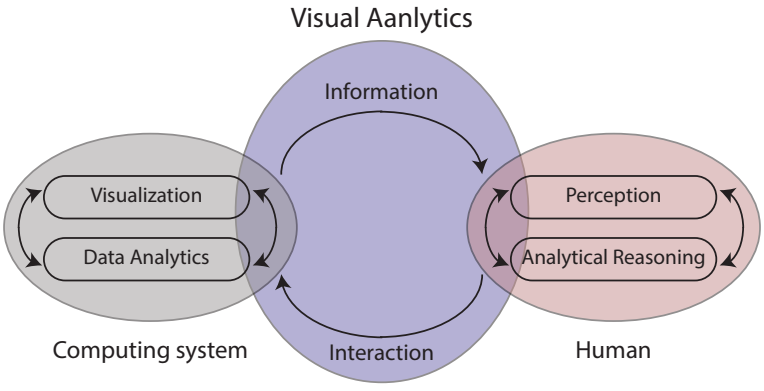

Figure 1.3: ***Visual Analytics process.*** *Starting from the creation of visualizations utilizing data analytics methods, provides insights to the expert. The expert, through her/his analytical reasoning, perceive the insights and interacts accordingly with the system to continue the data exploration.*

visualizations. The role of the experts (i.e., humans) in such a system is pivotal, as they interactively steer the data exploration [36].

Visual Analytics systems are being used in various scientific domains to facilitate the analysis of large amounts of complex data. Besides the exploration of tissue images, Visual Analytics is being used in Medicine for the exploration of public health data [37]. Moreover, Visual Analytics systems are being used for the exploration of the most efficient stock trading algorithm [38], exploration of high resolution remote sensing imagery for the support of precision agriculture [39] or the similarity exploration of texts [40].

Even though Visual Analytics is a valuable tool for the analysis of complex problems in large information spaces, it does not come without restrictions. Modern applications (e.g. more sensitive sensors, large scale scientific experiments) has resulted in the accumulation of enormous amounts of data. Following the increase of available data, algorithms have evolved to live up with the analysis requirements. Nonetheless, not all of these algorithms are appropriate to serve the main Visual Analytics tasks, namely the illustration of high-level abstractions synchronously with low-level details of the data [41]. Also, the combination of heterogeneous types of data is another limiting factor for many Visual Analytics applications. For scientific domains where multiple different data sources are combined is challenging for Visual Analytics approaches to create effective solutions. Furthermore, many restrictions are imposed from the involvement of the user in Visual Analytics approaches; the necessity for users to supervise 24/7 applications which include data streams and require their immediate response and supervision, or the required training in order to handle properly Visual Analytics tool and interpret any findings or results. Finally, a serious challenge which leads many times to the disuse of Visual Analytics tools is their limited compatibility with existing systems that the users utilize in their daily routine.

From the aforementioned restrictions and challenges it is evident that proper evaluation of the Visual Analytics systems should be performed. According to Keim et al. [10], Visual Analytics is both a science and a technique and thus should be evaluated based on

effectiveness, *efficiency* and *user satisfaction*. In other words, a Visual Analytics system should be evaluated based on its ability to enable users to fulfill the tasks that originally the system has been developed for, given a finite amount of resources. The evaluation includes the data, the tasks definition, the stakeholders and the Visual Analytics system itself. The evaluation of the data should take into account the heterogeneity and suitability of the data to answer the research questions. The defined tasks should be evaluated according to their complexity. The evaluation of the stakeholders should not be limited to the users, but also expanded to the the developers of the system. The evaluation of the Visual Analytics systems should include the graphical representations, the suitability and efficacy of the selected technologies. In general, through the evaluation process it is easier to deduct conclusions about precise well-defined problems (e.g. the efficacy of a visual analytics system to cluster the data from a two-dimensional scatter plot) than generic findings (e.g. the amount of saved time if we use one technique in favor of one other). Çöltekin et al. [42] in their research are trying to discover such generic conclusions recording and afterwards analysing the movement of the user's eyes. At the end of the day, proper evaluation of the Visual Analytics systems can enable the comparison of existing approaches and eventually lead to development of innovative methods and techniques.

1.4. CONTRIBUTION & OUTLINE

The main contribution of our work is the development of an end-to-end exploratory data analysis methodology , (Figure 1.4) addressing the main challenges researchers face during the analysis of highly multiplexed tissue images at cellular level, in close cooperation with the clinical researchers from Leiden University Medical Centre (LUMC). A birds-eye overview of our pipeline is described in our previous work [43]. More specifically, our pipeline covers the preprocessing of the data (Figure 1.4a), including the removal of the biological irrelevant intra-cohort heterogeneity and the segmentation of complex cellular structures, preparing the data for the main task of the analysis, the exploration of spatial cellular patterns. Then, we enable the researchers to visually explore the data, through two interactive and data-driven Visual Analytics frameworks; ImaCytE [44] (Figure 1.4b), for the identification of spatial cellular patterns in tissue samples and SpaCeCo [45] (Figure 1.4c), for the comparison of distinct clinical cohorts of patients. In particular:

- A workflow focused on the normalization of cohorts of images derived from Imaging Mass Cytometry which entails the semi-supervised classification of pixels, as either actual signal or background and the assignment of a cell's expression value, as the fraction of actual signal pixels that belong to its segmented area.

- An automated algorithm for the segmentation of microglia, a complex cellular structure with a main circular part and many branches, based on its nucleus and membrane expression markers utilizing an advanced level-set-based approach [46].

- ImaCytE; A Visual Analytics framework, which was developed for the end-to-end, in depth, analysis of individual images derived from Imaging Mass Cytometry, but can be applied to any spatial omics modality offering cellular resolution. The main novelty of our framework is not only the interactive identification of cohesive cell types, but through motifs the stratification of the cell types into subgroups with

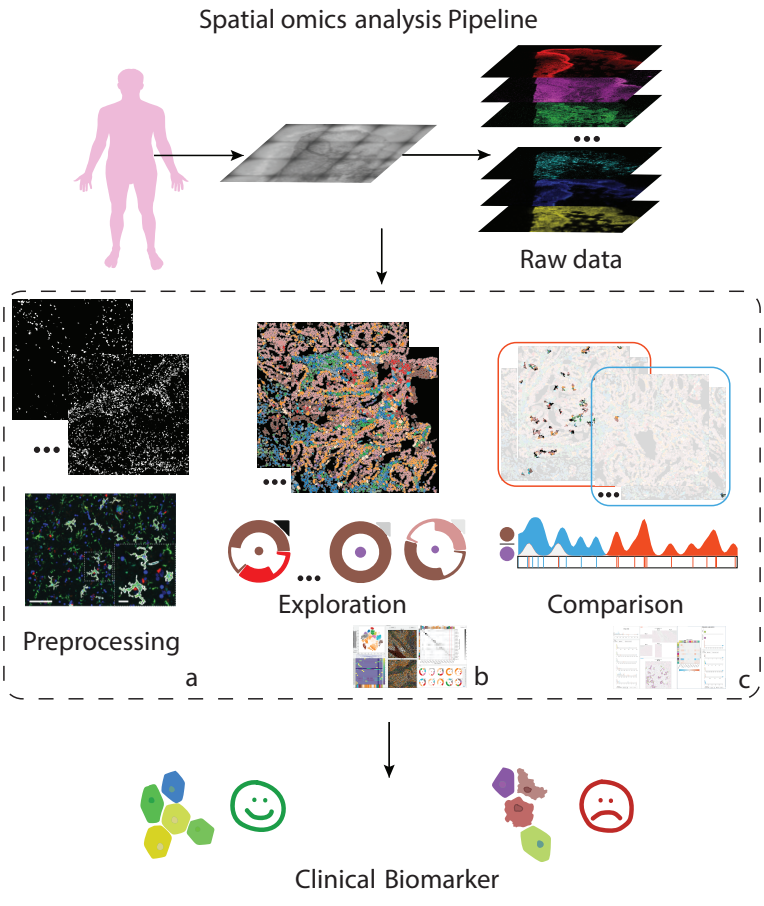

Figure 1.4: *An end-to-end pipeline for the analysis of highly multiplexed tissue images at cellular resolution, covering from (a) the preprocessing over (b) the exploration to (c) the comparison of the data.*

unique microenvironment characteristics and their inspection interactively in the tissue.

- SpaCeCo; A Visual Analytics tool, for the comparison of cohorts with labelled cellular images, based on both their cell types and the spatial patterns they form. Moreover, the tool enables in every step of the comparison the identification of outliers in each cohort and location of any finding in the tissue.

The dissertation continues with Chapter 2, which presents an overview of the preprocessing part of our pipeline, including the data normalization and the segmentation of complex cellular structures. Chapter 3 and 4 form the core part of our analysis pipeline and explain in depth the concepts and functionality of our two Visual Analytics frameworks; ImaCytE and SpaCeCo. Chapter 5 outlines a study in Alzheimer's disease, where major

parts of our pipeline have been utilized for discovery research. Finally, Chapter 6 summarizes this work and reflects on the lessons learnt during this research and upon directions for future work.

REFERENCES

[1] D. G. Scarpelli and W. Burrows, "Disease's definition." https://www.britannica.com/science/disease, 1976–98. [Online; Accessed: 2021-07-14].

[2] P. Mazzarello, "A unifying concept: the history of cell theory," *Nature Cell Biology*, vol. 1, no. 1, pp. E13–E15, 1999.

[3] A. Regev, S. A. Teichmann, E. S. Lander, I. Amit, C. Benoist, E. Birney, B. Bodenmiller, P. Campbell, P. Carninci, M. Clatworthy, *et al.*, "Science forum: the human cell atlas," *Elife*, vol. 6, p. e27041, 2017.

[4] C. Giesen, H. A. Wang, D. Schapiro, N. Zivanovic, A. Jacobs, B. Hattendorf, P. J. Schüffler, D. Grolimund, J. M. Buhmann, S. Brandt, Z. Varga, P. J. Wild, D. Günther, and B. Bodenmiller, "Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry," *Nature Methods*, vol. 11, no. 4, pp. 417–422, 2014.

[5] Y. Goltsev, N. Samusik, J. Kennedy-Darling, S. Bhate, M. Hale, G. Vazquez, S. Black, and G. P. Nolan, "Deep profiling of mouse splenic architecture with codex multiplexed imaging," *Cell*, vol. 174, no. 4, pp. 968–981, 2018.

[6] R. Ke, M. Mignardi, A. Pacureanu, J. Svedlund, J. Botling, C. Wählby, and M. Nilsson, "In situ sequencing for rna analysis in preserved tissue and cells," *Nature Methods*, vol. 10, no. 9, pp. 857–860, 2013.

[7] J. H. Lee, E. R. Daugharthy, J. Scheiman, R. Kalhor, T. C. Ferrante, R. Terry, B. M. Turczyk, J. L. Yang, H. S. Lee, J. Aach, K. Zhang, and G. M. Church, "Fluorescent in situ sequencing (FISSEQ) of rna for gene expression profiling in intact cells and tissues," *Nature Protocols*, vol. 10, no. 3, pp. 442–458, 2015.

[8] D. Schapiro, H. W. Jackson, S. Raghuraman, J. R. Fischer, V. R. T. Zanotelli, D. Schulz, C. Giesen, R. Catena, Z. Varga, and B. Bodenmiller, "histoCAT: analysis of cell phenotypes and interactions in multiplex image cytometry data," *Nat Methods*, vol. 14, no. 9, pp. 873–876, 2017.

[9] Y. Yuan, H. Failmezger, O. M. Rueda, H. Raza Ali, S. Gräf, S. F. Chin, R. F. Schwarz, C. Curtis, M. J. Dunning, H. Bardwell, N. Johnson, S. Doyle, G. Turashvili, E. Provenzano, S. Aparicio, C. Caldas, and F. Markowetz, "Quantitative image analysis of cellular heterogeneity in breast tumors complements genomic profiling," *Science Translational Medicine*, vol. 4, no. 157, pp. 157ra143–157ra143, 2012.

[10] D. Keim, J. Kohlhammer, G. Ellis, and F. Mansmann, "Mastering the information age: solving problems with visual analytics," Goslar: Eurographics Association, 2010.

[11] J. Eberwine, H. Yeh, K. Miyashiro, Y. Cao, S. Nair, R. Finnell, M. Zettel, and P. Coleman, "Analysis of gene expression in single live neurons," *Proceedings of the National Academy of Sciences*, vol. 89, no. 7, pp. 3010–3014, 1992.

[12] R. N. Van Gelder, M. E. von Zastrow, A. Yool, W. C. Dement, J. D. Barchas, and J. H. Eberwine, "Amplified rna synthesized from limited quantities of heterogeneous cdna," *Proceedings of the National Academy of Sciences*, vol. 87, no. 5, pp. 1663–1667, 1990.

[13] A. McKenna, G. M. Findlay, J. A. Gagnon, M. S. Horwitz, A. F. Schier, and J. Shendure, "Whole-organism lineage tracing by combinatorial and cumulative genome editing," *Science*, vol. 353, no. 6298, 2016.

[14] K. H. Chen, A. N. Boettiger, J. R. Moffitt, S. Wang, and X. Zhuang, "Spatially resolved, highly multiplexed rna profiling in single cells," *Science*, vol. 348, no. 6233, 2015.

[15] S. Shah, E. Lubeck, W. Zhou, and L. Cai, "In situ transcription profiling of single cells reveals spatial organization of cells in the mouse hippocampus," *Neuron*, vol. 92, no. 2, pp. 342–357, 2016.

[16] M. Angelo, S. C. Bendall, R. Finck, M. B. Hale, C. Hitzman, A. D. Borowsky, R. M. Levenson, J. B. Lowe, S. D. Liu, S. Zhao, *et al.*, "Multiplexed ion beam imaging of human breast tumors," *Nature medicine*, vol. 20, no. 4, pp. 436–442, 2014.

[17] M. J. Gerdes, C. J. Sevinsky, A. Sood, S. Adak, M. O. Bello, A. Bordwell, A. Can, A. Corwin, S. Dinn, R. J. Filkins, *et al.*, "Highly multiplexed single-cell analysis of formalin-fixed, paraffin-embedded cancer tissue," *Proceedings of the National Academy of Sciences*, vol. 110, no. 29, pp. 11982–11987, 2013.

[18] M. E. Ijsselsteijn, T. P. Brouwer, Z. Abdulrahman, E. Reidy, A. Ramalheiro, A. M. Heeren, A. Vahrmeijer, E. S. Jordanova, and N. F. de Miranda, "Cancer immunophenotyping by seven-colour multispectral imaging without tyramide signal amplification," *The Journal of Pathology: Clinical Research*, vol. 5, pp. 3–11, 2019.

[19] D. R. Bandura, V. I. Baranov, O. I. Ornatsky, A. Antonov, R. Kinach, X. Lou, S. Pavlov, S. Vorobiev, J. E. Dick, and S. D. Tanner, "Mass cytometry: Technique for real time single cell multitarget immunoassay based on inductively coupled plasma time-of-flight mass spectrometry," *Analytical Chemistry*, vol. 81, no. 16, pp. 6813–6822, 2009.

[20] M. Bibbo and D. Wilbur, *Comprehensive Cytopathology E-Book*. Elsevier Health Sciences, 2014.

[21] J. Seuma, J. Bunch, A. Cox, C. McLeod, J. Bell, and C. Murray, "Combination of immunohistochemistry and laser ablation icp mass spectrometry for imaging of cancer biomarkers," *Proteomics*, vol. 8, no. 18, pp. 3775–3784, 2008.

[22] J.-P. Fortin, N. Cullen, Y. I. Sheline, W. D. Taylor, I. Aselcioglu, P. A. Cook, P. Adams, C. Cooper, M. Fava, P. J. McGrath, *et al.*, "Harmonization of cortical thickness measurements across scanners and sites," *Neuroimage*, vol. 167, pp. 104–120, 2018.

[23] S. Beucher, "Use of watersheds in contour detection," in *Proceedings of the International Workshop on Image Processing*, CCETT, 1979.

[24] S. Osher and J. A. Sethian, "Fronts propagating with curvature-dependent speed: Algorithms based on hamilton-jacobi formulations," *Journal of Computational Physics*, vol. 79, no. 1, pp. 12–49, 1988.

[25] N. F. Greenwald, G. Miller, E. Moen, A. Kong, A. Kagel, C. C. Fullaway, B. J. McIntosh, K. Leow, M. S. Schwartz, T. Dougherty, C. Pavelchek, S. Cui, I. Camplisson, O. Bar-Tal, J. Singh, M. Fong, G. Chaudhry, Z. Abraham, J. Moseley, S. Warshawsky, E. Soon, S. Greenbaum, T. Risom, T. Hollmann, L. Keren, W. Graf, M. Angelo, and D. Van Valen, "Whole-cell segmentation of tissue images with human-level performance using large-scale data annotation and deep learning," *bioRxiv*, 2021.

[26] S. Berg, D. Kutra, T. Kroeger, C. N. Straehle, B. X. Kausler, C. Haubold, M. Schiegg, J. Ales, T. Beier, M. Rudy, *et al.*, "Ilastik: interactive machine learning for (bio) image analysis," *Nature Methods*, vol. 16, no. 12, pp. 1226–1232, 2019.

[27] M. Abdolhoseini, M. G. Kluge, F. R. Walker, and S. J. Johnson, "Segmentation, tracing, and quantification of microglial cells from 3d image stacks," *Scientific reports*, vol. 9, no. 1, pp. 1–10, 2019.

[28] PerkinElmer, "inForm Advanced Image Analysis Software - For Accurately Quantifying Biomarkers in Tissue Sections." https://tinyurl.com/y2ckm848, 2012. [Online; Accessed: 2021-07-14].

[29] J. W. Tukey, "The Future of Data Analysis," *The Annals of Mathematical Statistics*, vol. 33, no. 1, pp. 1 – 67, 1962.

[30] Wikipedia contributors, "Anscombe's quartet — Wikipedia, the free encyclopedia." https://en.wikipedia.org/wiki/Anscombe [Online; Accessed 2021-06-10].

[31] F. J. Anscombe, "Graphs in statistical analysis," *The American Statistician*, vol. 27, no. 1, pp. 17–21, 1973.

[32] C. A. Pickover and S. K. Tewksbury, *Frontiers of scientific visualization*. John Wiley & Sons, Inc., 1994.

[33] T. Munzner, *Visualization analysis and design*. A K Peters/CRC press, 2014.

[34] A. Endert, M. S. Hossain, N. Ramakrishnan, C. North, P. Fiaux, and C. Andrews, "The human is the loop: new directions for visual analytics," *Journal of intelligent information systems*, vol. 43, no. 3, pp. 411–435, 2014.

[35] D. A. Keim, F. Mansmann, and J. Thomas, "Visual analytics: How much visualization and how much analytics?," *SIGKDD Explor. Newsl.*, vol. 11, p. 5–8, may 2010.

[36] R. Raidou, *Visual analytics for digital radiotherapy : towards a comprehensible pipeline*. PhD thesis, Technische Universiteit Eindhoven, 2017.

[37] B. Preim, P. Klemm, H. Hauser, K. Hegenscheid, S. Oeltze, K. Toennies, and H. Völzke, "Visual analytics of image-centric cohort studies in epidemiology," in *Visualization in Medicine and Life Sciences III. Mathematics and Visualization.* (L. Linsen, B. Hamann, and H. C. Hege, eds.), pp. 221–248, Springer, 2016.

[38] K. Tsang, H. Li, F. Lam, Y. Mu, Y. Wang, and H. Qu, "Tradao: A visual analytics system for trading algorithm optimization," in *Proceedings of the 32nd IEEE Visualization Conference (IEEE VIS)*, pp. 61–65, 2020.

[39] M. P. Wachowiak, D. F. Walters, J. M. Kovacs, R. Wachowiak-Smolíková, and A. L. James, "Visual analytics and remote sensing imagery to support community-based research for precision agriculture in emerging areas," *Computers and Electronics in Agriculture*, vol. 143, pp. 149–164, 2017.

[40] A. Abdul-Rahman, G. Roe, M. Olsen, C. Gladstone, R. Whaling, N. Cronk, R. Morrissey, and M. Chen, "Constructive visual analytics for text similarity detection," in *Computer Graphics Forum*, vol. 36, pp. 237–248, Wiley Online Library, 2017.

[41] D. A. Keim, F. Mansmann, J. Schneidewind, and H. Ziegler, "Challenges in visual data analysis," in *Tenth International Conference on Information Visualisation (IV'06)*, pp. 9–16, IEEE, 2006.

[42] A. Çöltekin, S. Fabrikant, and M. Lacayo, "Exploring the efficiency of users' visual analytics strategies based on sequence analysis of eye movement recordings," *International Journal of Geographical Information Science*, vol. 24, no. 10, pp. 1559–1575, 2010.

[43] A. Somarakis, M. E. Ijsselsteijn, B. Kenkhuis, V. v. Unen, S. J. Luk, F. Koning, L. v. d. Weerd, N. F. C. C. d. Miranda, B. P. F. Lelieveldt, and T. Höllt, "Visual Analysis of Tissue Images at Cellular Level," in *EuroVis 2021 - Dirk Bartz Prize* (S. Oeltze-Jafra and R. G. Raidou, eds.), The Eurographics Association, 2021.

[44] A. Somarakis and T. Höllt, "ImaCytE open source software." https://www.doi.org/10.5281/zenodo.3345951, 2019. [Online; Accessed: 2021-07-14].

[45] A. Somarakis and T. Höllt, "SpaCeCo open source software." https://www.doi.org/10.5281/zenodo.3885814, 2020. [Online; Accessed: 2021-07-14].

[46] O. Dzyubachyk, W. A. Van Cappellen, J. Essers, W. J. Niessen, and E. Meijering, "Advanced level-set-based cell tracking in time-lapse fluorescence microscopy," *IEEE Transactions on Medical Imaging*, vol. 29, no. 3, pp. 852–867, 2010.

2

Preprocessing of Highly Multiplexed Cellular Images

2.1. Normalization strategies for Imaging Mass Cytometry data

This chapter was adapted from:

M. E. Ijsselsteijn*, **A. Somarakis***, B. Lelieveldt, T. Höllt+, and N. F. de Miranda++, "Semi-automated background removal limits data loss and normalizes imaging mass cytometry data," *Cytometry Part A*, vol. 99, no. 12, pp. 1187–1197, 2021.

*,+ Contributed equally

2.1.1. ABSTRACT

Imaging Mass Cytometry (IMC) allows the detection of multiple antigens (approximately 40 markers) combined with spatial information, making it a unique tool for the evaluation of complex biological systems. Due to its widespread availability and retained tissue morphology, formalin-fixed, paraffin-embedded (FFPE) tissues are often a material of choice for IMC studies. However, antibody performance and signal to noise ratios can differ considerably between FFPE tissues as a consequence of variations in tissue processing, including fixation. In contrast to batch effects caused by differences in the immunodetection procedure, variations in tissue processing are difficult to control. We investigated the effect of immunodetection-related signal intensity fluctuations on IMC analysis and phenotype identification, in a cohort of twelve colorectal cancer tissues. Furthermore, we explored different normalisation strategies and propose a workflow to normalise IMC data by semi-automated background removal, using publicly available tools. This workflow can be directly applied to previously acquired datasets and considerably improves the quality of IMC data, thereby supporting the analysis and comparison of multiple samples.

2.1.2. INTRODUCTION

Mass cytometry has advanced as an important technology for the characterisation of cellular contextures in health and disease [1–6]. A major advantage of mass cytometry is its ability to simultaneously interrogate over 40 markers. The high-level of multiplexing is made possible via the use of antibodies conjugated to heavy metal isotopes rather than fluorescent tags [7]. Cells are labelled with these and led into a CyTOF (Cytometry by time-of-flight) instrument, where heavy metal abundance is measured, per cell, by time-of-flight mass spectrometry [8]. Technological advancements in the field have made it possible to image tissue sections as opposed to single cells, allowing for the incorporation of spatial information [9]. Imaging Mass Cytometry (IMC) allows the analysis of, among others, archival tissue samples in the form of formalin-fixed paraffin-embedded (FFPE) or snap-frozen (FF) tissue. Tissue sections are labelled with metal-conjugated antibodies and ablated in small portions (typically $1\mu m^2$ = 1 pixel). The ablated tissue is then analysed with the CyTOF instrument. The pixel data is processed into an image, thereby allowing the visualization of phenotypes and incorporation of spatial information in subsequent analyses. IMC users have already contributed with a number of studies aimed at optimising the use of this technology, including: a strategy to address signal spill-over during heavy metal detection [10] as well as methodologies to aid the implementation of large antibody panels for FFPE [11] or snap-frozen [12] tissues. Schulz and colleagues demonstrated the potential of combining protein and RNA in situ detection with IMC [13]. Furthermore, IMC has been used to comprehensively study tissue architectures and cellular composition of breast cancers [14] and pancreatic tissues affected by type 1 diabetes [15, 16], among other applications. The increasingly widespread application of IMC for the characterisation of tissues is accompanied by the need to develop analytical tools that can handle large and complex datasets where, for instance, signal to noise ratio fluctuates across samples. The general pipeline for IMC analysis involves the creation of cell segmentation masks with ilastik [17] and CellProfiler [18], after which the resulting image-stacks and masks are processed by dedicated software packages like HistoCAT [19] or ImaCytE [20]. The majority of current IMC studies make use of FFPE tissues, due to their widespread

availability in tissue archives and good morphology after fixation. For the interpretation of immunohistochemistry data on FFPE tissues, it has long been recognised that antibody performance and signal detection can vary considerably between specimens. This can be explained by the use of different fixation times, size of tissue during fixation, dehydration of the tissue after fixation, the age of the FFPE tissue block or how long the tissue slides have been stored before immunodetection [21–24]. Moreover, particularly impactful and difficult to control, is the ischemia period that concerns the time between the collection of a tissue and its fixation. Ischemia can cause a number of artefacts due to autolysis, protein degradation, or the drying of the outer layer of the tissue [22–26]. Therefore, the comparison of intensities of antibody signal between different FFPE tissues is not general practice in the evaluation of immunohistochemistry results. In this work we investigated three methods for the processing of IMC data. We first analysed an IMC dataset without preprocessing and compared this to two normalisation strategies: background identification to correct for variations in signal intensity and background between tissues, using manual thresholding or a semi-automated method. Both approaches were followed by per-pixel binarization of marker intensity to overcome differences in immunodetection intensity between tissues. After comparing the three approaches we propose a workflow for the analysis of tissues that makes use of publicly available tools to generate processed IMC data. Importantly, we implemented a normalisation strategy that overcomes immunodetection intensity variations across samples and considerably improves the quality of IMC data.

2.1.3. Methodology

2.1.3.1. Tissue Material

FFPE blocks from 12 colorectal cancers were obtained from the department of Pathology of the Leiden University Medical Centre (Leiden, The Netherlands). Samples were anonymised and handled according to the medical ethical guidelines described in the Code of Conduct for Proper Secondary Use of Human Tissue of the Dutch Federation of Biomedical Scientific Societies. Colorectal cancer tissues were cut into 4 µm sections and placed on silane-coated glass slides (VWR, Radnor, PA, USA).

2.1.3.2. Imaging Mass Cytometry immunodetection and acquisition

Antibodies employed in this study were conjugated to purified lanthanide metals using the Maxpar antibody labelling kit and protocol (Fluidigm, San Francisco, CA, USA). Antibodies were eluted in 50 µl antibody stabilizer solution (Candor Bioscience, Wangen im Allgäu, Germany) supplemented with 0,05 % sodium azide and 50 µl W-buffer (Fluidigm). After conjugation, all antibodies were tested by IHC on 4 µm tonsil tissue to confirm that the labelling process did not affect antibody performance. IMC immunodetection was performed following the methodology published previously by our lab [11] using the antibodies and conditions described in supplementary table 1. Tissue sections were ablated within a week after immunodetection by using the Hyperion mass cytometry imaging system (Fluidigm). The Hyperion was autotuned using a 3-element tuning slide (Fluidigm) as described in the Hyperion imaging system user guide. In addition to the successful tuning requirements of the Hyperion imaging system, a minimum detection of 1500 mean duals of 175Lu was required, to control for variations in the plasma-line positioning. Four 1000x1000 µm regions of interest per sample were selected based on haematoxylin and

eosin (H&E) stains on consecutive slides and ablated at 200 Hz. Data was exported as MCD files and txt files and visualised using the Fluidigm MCD viewer. For downstream analysis, the MCD files were transformed to either 32-bit multi-tiff or single-marker tiff images in the MCD viewer.

2.1.3.3. CREATION OF SINGLE CELL MASKS

For each sample, one tiff image was exported from the MCD viewer, combining the keratin and vimentin expression as well as DNA detection. ilastik [17] (v1.3.3) was used to create masks for nuclei (based on the DNA signal), cytoplasm/membrane (based on keratin and vimentin expression) and background (based on the absence of signal in the DNA, keratin and vimentin image). ilastik's random forest classifier was trained using manually assigned pixels that underwent Gaussian smoothing (ilastik feature settings: 0.3, 0.7 and 1.0 sigma for colour/intensity, edge and texture). Training was performed on 12 images (one representative image per sample) after which the classifier was applied to all images in the dataset and data was exported as probability maps indicating the likelihood of each pixel corresponding to nucleus, cytoplasm/membrane or background. In CellProfiler [18] (v2.2.0) the probability maps were used to create single cell masks for all samples. All masks were compared to the original IMC images to validate the cell segmentation procedure.

2.1.3.4. BACKGROUND IDENTIFICATION AND BINARIZATION

To address variations in immunodetection signal intensity between samples, two normalisation approaches were applied: 1) manual background identification in MCD viewer or 2) semi-automated background identification in ilastik, both followed by binarization of pixel values.

- Manual background identification was done using the MCD viewer by inspecting each marker and setting a minimum intensity/mean duals threshold to remove background noise. The threshold was identified by visual inspection, based on the user's knowledge of the expected immunodetection pattern and corresponding IHCs of the protein in question. After setting a threshold for each marker, the data was saved as txt files containing all previously defined thresholds. This process was repeated for all images. Together with the multi-tiff images and the cell segmentation masks, the threshold txt files were loaded into ImaCytE. In ImaCytE, the thresholds were applied to the images and pixel intensity values were binarized (i.e., all pixels below the threshold were set to 0 and all pixels above threshold were set to 1). Normalised cell intensities were then defined as the frequency of positive pixels, per cell.

- Semi-automated background identification was done on single marker tiff images, exported from MCD viewer. The images corresponding to a single marker across the entire cohort were loaded into ilastik and a small amount (i.e., approximately 1 %) of pixels were assigned to either 'signal' or 'background' in 12 images (one representative image per sample). To facilitate pixel annotation, outliers were removed from the images through saturation of all pixels with values lower than the 1st and higher than the 99th percentile using MATLAB. Then, after Gaussian

smoothing (ilastik feature settings: 0.7, 1.0 and 1.6 sigma for colour/intensity, edge and texture), the random forest classifier automatically classified signal and background pixels. After training on at least 12 images (1 for each sample), the classifier was applied to all images in the dataset and the data was exported as binary expression maps with the 'background' pixels set to 0 and the 'signal' pixels set to 1. This approach was repeated for each marker. A folder was created for each image containing the binary expression maps of all markers and the previously created cell segmentation masks. These were loaded into ImaCytE and for each marker the relative frequency of positive pixels in a cell was visualised on the cell mask.

2.1.3.5. SINGLE CELL CLUSTERING AND PHENOTYPE CALLING

Single cell data was obtained by processing cell segmentation masks with their corresponding pixel intensity files in ImaCytE, for the generation of FCS files. For the analyses without preprocessing or with manual background identification, multi-tiff images and their corresponding cell segmentation masks were employed. For the analysis with semi-automated background identification, segmentation masks were loaded together with binary expression maps of each marker. Single-cell FCS files containing mean pixel values per cell (for the non-normalised dataset) or relative frequency of positive pixels per cell (for the normalised dataset) were then exported from ImaCytE and analysed by t-SNE [27] (t-distributed stochastic neighbor embedding) in Cytosplore [28]. Cells forming visual neighbourhoods in the t-SNE embedding were grouped using Mean-shift clustering and exported as separate FCS files. The resulting subsets were imported back in ImaCytE for visualisation of subsets in the segmentation masks and localisation was compared to original MCD images to validate the obtained clusters.

2.1.4. RESULTS

2.1.4.1. VARIATION IN SIGNAL TO NOISE RATIO BETWEEN FFPE SAMPLES INFLUENCES UNSUPERVISED IMAGE ANALYSIS

Immunodetection in FFPE tissues is complicated by variations in antigen availability and accessibility across samples due to tissue processing and fixation procedures. To understand its implications to the quality of IMC data, we analysed 48 images generated from 12 colorectal cancer samples. FFPE tissues were labelled with a 30 antibody panel and four regions of interest (ROI) of 1 mm2 were ablated, per tumour, by the Hyperion imaging mass cytometer. CD45 and CD4 were excluded from further analysis due to poor signal detection. Further visual inspection of the images using the MCD viewer showed that large differences exist in immunodetection intensity of the same antigen between tissues, which cannot always be explained by biological variation (Figure 2.1a). To determine the impact of these fluctuations on downstream analyses, the added-value of two normalisation approaches was investigated in comparison to the IMC analysis pipeline without preprocessing (Figure 2.1b). In short, cell masks were created using ilastik and CellProfiler and were loaded into ImaCytE combined with the raw or normalised images in order to define relative marker expression per cell. FCS files were produced, and clustering of cells was performed by t-SNE to identify cell subsets, using Cytosplore. Next, the phenotypes were projected back onto the cell masks in ImaCytE for visualisation and spatial analysis. First, we visualised immunodetection signal intensities, without

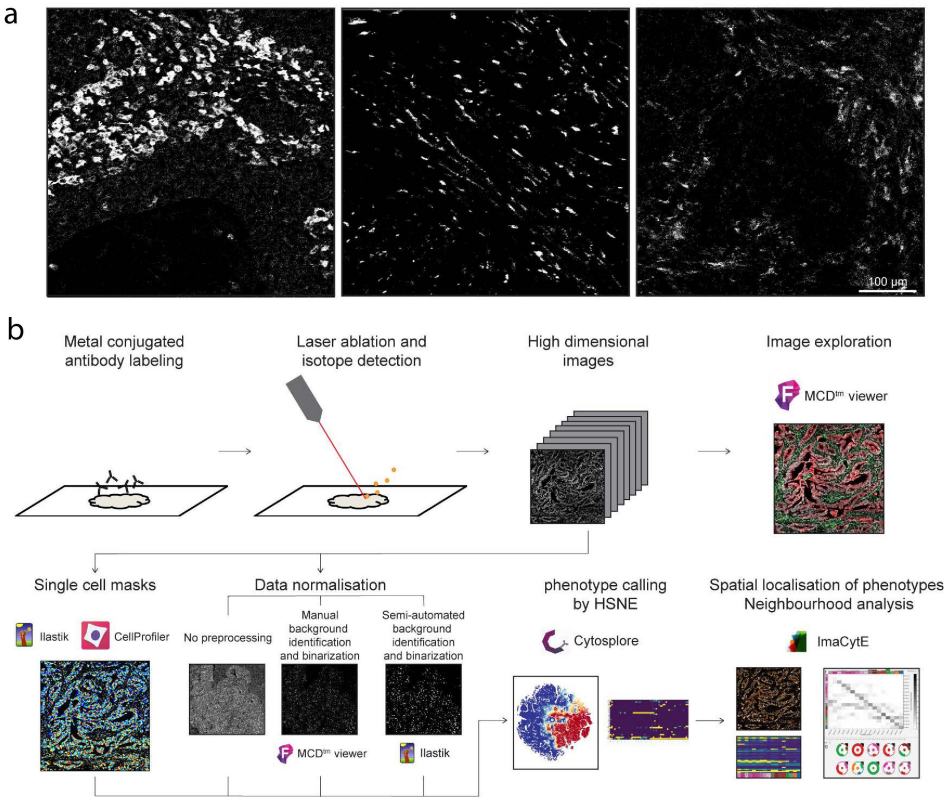

Figure 2.1: *(a) CD163 expression patterns in three samples. A signal range of 3-20 mean dual counts was set for all images in the MCD viewer, but differences in signal intensity and background are observed between images. (b) Workflow for IMC FFPE imaging and data analysis, including the three tested data processing approaches where either no preprocessing, manual background identification and binarization or semi-automated background identification and binarization were performed.*

normalisation, per antibody, on all cell mask overlays where antibody signal was displayed as mean pixel intensity (Figure 2.2a and Figure 2.2b, lower panel). Where differences were observed, the original IMC images were inspected, showing that fluctuations in intensity on the cell masks generally corresponded to variations in signal to noise ratios. This resulted in either overestimation (Figure 2.2a) or underestimation (Figure 2.2b) of cells positive for a marker with variable signal to noise ratios between samples. Furthermore, alongside small differences in marker expression due to, for instance, signal spillover from neighbouring cells, high variability in signal intensity between samples could potentially result in similar immune cell subsets being assigned to distinct immune cell populations when using automated clustering approaches. To test this, a t-SNE embedding was computed using the single cell marker expression data extracted from 48 images

The embedding contained 393727 cells and was visualised in a two-dimensional scatterplot with sample IDs and expression of each marker shown by colour coding. It was observed that cells with a similar marker profile were scattered throughout the t-SNE em-

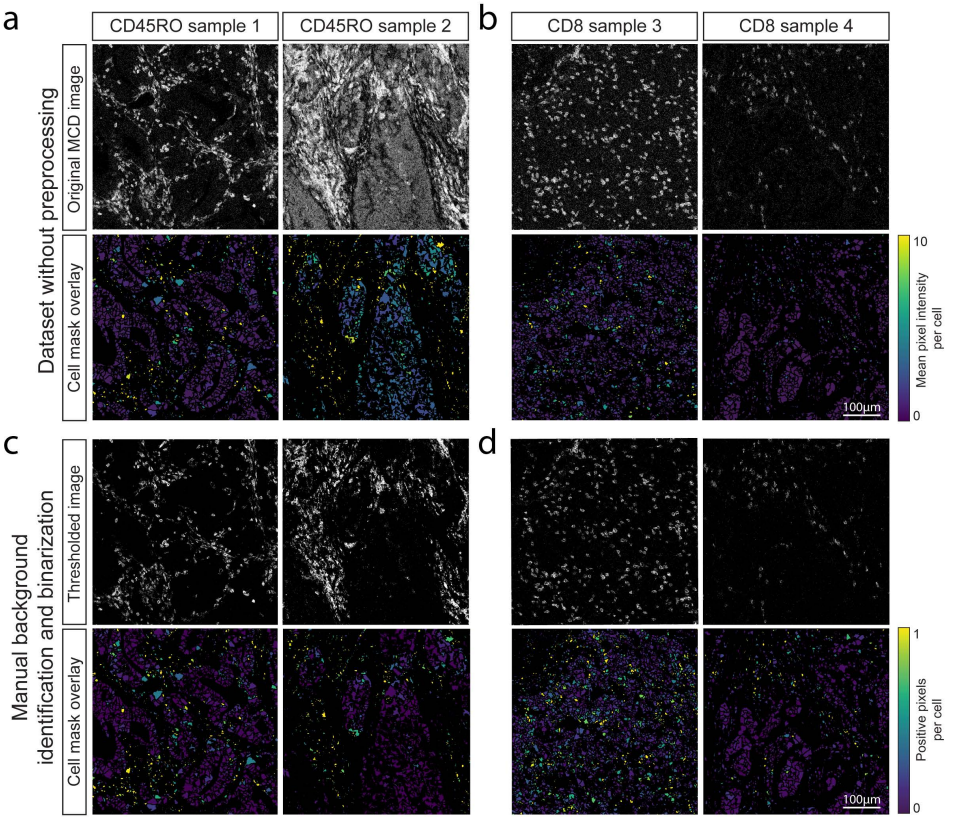

Figure 2.2: *(a) Comparison between original MCD image and cell expression after mask overlay. Variation between images occurs due to differences in background as seen for CD45RO between sample 1 and 2 and (b) variation in signal intensity as observed for CD8 between sample 3 and 4. Signal in the cell mask ranged from 0 – 10 mean duals. (c) comparison of CD45RO and (d) CD8 immunodetection in two thresholded MCD images and the mask overlay after manual thresholding and pixel binarization. Signal intensity ranges between 0 and 1 due to the visualisation as relative frequency of positive pixels per cell.*

bedding rather than clustering together (Figure 2.2a, Figure 2.3a, Figure 2.4). Furthermore, cells also tended to group according to their sample of origin. These observations led to the hypothesis that the intensity range and signal to noise variation between samples can overshadow cell type differences and similarities. Finally, cells positive for FOXP3, CD20 or CD103, markers with a low signal to noise ratio, did not form groups in the t-SNE analysis (Figure 2.4).

2.1.4.2. MANUAL BACKGROUND IDENTIFICATION AND BINARIZATION NORMALISES IMC INTER-SAMPLE VARIATION FOR AUTOMATED DOWNSTREAM ANALYSIS

A methodology was devised to test whether the observed immunodetection variation could be overcome by normalising the IMC data, while minimising data loss, for downstream

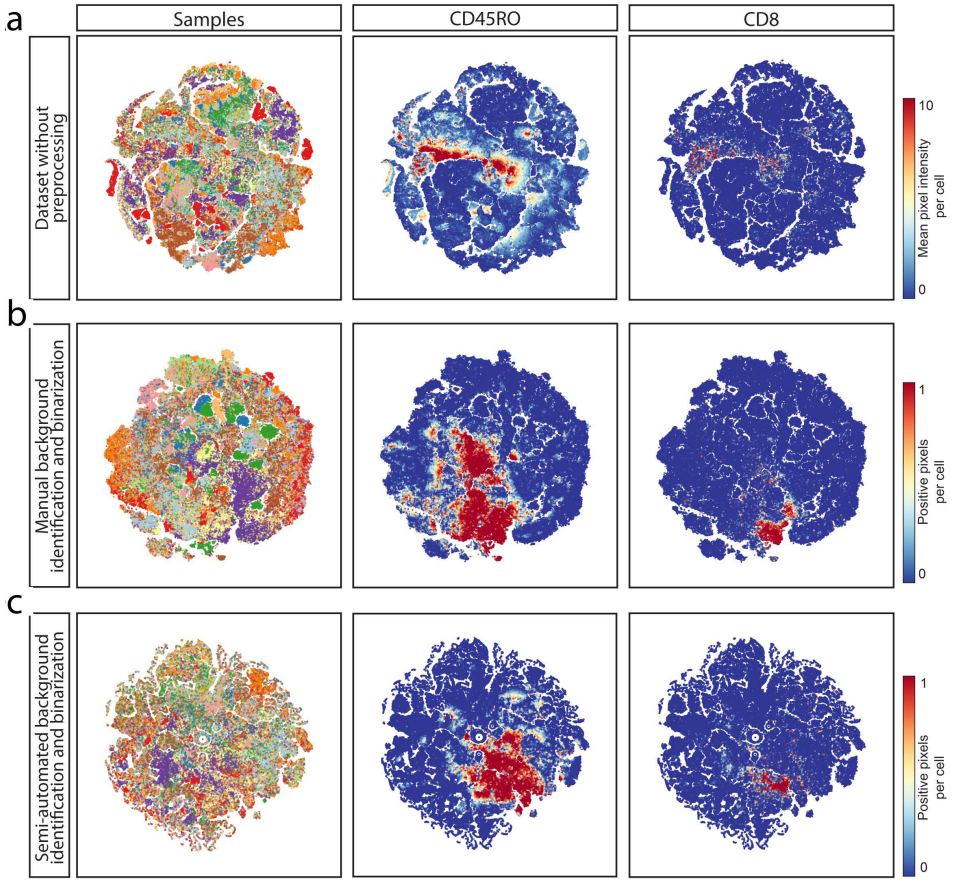

Figure 2.3: *(a) t-SNE analysis embedding of single cell data extracted from all 48 images without preprocessing. Each dot marks a cell and is coloured by sample of origin.To determine the effect of data normalisation, t-SNE analysis embeddings were generated for the same dataset after (b) manual background identification and binarization and after (c) semi-automated identification and binarization.*

analysis. This approach utilised a user-defined minimum signal threshold for each marker followed by pixel binarization of the dataset. To confirm whether this approach was sufficient for reliable downstream analysis, we visualised the percentage of positive pixels on the cell masks. Indeed, setting a minimum signal threshold overcame the variation between samples (Figure 2.2c, d), compared to images obtained without preprocessing of the data (Figure 2.2a, b). However, for markers with a low signal to noise ratio, it was observed that implementing a threshold not only filtered out the background but also signal corresponding to cells expressing the marker of interest (Figure 2.5a-c) . To further assess the effect of thresholding and binarization on downstream analysis, t-SNE embedding, as described in the previous section, was performed on the single cell data obtained after manual background identification and binarization (Figure 2.3b, Figure 2.4). In contrast to the t-SNE embedding of data without preprocessing, cells with a similar marker

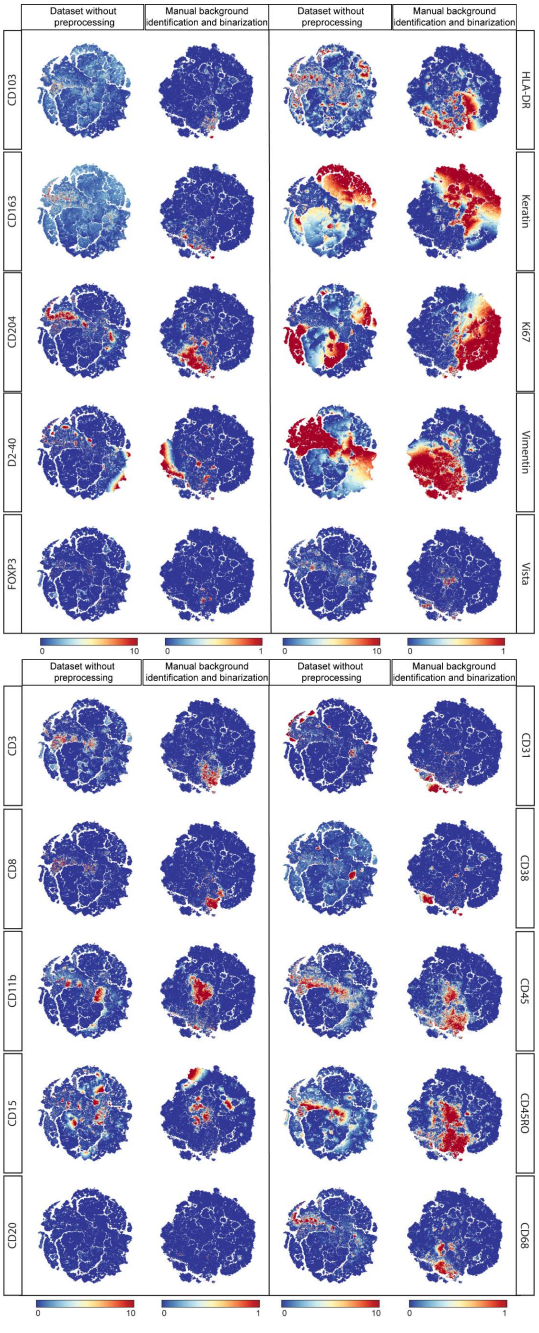

Figure 2.4: *t-SNE analysis on single cells extracted from dataset without preprocessing and the dataset generated with manual background identification and binarization.*

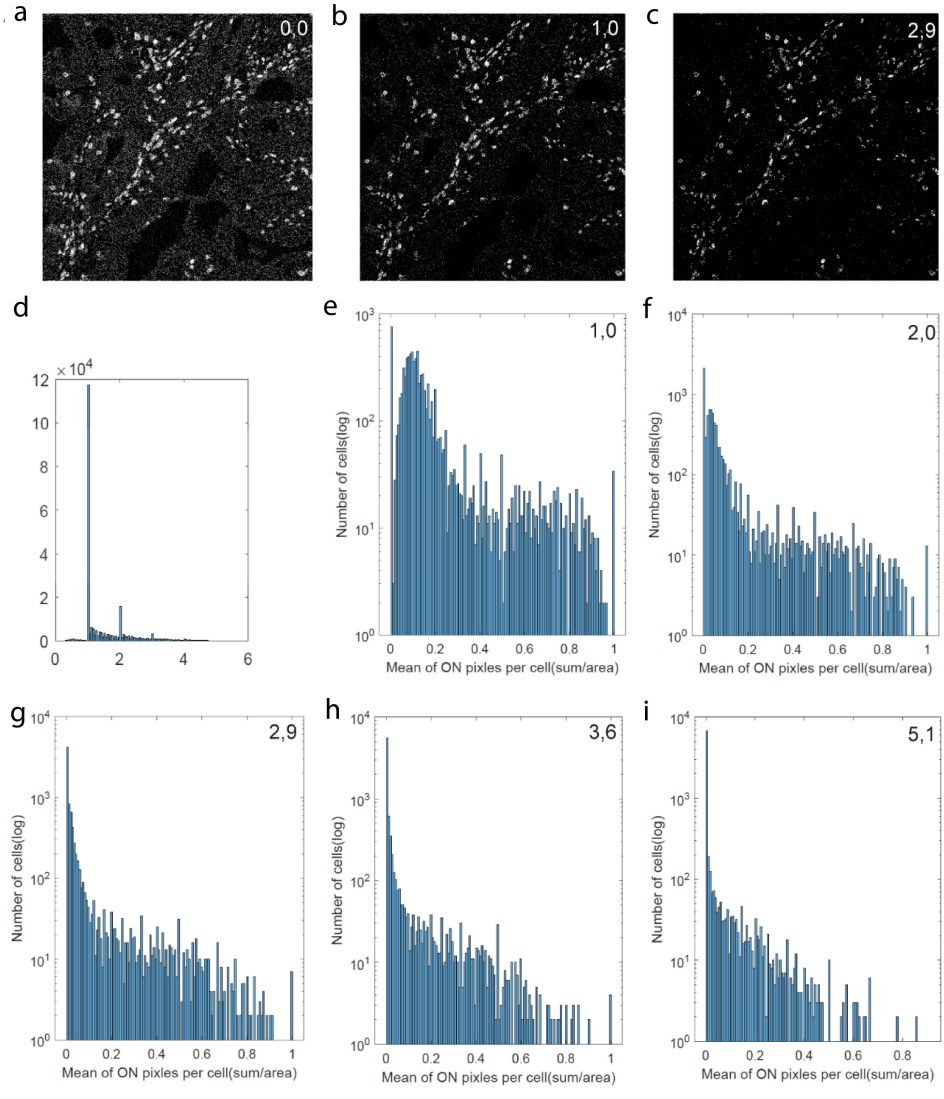

Figure 2.5: *(a-c) CD3 expression in a representative image with a pixel threshold of 0, 1 or 2,9. (d) pixel data plotted in a histogram displaying the relation between pixel intensity (x-axis) and frequency (y-axis). (e) Mean CD3 positive pixels per cell with a cut off threshold at 1. All pixels below threshold were set to 0 and above threshold at 1 and used to determine the mean intensity per cell. (f-i) histograms of CD3 expression with different threshold between 2 and 5*

profile clustered together (as observed for CD8, Figure 2.4). Furthermore, the distinction between positive and negative cells for a specific marker was clearer (as observed for CD163, Figure 2.4). (Figure 2.3b, Figure 2.4). Sample specific clustering was largely resolved but some sample-related bias remained (Figure 2.6b). Further inspection of the t-SNE embedding showed that the cells in those clusters were keratin-positive (a marker

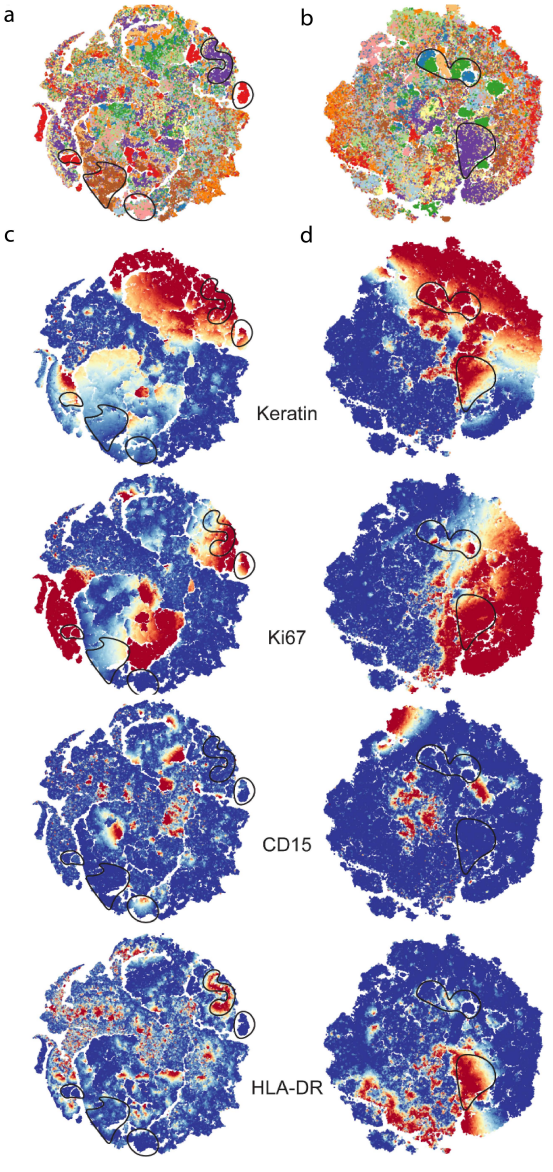

Figure 2.6: *(a) t-SNE embedding of single cell data extracted from dataset without pre-processing. Each colour represents a different sample and cells cluster together by phenotype. Sample-specific clustering is encircled. (b) t-SNE embedding of single cell data extracted from dataset generated with manual background identification and binarization. Each colour represents a different sample and sample specific clustering is encircled. (c and d) Keratin, Ki67, CD15 and HLA-DR expression overlay on the t-SNE from data without pre-processing (c) and data that underwent manual background identification and binarization (d).*

for epithelial cells) with varying combinations of HLA-DR, Ki67 and CD15 expression, markers that are often differentially expressed between cancer cells, which could, in part, explain the sample-specific clustering of the tumour cells (Figure 2.6c, d). In line with the observations made during visual inspection, low numbers of cells positive for dim markers were observed (for instance CD20, FOXP3, Figure 2.4), albeit higher than the number of positive cells observed in t-SNE of the dataset without normalisation (Figure 2.3a, b and Figure 2.4). Thus, manual background identification and binarization of pixel intensity largely resolved sample-specific clustering and allowed for comparison between samples, but did not resolve the presence of false negatives. Although manual background identification was found to overcome some of the challenges of analysing FFPE IMC data, its major disadvantages are that it is time consuming and subject to errors as it requires vast knowledge of the expected immunodetection patterns of each marker and high inter-user variability is inevitable. Furthermore, while background identification through thresholding removes background noise, a portion of specific signal can be lost, particularly when the signal to noise ratio is low, resulting in false negatives. Therefore, we set out to investigate if an automated and unbiased approach could replace manual thresholding. We first visualised the pixel data in histograms for each marker to assess if pixel intensity was bimodally distributed in order to set an automatic threshold between negative and positive pixels. However, no bimodal distribution but a negative correlation between number of pixels and signal intensity was observed, possibly in part due to detector noise of the mass cytometer (Figure 2.5d). We then investigated whether grouping pixel intensities per cell after applying cell masks onto the images allowed the definition of a threshold that separated positive and negative signals for a given marker. Initially, we set the threshold value at 1 mean duals and regarded all cells above this value as positive for a marker. Then, we visualised the data on cell level by plotting the percentage of positive pixels within a cell (Figure 2.4e). Also, at cell level, no clear bimodal distribution was observed. Similarly, cut-off threshold values between 2 and 5 resulted in similar distributions (Figure 2.5f-i). Moreover, a threshold of 2,9 mean duals was comparable to the cut-off chosen during manual thresholding but this could not be deduced from the cell-based value distribution (Figure 2.5c, g). Thus, an automated approach to determine a precise cut-off value could not be established. Furthermore, setting a single-value threshold, as was also observed with manual thresholding, causes a trade-off between the removal of background and low intensity true signal and does not overcome case-specific background signal as was observed for some images and markers (e.g. CD45RO, Figure 2.2a).

2.1.4.3. SEMI-AUTOMATED BACKGROUND IDENTIFICATION LIMITS LOSS OF DATA AND NORMALISES THE IMAGES FOR DOWNSTREAM ANALYSIS OF IMC DATA

To correct for both technical noise and sample-specific background signal, a semi-automated background identification approach based on ilastik's pixel classification algorithm was employed. First, outliers were removed by saturating all pixels below the 1st and above the 99th percentile. This slightly improved the signal to noise compared to raw images, due to the removal of the brightest pixels, but variation in signal intensities between samples remained (Figure 2.7) as well as background noise (Figure 2.8). Thus, percentile normalisation alone was not sufficient to normalise the data. Next, pixels corresponding to either 'background' or 'signal' were labelled for each marker and used to train a random

Figure 2.7: *CD163 expression pattern in the three samples after saturation of pixels below the 1st and above the 99th percentile.*

forest classifier in ilastik. After training on 12 images per marker (or 1 image per sample) the algorithm was applied to all images in the dataset, to create binary signal masks for

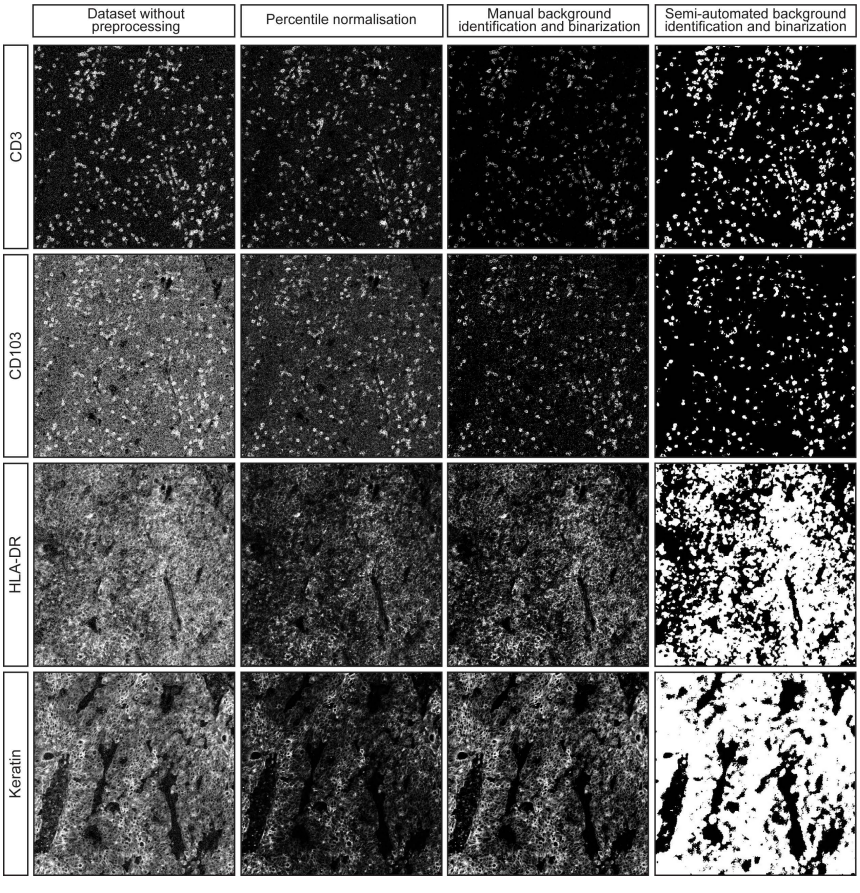

Figure 2.8: *Representative images of CD3, CD103, HLA-DR and Keratin signal without preprocessing, percentile normalisation, manual background removal or semi-automated background removal, both followed by pixel binarization*

each marker. Of note, the algorithm takes into account pixel intensity but also patterns of neighbouring pixels, which could aid the identification of background without losing true signal. Comparison of the images without preprocessing, after percentile normalisation, manual background identification or the semi-automated background identification, showed that the latter approach could be successfully applied to retain true signal while removing a substantial amount of background noise (Figure 2.8). To further confirm the validity of this approach, a t-SNE embedding was computed from the single cell data extracted from the binary signal masks (Figure 2.3c, Figure 2.9). Inspection of sample and marker overlays on the t-SNE embedding, using colour-coding, showed cells grouped by the percentage of positive pixels for each marker rather than per sample, similar to the manual thresholding (Figure 2.3b). Furthermore, the signal to noise ratio was higher and cells that were positive for low intensity markers, such as FOXP3 and CD103, were identified by the semi-automated background removal in contrast to manual thresholding (Figure 2.4, Figure 2.9).

2.1.4.4. BACKGROUND REMOVAL AND BINARIZATION COMBINED WITH THE PROPOSED DOWNSTREAM ANALYSIS PIPELINE ALLOWS PHENOTYPING OF THE TUMOUR IMMUNE MICROENVIRONMENT

To demonstrate the added value of performing background identification and binarization of pixel intensity for the definition of immunophenotypes, clusters of cells with a comparable marker profile were identified, by applying Gaussian mean shift clustering on the t-SNE embedding, computed in the previous section. Next, clusters were mapped back onto the segmentation masks in ImaCytE. A proliferating and non-proliferating tumour cluster was identified through the expression of keratin and distinguished by Ki-67 (Figure 2.10a). Five myeloid clusters were identified by their CD68 expression and differentiated by CD204, CD163 and HLA-DR expression. Furthermore, five lymphoid-cell clusters could be identified where two clusters were CD8 positive, thus, corresponding to cytotoxic T cells. Of note, one of these clusters also showed a positive signal for keratin indicating that these cells were located directly adjacent to epithelial cancer cells. Three clusters were CD8 negative and considered to be mostly composed of CD4 T cells. One of the clusters corresponded to regulatory T cells ($FOXP3^+$) and the other two clusters were differentiated by Ki-67 expression. The tumour, myeloid and lymphoid clusters were each mapped back onto the images and compared to the original MCD images in the MCD viewer (Figure 2.10b). Indeed, the number and location of positive cells for each phenotype was comparable between the images overlaid with phenotype masks and the original MCD images. Thus, semi-automated background identification using ilastik combined with binarization is applicable to normalise IMC datasets derived from archival samples and allows for the identification and localisation of biologically relevant phenotypes.

2.1.5. DISCUSSION

With the rise of mass cytometry for the characterisation of cellular contextures in health and disease, IMC has surfaced as a valuable tool to investigate immunophenotypes while preserving spatial information. IMC allows the simultaneous investigation of over 40 markers thereby generating complex datasets that require analysis tools that combine deep immunophenotyping data with spatial localisation and neighbourhood analysis. However,

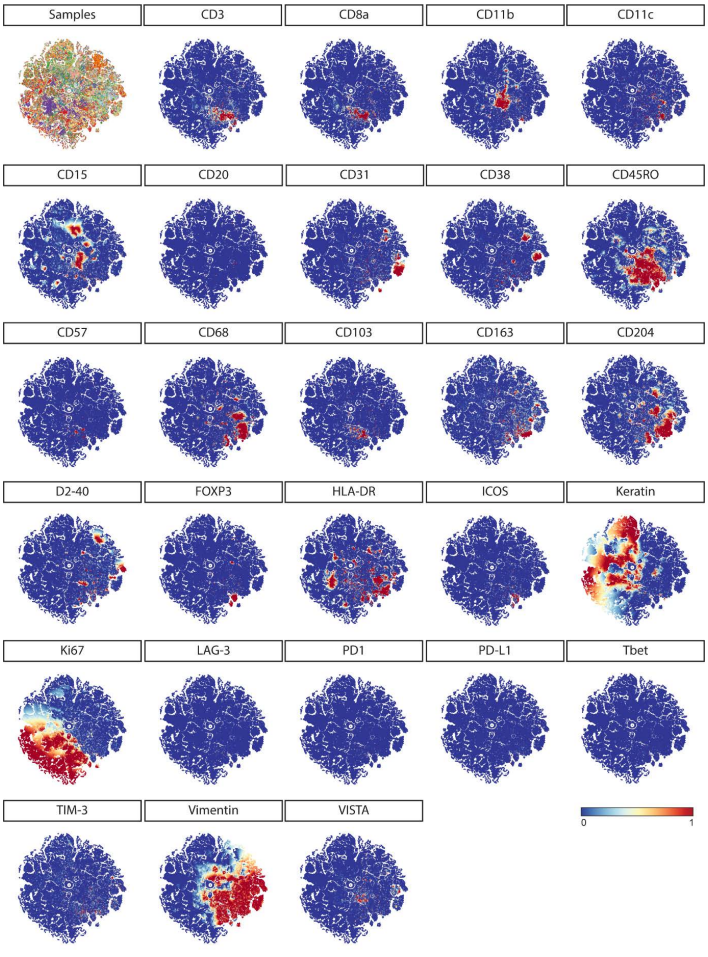

Figure 2.9: *t-SNE analysis on single cells extracted from the dataset generated through semi-automated background identification and binarization. Data is shown in a range of 0 (blue) to 1 (red).*

before data interpretation, non-biological variation of signal intensities between tissues should be dealt with. Technical noise is consistent in each image and will therefore not influence downstream analysis as long as the signal to noise ratio is high, and can be addressed by optimising the wet-lab procedures and performing the immunodetection of tissues in a single experiment. However, sample-specific background, related to tissue processing procedures, is impossible to address during the immunodetection procedures. FFPE tissue is often the tissue of choice for IMC due to its accessibility and good morphology. Differences in ischemia time, tissue fixation procedures and age of the samples affects immunodetection and tissue-related background, prompting the need to normalise the data before analysis. Furthermore, and in contrast to single cell technologies, spatial

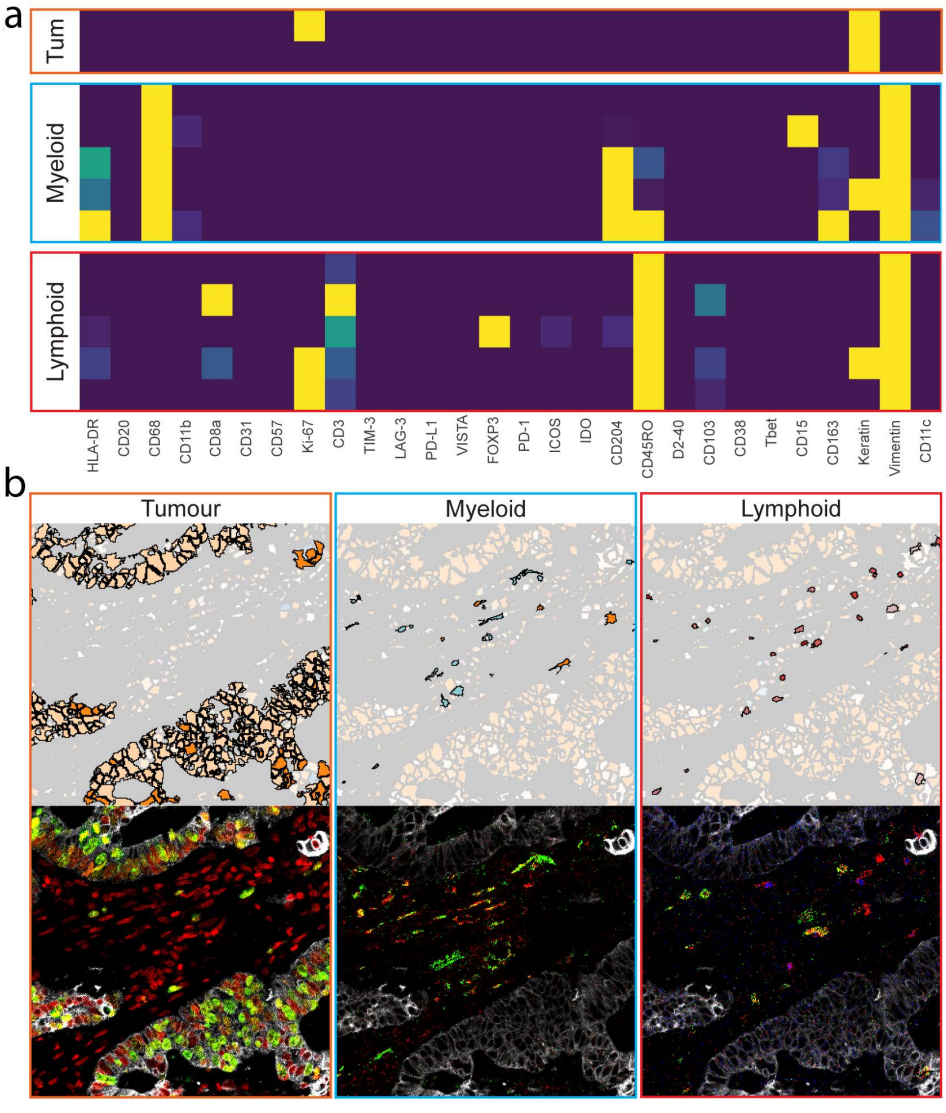

Figure 2.10: *(a) Lymphoid, myeloid and tumour (tum) clusters identified in the t-SNE analysis from Figure 2.3c . (b) the clusters were mapped back onto a representative image and compared to their corresponding cell types in the MCD files. The images contain the following markers: for tumour keratin (white), Ki67 (green) and DNA (red), for myeloid cells CD68 (red) and CD163 (green) and for lymphoid cells CD3 (red), CD8 (green) and FOXP3 (blue). To improve visibility a lower threshold of 1 mean dual counts was set.*

technologies, making use of tissue sections, require that cells are cut at different planes which influences the prevalence and intensity of a marker of interest. In light of this, it is preferable to classify cells into positive and negative categories for a given marker

rather than relying excessively on intensity differences across cells for the identification of phenotypes.

Due to tissue-specific variations in signal intensities, direct phenotyping of IMC data, without normalisation, can lead to cells clustering per sample rather than phenotypes as the variation in overall signal intensity overshadows the differences between cell types. Furthermore, the signal range of the more abundant structural markers (e.g. keratin, vimentin) is much wider compared to scarcer, but important targets of investigation (e.g. co-receptors on T cells). This difference influences downstream computational methods such as clustering or dimensionality reduction algorithms and hinders the detection of specific cell subsets.

In this work, we enable the analysis of large IMC datasets and improve the detection of cells expressing lowly abundant proteins by removing background noise and binarizing each sample's pixel values. Binarizing was done by assigning the value of 1 to all pixels determined to be positive for a given marker and the value of 0 to all other pixels. More specifically, we tested two different normalisation approaches. First, a manual method was utilised where a minimum threshold corresponding to signal was set for each marker in all images. All pixels below threshold were set to 0 and all pixels above were set to 1. Then, marker expression per cell was defined as the percentage of positive pixels per cell. This approach indeed partially overcame variation between tissues and allowed for t-SNE-guided phenotype identification. However, manual thresholding is labour intensive and relies on vast knowledge of the expected immunodetection patterns for each antibody, leading to potentially biased results. In general, thresholding is a trade-off between the removal of background and loss of true signal and therefore can result in the frequent designation of false negatives. An automated approach could overcome these challenges but large variations between samples and markers make fully automated unsupervised methods infeasible and the lack of labelled datasets is prohibitive for supervised machine learning approaches. Therefore, we propose a, semi-automated, methodology using ilastik where we first annotate representative pixels either as actual signal or background noise and then a random forest classifier is run to categorise the whole dataset, based on these categories. This approach is faster, less subjective and results in data comparable to manual thresholding. Furthermore, loss of true signal is less frequently observed, and dim markers are more clearly represented after semi-automated background identification and pixel intensity binarization. A drawback of signal binarization is the potential loss of information on biologically relevant, signal intensity variations for a given marker. However, it is currently challenging to distinguish between biological and technical causes for signal intensity variations in in situ imaging approaches that utilise FFPE tissue. Thus, to limit the loss of biologically relevant information, we chose to binarize the data at pixel level which still allows the evaluation of differences in number of positive pixels per cell. Finally, accurate cell segmentation remains an important challenge to address in molecular imaging. Current methodologies, like the ones employed here, are not flawless, particularly in areas where cells are densely packed. Therefore, we have performed cell segmentation on raw images and upstream of normalisation approaches in order to allow their comparison independently of segmentation.

In recent years, great advancements have been made in the analysis and interpretation of single cell (mass cytometry) data and a number of developed tools have also proven useful

for the analysis of IMC datasets. However, it is essential that such knowledge is combined with the accumulated experience in the immunohistochemistry and imaging fields to best address immunodetection variation amongst samples and to deal with technical artefacts. The here described normalisation methodology enables the comparative analysis of datasets generated from different tissues and it supports the identification of less abundant cellular subsets. Furthermore, the methodology does not require adaptation in immunodetection procedure and can, thus, be directly applied on available datasets. In sum, this work has the potential to directly aid research groups in their analysis and interpretation of Imaging Mass Cytometry data.

Finally, we would like to declare no conflict of interest and data is available upon request.

2.1.6. ACKNOWLEDGEMENTS

Noel de Miranda is funded by the European Research Council (ERC) under the European Union's Horizon 2020 Research and Innovation Programme (grant agreement no. 852832). Antonios Somarakis received funding through Leiden University Data Science Research Programme.

2.1.7. SUPPLEMENTARY INFORMATION

Target	Clone	Metal	Incubation Time	Temperature	Dilution
CD45	D9M8I	^{115}In	Overnight	4°C	50
HLA-DR	TAL-1B5	^{141}Pr	5 hours	RT	100*
CD20	H1	^{142}Nd	Overnight	4°C	100
CD68	D4B9C	^{143}Nd	Overnight	4°C	100*
CD11b	D6X1N	^{144}Nd	5 hours	RT	100
CD4	EPR6855	^{145}Nd	5 hours	RT	50
CD8α	D8A8Y	^{146}Nd	5 hours	RT	50
CD31	89C2	^{147}Sm	Overnight	4°C	100*
CD57	HNK-1/Leu-7	^{151}Eu	Overnight	4°C	100*
Ki-67	8D5	^{152}Sm	Overnight	4°C	100*
CD3	D7A6E	^{153}Eu	Overnight	4°C	50
TIM-3	D5D5R	^{154}Sm	5 hours	RT	100
LAG3	D2G4O	^{155}Gd	5 hours	RT	50
PD-L1	E1L3N	^{156}Gd	Overnight	4°C	50
VISTA	D1L2G	^{158}Gd	5 hours	RT	100
FoxP3	D6O8R	^{159}Tb	Overnight	4°C	50
PD-1	D4W2J	^{160}Gd	5 hours	RT	50
ICOS	D1K2T	^{161}Dy	5 hours	RT	50
IDO	D5J4E	^{162}Dy	Overnight	4°C	100
CD204	J5HTR3	^{164}Dy	5 hours	RT	50
CD45RO	UCHL1	^{165}Ho	Overnight	4°C	100*
D2-40	D2-40	^{166}Er	Overnight	4°C	100*
CD103	EPR4166(2)	^{168}Er	5 hours	RT	50
CD38	EPR4106	^{169}Tm	Overnight	4°C	100*
T-bet	4B10	^{170}Er	5 hours	RT	50
CD15	BRA-4F1	^{171}Yb	Overnight	4°C	100*
CD163	EPR14643-36	^{173}Yb	5 hours	RT	50
CD11c	EP1347Y	^{176}Yb	5 hours	RT	100
Vimentin	D21H3	^{194}Pt	Overnight	4°C	50
Pan-Keratin	AE1/AE3 and C11	^{198}Pt	Overnight	4°C	50

Table 2.1: *Imaging Mass Cytometry antibody panel. Antibodies with a * had prediluted stock solutions.*

2.2. Segmentation of microglia cells

This chapter was adapted from:

B. Kenkhuis*, **A. Somarakis***, L. de Haan, O. Dzyubachyk, M. E. IJsselsteijn, N. F. de Miranda, B. P. Lelieveldt, J. Dijkstra, W. M. van Roon-Mom, T. Höllt, and L. van der Weerd, "Iron loading is a prominent feature of activated microglia in Alzheimer's disease patients ," *Acta Neuropathologica Communications*, vol. 9, no. 1, pp. 1–15, 2021.

* Contributed equally

2.2.1. INTRODUCTION

Identification of the different microglia types was based on the amount of expressed proteins over the segmented area (Figure 2.11a). Hence, accurate segmentation of the whole microglia cell area is of paramount importance for our method. Solutions currently available for microglia cell segmentation (Abdolhoseini et al. [29], Inform [30]) typically fall short of capturing the whole microglia area (Figure 2.11b). These are focused on either capturing the skeleton of the cells, without properly identifying the cell boundaries (Figure 2.11c), or segmenting the microglia's soma excluding their processes, which in the acquired 2D images are typically detached from the soma ((Figure 2.11d). As a result, a novel segmentation algorithm for this type of data was developed.

2.2.2. METHODOLOGY

Identification of the entire cytoplasmic area of microglia cells is error-prone, especially in regions close to Aβ-plaques, where microglia cells are densely packed. This problem was tackled by starting with the identification of the microglia's soma. This part the of the microglia cells should overlap with its nucleus and shows high intensity values, making it easily discernible. Segmentation of microglia nuclei and somas was performed using a customized level-set-based cell segmentation method [31]. The main algorithm parameters are the weight for the energy terms minimizing the perimeter (ν) and the area (μ), which are empirically selected for each segmentation task. Larger parameter values correspond to smoother segmentation results. For the microglia nucleus segmentation, the DNA component image was used as input, and level-set parameters were set to ν=2 and μ=3. Similarly, for the microglia soma segmentation the summation of the intensity values of the membrane (TMEM, PRY12, FTL, Iba) component images was utilized as input and level-set parameters were set to ν=2 and μ=3. In both cases, level-sets were initialized with regions obtained using the Otsu thresholding method [32] which is robust to intensity variation between images originating from the white and grey matter. Additionally, somas and nuclei with a total area smaller than 50 and 30 pixels, respectively, were removed.

For the extension of the obtained segmentation to the whole cytoplasmic area, the approach previously described for soma was repeated with less strong regularization (v=2, μ=2). The result of this step was a finer segmentation capturing microglia areas that are less bright than the soma. Connected components overlapping with the previously identified somas were regarded as microglia cells, whereas not overlapping components were considered as possible detached processes. At this step, in case a blood vessel was identified in an image, the Li thresholding method [33] was chosen over Otsu for the initialization of the level-sets algorithm, as it is less sensitive to the high intensity pixels representing the vessel. Vessels were defined as components larger than 4000 pixels, after Otsu thresholding of the autofluorescent component image.

For correct identification of microglia cells in the proximity of Aβ-plaques, the watershed segmentation was applied specifically to those cells whose cytoplasmic area is shared among multiple microglia somas [34]. Aβ-plaque identification was performed employing a semi-supervised approach using Ilastik [35].

Finally, branches identified within a 10 pixel radius from the region corresponding to each identified microglia soma were identified as detached processes and assigned to the microglia cell.

2.2.3. RESULTS

A sample resulting segmentation of entire microglial cells is illustrated in Figure 2.11e. For the evaluation of our algorithm, 186 cells were manually segmented, in 7 images from different subjects and regions. Our proposed segmentation framework outperformed the available segmentation solutions correctly capturing 153 cells (Inform: 12 cells, Abdolhoseini et al., 2019: 49 cells compare Figure 2.11g), with false positive 33 cells (Figure 2.11h) and false negatives 40 cells (Figure 2.11i). Among the correctly identified cells, median Dice's similarity index [36] of 0.8 was achieved (Figure 2.11f).

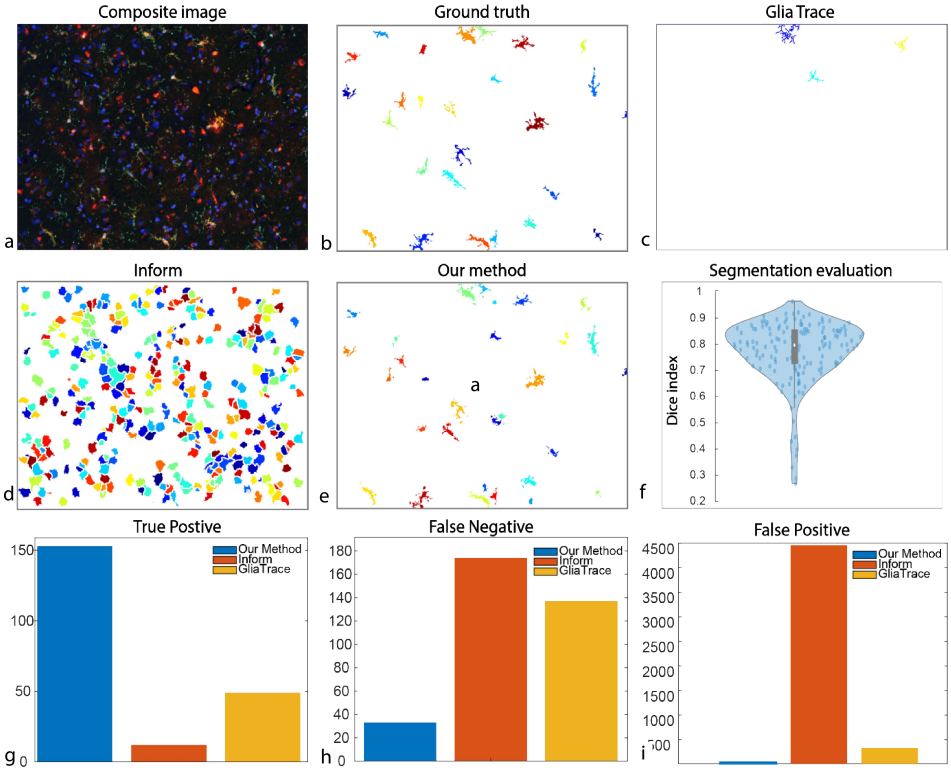

Figure 2.11: *(a) Example of multispectral immunofluorescence data extracted from the grey matter of an AD-patient. The segmentation mask derived: manually by the specialist (b), using the GliaTrace toolbox (c), automatically from the "Inform" image acquisition software (d) and from our proposed segmentation pipeline (e). (f) Dice's coefficient for each of the 156 cells segmented by our method. Number of correctly identified (g), missing (h) and falsely identified cells (i).*

REFERENCES

[1] N. L. de Vries, V. van Unen, M. E. Ijsselsteijn, T. Abdelaal, R. van der Breggen, A. Farina Sarasqueta, A. Mahfouz, K. Peeters, T. Hollt, B. P. F. Lelieveldt, F. Koning, and N. de Miranda, "High-dimensional cytometric analysis of colorectal cancer reveals novel mediators of antitumour immunity," *Gut*, vol. 69, no. 4, pp. 691–703, 2020.

[2] S. J. S. Rubin, L. Bai, Y. Haileselassie, G. Garay, C. Yun, L. Becker, S. E. Streett, S. R. Sinha, and A. Habtezion, "Mass cytometry reveals systemic and local immune signatures that distinguish inflammatory bowel diseases," *Nat Commun*, vol. 10, no. 1, p. 2686, 2019.

[3] N. Li, V. van Unen, T. Abdelaal, N. Guo, S. A. Kasatskaya, K. Ladell, J. E. McLaren, E. S. Egorov, M. Izraelson, S. M. Chuva de Sousa Lopes, T. Hollt, O. V. Britanova, J. Eggermont, N. de Miranda, D. M. Chudakov, D. A. Price, B. P. F. Lelieveldt, and F. Koning, "Memory cd4(+) t cells are generated in the human fetal intestine," *Nat Immunol*, vol. 20, no. 3, pp. 301–312, 2019.

[4] S. C. Bendall, E. F. Simonds, P. Qiu, A. D. Amir el, P. O. Krutzik, R. Finck, R. V. Bruggner, R. Melamed, A. Trejo, O. I. Ornatsky, R. S. Balderas, S. K. Plevritis, K. Sachs, D. Pe'er, S. D. Tanner, and G. P. Nolan, "Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum," *Science*, vol. 332, no. 6030, pp. 687–96, 2011.

[5] E. W. Newell, N. Sigal, S. C. Bendall, G. P. Nolan, and M. M. Davis, "Cytometry by time-of-flight shows combinatorial cytokine expression and virus-specific cell niches within a continuum of cd8+ t cell phenotypes," *Immunity*, vol. 36, no. 1, pp. 142–52, 2012.

[6] S. C. Wei, J. H. Levine, A. P. Cogdill, Y. Zhao, N. A. S. Anang, M. C. Andrews, P. Sharma, J. Wang, J. A. Wargo, D. Pe'er, and J. P. Allison, "Distinct cellular mechanisms underlie anti-ctla-4 and anti-pd-1 checkpoint blockade," *Cell*, vol. 170, no. 6, pp. 1120–1133 e17, 2017.

[7] D. R. Bandura, V. I. Baranov, O. I. Ornatsky, A. Antonov, R. Kinach, X. Lou, S. Pavlov, S. Vorobiev, J. E. Dick, and S. D. Tanner, "Mass cytometry: technique for real time single cell multitarget immunoassay based on inductively coupled plasma time-of-flight mass spectrometry," *Anal Chem*, vol. 81, no. 16, pp. 6813–22, 2009.

[8] O. I. Ornatsky, R. Kinach, D. R. Bandura, X. Lou, S. D. Tanner, V. I. Baranov, M. Nitz, and M. A. Winnik, "Development of analytical methods for multiplex bio-assay with inductively coupled plasma mass spectrometry," *J Anal At Spectrom*, vol. 23, no. 4, pp. 463–469, 2008.

[9] C. Giesen, H. A. Wang, D. Schapiro, N. Zivanovic, A. Jacobs, B. Hattendorf, P. J. Schuffler, D. Grolimund, J. M. Buhmann, S. Brandt, Z. Varga, P. J. Wild, D. Gunther, and B. Bodenmiller, "Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry," *Nat Methods*, vol. 11, no. 4, pp. 417–22, 2014.

[10] S. Chevrier, H. L. Crowell, V. R. T. Zanotelli, S. Engler, M. D. Robinson, and B. Bodenmiller, "Compensation of signal spillover in suspension and imaging mass cytometry," *Cell Syst*, vol. 6, no. 5, pp. 612–620 e5, 2018.

[11] M. E. Ijsselsteijn, R. van der Breggen, A. Farina Sarasqueta, F. Koning, and N. de Miranda, "A 40-marker panel for high dimensional characterization of cancer immune microenvironments by imaging mass cytometry," *Front Immunol*, vol. 10, p. 2534, 2019.

[12] N. Guo, V. van Unen, M. E. Ijsselsteijn, L. F. Ouboter, A. E. van der Meulen, S. M. Chuva de Sousa Lopes, N. de Miranda, F. Koning, and N. Li, "A 34-marker panel for imaging mass cytometric analysis of human snap-frozen tissue," *Front Immunol*, vol. 11, p. 1466, 2020.

[13] D. Schulz, V. R. T. Zanotelli, J. R. Fischer, D. Schapiro, S. Engler, X. K. Lun, H. W. Jackson, and B. Bodenmiller, "Simultaneous multiplexed imaging of mrna and proteins with subcellular resolution in breast cancer tissue samples by mass cytometry," *Cell Syst*, vol. 6, no. 4, p. 531, 2018.

[14] H. W. Jackson, J. R. Fischer, V. R. T. Zanotelli, H. R. Ali, R. Mechera, S. D. Soysal, H. Moch, S. Muenst, Z. Varga, W. P. Weber, and B. Bodenmiller, "The single-cell pathology landscape of breast cancer," *Nature*, vol. 578, no. 7796, pp. 615–620, 2020.

[15] Y. J. Wang, D. Traum, J. Schug, L. Gao, C. Liu, H. Consortium, M. A. Atkinson, A. C. Powers, M. D. Feldman, A. Naji, K. M. Chang, and K. H. Kaestner, "Multiplexed in situ imaging mass cytometry analysis of the human endocrine pancreas and immune system in type 1 diabetes," *Cell Metab*, vol. 29, no. 3, pp. 769–783 e4, 2019.

[16] N. Damond, S. Engler, V. R. T. Zanotelli, D. Schapiro, C. H. Wasserfall, I. Kusmartseva, H. S. Nick, F. Thorel, P. L. Herrera, M. A. Atkinson, and B. Bodenmiller, "A map of human type 1 diabetes progression by imaging mass cytometry," *Cell Metab*, vol. 29, no. 3, pp. 755–768 e5, 2019.

[17] S. Berg, D. Kutra, T. Kroeger, C. N. Straehle, B. X. Kausler, C. Haubold, M. Schiegg, J. Ales, T. Beier, M. Rudy, K. Eren, J. I. Cervantes, B. Xu, F. Beuttenmueller, A. Wolny, C. Zhang, U. Koethe, F. A. Hamprecht, and A. Kreshuk, "ilastik: interactive machine learning for (bio)image analysis," *Nat Methods*, vol. 16, no. 12, pp. 1226–1232, 2019.

[18] A. E. Carpenter, T. R. Jones, M. R. Lamprecht, C. Clarke, I. H. Kang, O. Friman, D. A. Guertin, J. H. Chang, R. A. Lindquist, J. Moffat, P. Golland, and D. M. Sabatini, "Cellprofiler: image analysis software for identifying and quantifying cell phenotypes," *Genome Biol*, vol. 7, no. 10, p. R100, 2006.

[19] D. Schapiro, H. W. Jackson, S. Raghuraman, J. R. Fischer, V. R. T. Zanotelli, D. Schulz, C. Giesen, R. Catena, Z. Varga, and B. Bodenmiller, "histoCAT: analysis of cell phenotypes and interactions in multiplex image cytometry data," *Nat Methods*, vol. 14, no. 9, pp. 873–876, 2017.

[20] A. Somarakis, V. van Unen, F. Koning, B. P. Lelieveldt, and T. Höllt, "ImaCytE: visual exploration of cellular microenvironments for imaging mass cytometry data," *IEEE Transactions on Visualization and Computer Graphics*, vol. 27, no. 1, pp. 98–110, 2021.

[21] M. Economou, L. Schoni, C. Hammer, J. A. Galvan, D. E. Mueller, and I. Zlobec, "Proper paraffin slide storage is crucial for translational research projects involving immunohistochemistry stains," *Clin Transl Med*, vol. 3, no. 1, p. 4, 2014.

[22] M. Werner, A. Chott, A. Fabiano, and H. Battifora, "Effect of formalin tissue fixation and processing on immunohistochemistry," *Am J Surg Pathol*, vol. 24, no. 7, pp. 1016–9, 2000.

[23] G. O'Hurley, E. Sjostedt, A. Rahman, B. Li, C. Kampf, F. Ponten, W. M. Gallagher, and C. Lindskog, "Garbage in, garbage out: a critical evaluation of strategies used for validation of immunohistochemical biomarkers," *Mol Oncol*, vol. 8, no. 4, pp. 783–98, 2014.

[24] G. Bussolati and E. Leonardo, "Technical pitfalls potentially affecting diagnoses in immunohistochemistry," *J Clin Pathol*, vol. 61, no. 11, pp. 1184–92, 2008.

[25] C. R. Taylor and R. M. Levenson, "Quantification of immunohistochemistry–issues concerning methods, utility and semiquantitative assessment ii," *Histopathology*, vol. 49, no. 4, pp. 411–24, 2006.

[26] A. S. Leong, "Quantitation in immunohistology: fact or fiction? a discussion of variables that influence results," *Appl Immunohistochem Mol Morphol*, vol. 12, no. 1, pp. 1–7, 2004.

[27] V. van Unen, T. Hollt, N. Pezzotti, N. Li, M. J. T. Reinders, E. Eisemann, F. Koning, A. Vilanova, and B. P. F. Lelieveldt, "Visual analysis of mass cytometry data by hierarchical stochastic neighbour embedding reveals rare cell types," *Nat Commun*, vol. 8, no. 1, p. 1740, 2017.

[28] T. Höllt, N. Pezzotti, V. van Unen, F. Koning, E. Eisemann, B. P. F. Lelieveldt, and A. Vilanova, "Cytosplore: Interactive immune cell phenotyping for large single-cell datasets," *Computer Graphics Forum (Proceedings of EuroVis)*, vol. 35, no. 3, pp. 171–180, 2016.

[29] M. Abdolhoseini, M. G. Kluge, F. R. Walker, and S. J. Johnson, "Segmentation, tracing, and quantification of microglial cells from 3d image stacks," *Scientific reports*, vol. 9, no. 1, pp. 1–10, 2019.

[30] PerkinElmer, "inForm Advanced Image Analysis Software - For Accurately Quantifying Biomarkers in Tissue Sections." https://tinyurl.com/y2ckm848, 2012. [Online; Accessed: 2021-07-14].

[31] O. Dzyubachyk, W. A. Van Cappellen, J. Essers, W. J. Niessen, and E. Meijering, "Advanced level-set-based cell tracking in time-lapse fluorescence microscopy," *IEEE Transactions on Medical Imaging*, vol. 29, no. 3, pp. 852–867, 2010.

[32] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.

[33] C. Li and C. Lee, "Minimum cross entropy thresholding," *Pattern Recognition*, vol. 26, no. 4, pp. 617–625, 1993.

[34] S. Beucher, "Use of watersheds in contour detection," in *Proceedings of the International Workshop on Image Processing*, CCETT, 1979.

[35] C. Sommer, C. Straehle, U. Koethe, and F. A. Hamprecht, "Ilastik: Interactive learning and segmentation toolkit," in *Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on*, pp. 230–233, IEEE, 2011.

[36] L. R. Dice, "Measures of the amount of ecologic association between species," *Ecology*, vol. 26, no. 3, pp. 297–302, 1945.

3

Micro-Environment Exploration

This chapter was adapted from:

A. Somarakis, V. V. Unen, F. Koning, B. P. Lelieveldt, and T. Höllt, "ImaCytE: Visual Exploration of Cellular Micro-Environments for Imaging Mass Cytometry Data," *IEEE Transactions on Visualization and Computer Graphics*, vol. 27, no. 1, pp. 98–110, 2021.

3.1. ABSTRACT

Tissue functionality is determined by the characteristics of tissue-resident cells and their interactions within their microenvironment. Imaging Mass Cytometry offers the opportunity to distinguish cell types with high precision and link them to their spatial location in intact tissues at sub-cellular resolution. This technology produces large amounts of spatially-resolved high-dimensional data, which constitutes a serious challenge for the data analysis. We present an interactive visual analysis workflow for the end-to-end analysis of Imaging Mass Cytometry data that was developed in close collaboration with domain expert partners. We implemented the presented workflow in an interactive visual analysis tool; ImaCytE. Our workflow is designed to allow the user to discriminate cell types according to their protein expression profiles and analyze their cellular microenvironments, aiding in the formulation or verification of hypotheses on tissue architecture and function. Finally, we show the effectiveness of our workflow and ImaCytE through a case study performed by a collaborating specialist.

3.2. INTRODUCTION

Cells are the structural units of life and the main orchestrators of tissue function [1]. In recent years it has become clear that the phenotype and function of cells is co-determined by their location and interactions within the tissue context. Detailed analysis of the heterogeneity of cells provides information on tissue composition, while the spatial organization of cells provides information on tissue function. Both aspects are important for unraveling the complexity of tissue function and disease.

Conventional analysis of cell heterogeneity requires processing of tissue specimens into single-cell suspensions at the expense of spatial information. Traditional imaging analysis based on immunofluorescence is limited by the number of proteins that can be analyzed simultaneously and consequently the spatial location of only a few cell types can be revealed. A novel imaging modality, Imaging Mass Cytometry [2] was recently introduced allowing the measurement of the expression of up to 40 proteins simultaneously with a spatial resolution as low as one micrometer per pixel, preserving the tissue architecture at sub-cellular resolution [3]. The broad range of measured proteins allows the identification of a large variety of cell phenotypes within the tissue context. However, the wide range of discovered cell phenotypes in combination with the spatial resolution creates a highly complex system to analyze.

For analysis, each measured protein is typically interpreted as a dimension, defining every cell as a data-point in a high-dimensional space. The analysis of this high-dimensional space allows for the identification of distinct cell phenotypes, not known *a priori*, similar to non-spatial methods. In addition, the spatial resolution of the data allows biologists to localize the identified cells in the tissue architecture and form hypotheses based on the location of cells and their microenvironment. Here, we define the microenvironment of a cell as the directly adjacent cells in the tissue.

To explore the cells in their corresponding microenvironments, we first need to identify the different cell phenotypes existing in the analyzed tissue. Therefore, we extended our previous work [4] focused on the phenotype identification of non-spatial mass cytometry data to enable the exploration of cellular microenvironments and discover subsets of cell

phenotypes with unique microenvironment characteristics. We propose an interactive, data-driven workflow and consequently a tool that has been designed for the end-to-end exploratory analysis of Imaging Mass Cytometry data. The main contributions of this paper are:

1 An end-to-end workflow for the analysis of spatially resolved *–omics* data, including

- interactive definition of cohesive phenotypical groups,
- stratification of phenotypical groups based on their microenvironment characteristics,
- inspection of significant spatial interactions in the tissue, and
- motifs to group and visualize cells with a similar cellular microenvironment.

2 The implementation of the proposed workflow in an interactive visual analysis tool, we call ImaCytE.

The remainder of this paper is structured as follows. We present related work in Section 3.3, followed by a brief description of the biological background in Section 3.4. In Section 3.5 we identify and abstract tasks and present their corresponding designs. The effectiveness of our tool and workflow is shown in a case study in Section 3.6. We conclude in Section 3.7 and present open problems and potential directions for future work.

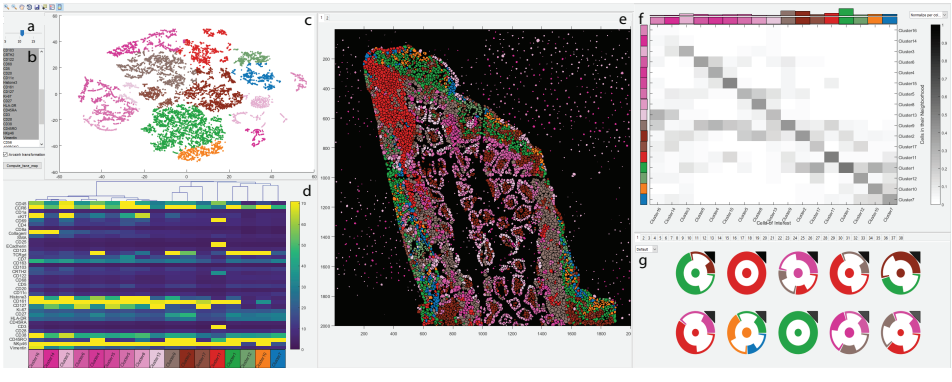

Figure 1: ***ImaCytE****. Screenshot of our integrated system including the settings panels (a,b), the embedding view (c), the heatmap view (d) for cluster visualization, the image view (e), the interactions heatmap (f), and the motif view (g).*

3.3. RELATED WORK

The routine acquisition of high-resolution spatially-resolved, *–omics* data is a relatively new development. Consequently, only few integrated analysis tools, leveraging the full complexity of such data exist. Multeesum [5] allows the comparative visualization of

spatio-temporal gene expression in fruit fly embryos acquired through photon microscopy. Abdemoula et al. [6], used a t-SNE-based workflow for the exploration of Imaging Mass Spectrometry data [7]. InsituNet [8] is using an interactive network-based visualization to illustrate the spatial co-expression of transcripts over a specified region of the tissue. Even though InSituNet allows the location of transcript interactions, one cannot explore spatial co-expression of transcript subsets. Compared to Imaging Mass Cytometry, all of these techniques deal with relatively low resolution data and none of them provides means to interpret microenvironment patterns.

Phenotype identification in non-spatially resolved Mass Cytometry [9] data has been a very active field of research, resulting in a large amount of tools. The most widely used ones include Vortex [10], FlowMaps [11], FlowSOM [12], and Phenograph [13], all employing unsupervised clustering. Amir et al. proposed viSNE [14], a manual technique for cell phenotype identification based on dimensionality reduction using t-SNE [15], through which local structure is visible at single-cell resolution. However, manual identification of phenotypically distinct clusters of cells is cumbersome and time-consuming. In previous work [4, 16], we developed Cytosplore, following the progressive visual analytics paradigm [17, 18], to allow interactive phenotype specification and adjustments. Here, we extend the concepts presented in the original work on Cytosplore to support the analysis of Imaging Mass Cytometry data.

To the best of our knowledge, the only integrated tool for phenotype identification and microenvironment analysis for Imaging Mass Cytometry data is histoCAT [19]. histoCAT provides a relatively fixed analysis pipeline, based on automated clustering for the phenotype identification and statistical analysis to identify samples with significant interactions among phenotypes. While histoCAT provides visualizations to present the results of the steps of the pipeline, it does not allow for fully interactive exploration, as we propose here. In particular, a key differentiator of our proposed workflow is that we link abstract information, such as protein expression or the identified phenotype, to the spatial information throughout all steps of the analysis. This makes an interactive, exploratory workflow possible that allows the user to make informed decisions at every step, including quality control, phenotype identification, and microenvironment characterization. This interactive workflow also allows for one of the main contributions of this work, the exploration of cell microenvironments and interactions. While histoCAT only allows classification of complete samples, based on automatically identified interaction patterns, here, we propose a glyph-based visual encoding to make the exploration of different patterns possible, link them to their location in tissue and finally stratify cell phenotypes further, based on their microenvironments.

To aggregate and compare different cell microenvironments we group them into motifs which we visualize with a glyph-based [20, 21] representation. Similar interaction patterns, such as protein interaction networks are often visualized with node-link diagrams. Such visualizations are insightful when the number of edges is comparable to the number of nodes or a pathway continuity should be depicted. Otherwise, the outcome is a complex and crowded network [22]. Our glyph representation for the cell microenvironment networks is inspired by Landesberger et al. [23] and Dunee et al. [24], who use motifs in order to simplify and enhance their network visualizations and for illustrating the network variance, respectively. However, these visualizations are mostly focused on the way that the data are

connected (i.e., star, clique etc.) and not on the type of the data (i.e. clusters), as is the case for the data presented in this work.

3.4. BACKGROUND

In recent years, high resolution and highly multiplexed imaging techniques have become available. State-of-the-art techniques like FISSEQ [25], smFISH [26], Padlock probes and RCA [27] allow transcriptome measurements at sub-cellular resolution. In combination with appropriate data analysis techniques these data are currently revolutionizing our perception of complex biological systems. Similarly, the newly introduced Imaging Mass Cytometry [2] produces multiplexed measurements of protein abundance at sub-cellular resolution.

3.4.1. DATA ACQUISITION

Imaging Mass Cytometry data acquisition consists of three steps. First, tissue sections are stained with antibodies conjugated to heavy metals that bind to specific proteins expressed by cells or the extracellular matrix. Second, the tissue sections are ablated spot by spot in a regular grid, where each spot represents a pixel in the resulting image. Currently the resolution of a pixel is $1\mu m$. Finally, the ablated material is guided to a traditional Mass Cytometer [9], where the abundance of each metal and thereby the expression of the corresponding protein are measured per pixel. Up to now, 40 different proteins can be simultaneously measured, but this number could increase up to 100 when appropriate metal reagents become available. Ultimately, the output of this process is a stack of gray-scale images, each depicting the abundance of one measured protein. While in principle, the values per pixel and protein correspond to the number of measured heavy metal ions, due to measurement limitations the actual values are continuous, to compensate for spillover between pixels and similar effects.

3.4.2. DATA PREPROCESSING

Although the output of the Imaging Mass Cytometry is a stack of gray-scale images with per-pixel measurements, ultimately the contained cells are of interest. In order to extract single-cell information from these images these need to be segmented in a preprocessing step. The segmentation process consists of two parts, as proposed by Shapiro et al. [19]. First, Ilastik [28], a semi-supervised machine learning technique is used to create a probabilistic classification per pixel. Subsequently, these classification masks are loaded into CellProfiler [29] which derives a segmentation mask. Based on the extracted segmentation mask, the protein abundance for each segmented cell can then be aggregated. The result is a simple table that contains the expression of all proteins per cell and the mask that relates every cell to a set of pixels in the original image. Based on the segmentation mask, we define the microenvironment of a segmented cell as the directly adjacent cells in image space.

3.5. IMACYTE

We designed and implemented our interactive visual analysis tool ImaCytE to support three main tasks;

T1 Quality Control,

T2 Cell Phenotype Identification, and

T3 Cell Microenvironment Exploration.

We follow Brehmer and Munzners task typology [30] for multi-level tasks and the extension for tasks regarding the high-dimensional data analysis [31]. In the following, we use a `mono-spaced font` when referencing this typology.

An overview of the three identified tasks is illustrated in Figure 2. During quality control (Figure 2a, Section 3.5.1) uninformative samples and proteins are identified and excluded from further analysis. Afterwards, the phenotype of each cell is identified (Figure 2b, Section 3.5.2), and finally the spatial interactions among the cells in the microenvironment are explored (Figure 2c, Section 3.5.3) by the user. In the following we discuss each of the tasks, and sub-tasks where available, in detail.

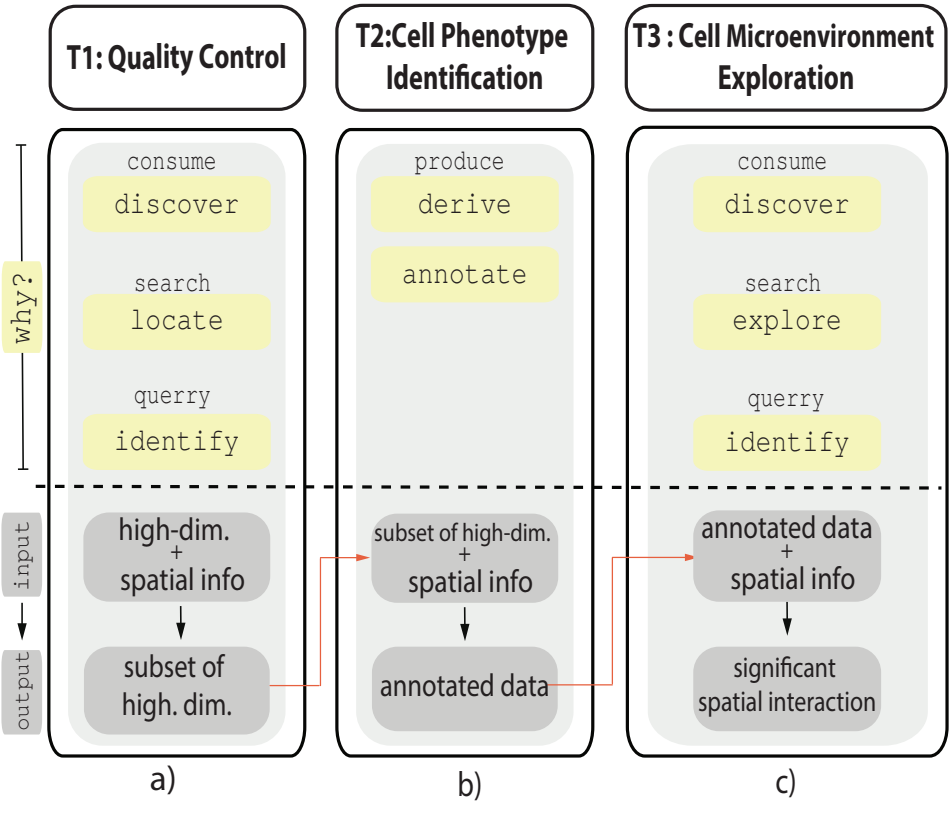

Figure 2: ***High-Level Task Overview.*** *(a) T1:Quality Control, (b) T2: Cell Phenotype Identification and (c) T3: Cell Microenvironment Exploration.*

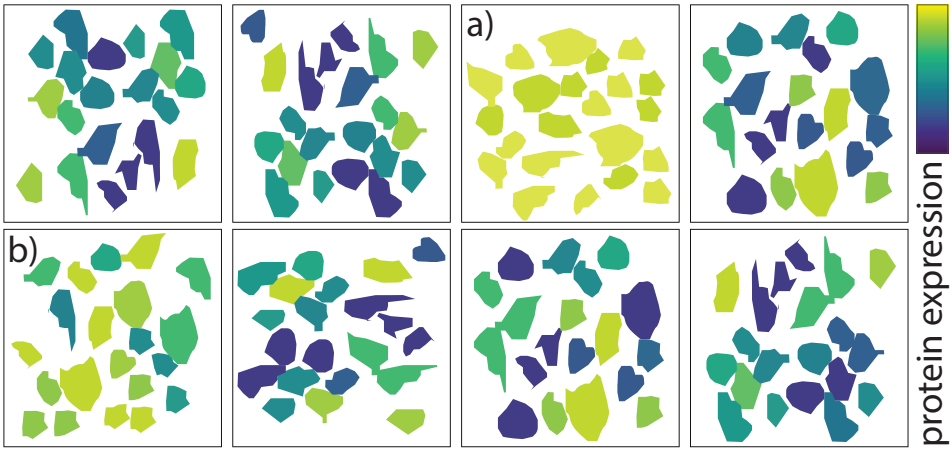

Figure 3: ***Batch Effects*** *visible in small multiples view. a) Overly expressed marker (all cells indicate close to maximum expression). b) Offset caused by background staining (no low expression cells).*

3.5.1. T1: Quality Control

Imaging Mass Cytometry data are acquired in patches where an edge typically measures approximately one millimeter. Larger tissue regions can be combined from multiple patches, or samples from different regions are combined for cohort analysis. Combining multiple samples in this fashion makes the data prone to batch effects, caused by differences in handling or preservation of the tissue and are indicated by variation in the staining efficiency between samples as illustrated in Figure 3. Typical examples are background staining [32] causing an offset in protein expression (Figure 3b) or overly abundant expression of a protein caused by universal binding of the protein (Figure 3a).

Typically, protein markers are validated separately by conventional immunochemistry before being used in Imaging Mass Cytometry, to make sure that expression patterns match those reported in literature. However, not all markers work equally well in the staining process, leaving uncertainty whether they are functional in Imaging Mass Cytometry. In order to evaluate their functionality, the user must be able to verify each marker and its expression patterns for all samples.

Abstraction. During Quality Control (Figure 2a) the user `discovers`, `locates`, and `identifies` samples as well as protein markers that need to be excluded from the analysis. The input is the complete dataset as described in Section 3.4.2 consisting of the cells and their high-dimensional marker expression in combination with their location in the tissue. The output is a set of samples and markers to be used for further analysis.

ImaCytE. The core of quality control is the visualization of marker expression for all cells in their spatial context. In the most simple case the user can select a sample alongside a protein marker to visualize the expression of that marker in the spatial context. We use the original cell segmentation mask and use color-coding to encode the values of a selected protein expression for a selected sample as more effective channels such as position [33] are already blocked by the spatial nature of the data. By default, we use the viridis [34] colormap since it provides a perceptually uniform representation of the data, minimizes

contrast artifacts, and is color blindness friendly. Inspecting one sample and one marker at a time to identify differences between samples/markers heavily relies on user memory and can be challenging. Therefore, we provide the user with the possibility to additionally select either the expression of a single protein for one or more samples (`discover` batch effects) or the expression of one or more proteins for a specific sample (`discover` broken markers) in a small multiples [35] view as illustrated in Figure 3. However, selection of multiple proteins alongside multiple samples is not possible.

In a typical exploration session, the user first selects a protein of interest and all samples for display in the small multiples image viewer. Then, she `locates` and `identifies` potential samples where the staining process was not effective. Such samples can then either be excluded from further analysis or adjusted in external tools. To verify working markers, the user selects a sample and displays the different markers for the same sample. In case a non-functional marker can be identified in this process it can be ignored for further computational steps, such as clustering or dimensionality reduction.

3.5.2. T2: CELL PHENOTYPE IDENTIFICATION

Task T2, illustrated in Figure 4, describes the identification and labeling of each cell with its corresponding phenotype. Here, the phenotype is defined by the expression of different proteins by the cell similar to conventional Mass Cytometry. Therefore, we adapt the phenotype identification process described in our earlier work [4] to the specific needs of Imaging Mass Cytometry. In particular, since we generally observe significantly fewer cells in Imaging Mass Cytometry (tens of thousands compared to millions in conventional Mass Cytometry), we can skip the lineage delineation step [4, Task 1]. Instead, we extend the phenotype identification step [4, Task 2] by a semantic grouping of subsets and the possibility to inspect the resulting phenotype definition directly in the tissue images.

As illustrated in Figure 4, T2 is further divided into three sub-tasks. The first step, T2.a, is dimensionality reduction of the data for visualization. In the second step, T2.b, we then cluster the dimensionality reduced data to define groups of phenotypically similar cells. Finally, in step T2.c, the clusters are verified, labeled, and semantically grouped. In the following, we describe each sub-task separately. Since tasks T2.a and T2.b are identical to our previous work we provide a brief description only.

3.5.2.1. T2.A: DIMENSIONALITY REDUCTION

Abstraction. Here, we `derive` a two-dimensional embedding of the cells for visualization. The input are the cells (data-points) originating from the samples selected in task T1 alongside the markers (high-dimensional space), identified to be functional in T1.

ImaCytE. The goal of the dimensionality reduction is to provide the user with a low-dimensional representation of the data that allows the visual identification of phenotypically similar cells. As described in our previous work [4], t-SNE-based approaches [15] are a good choice for this task, as they aim to preserve neighborhoods of the original space in the low-dimensional space. Linderman and Steinerberger show that t-SNE preserves well separated clusters in high-dimensional data and relates to spectral clustering in the high-dimensional space [36]. This means that cells that are similar according to their high-dimensional protein expression will be close together in the low-dimensional embedding, making such groups easy to identify, for example in a two-dimensional scatter plot. We use

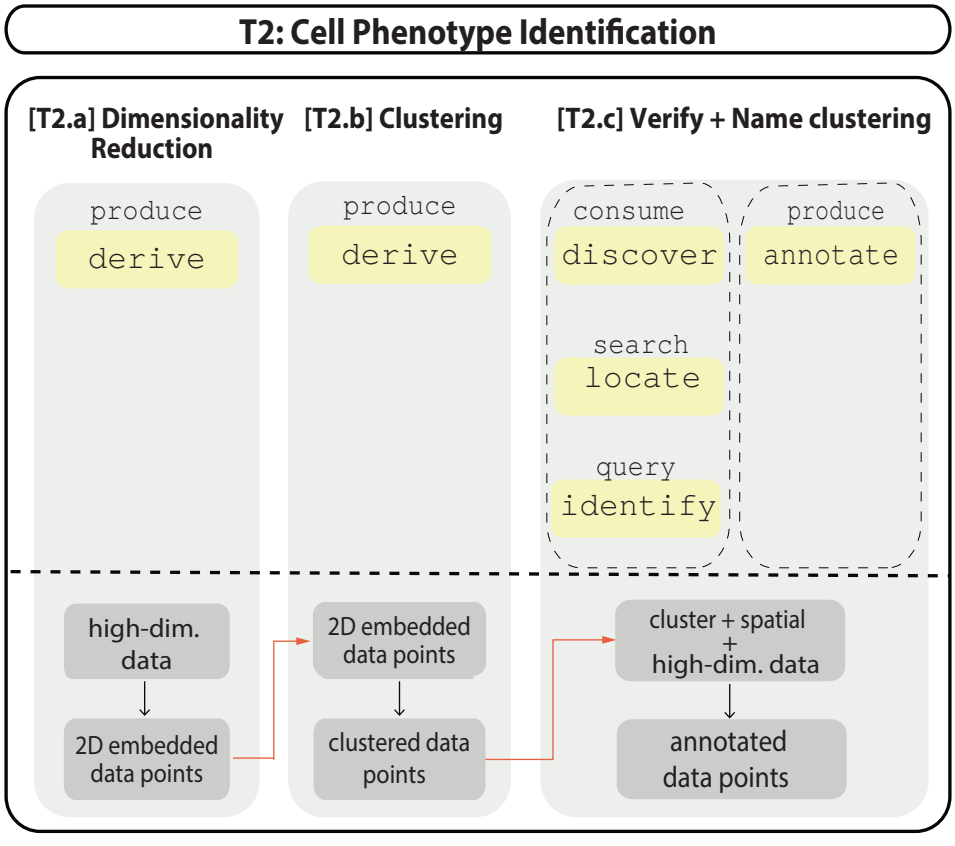

Figure 4: ***The Cell Phenotype Identification Task*** *is divided into three sub-tasks, dimensionality reduction, clustering, and verification and annotation.*

the A-tSNE algorithm [37] to `derive` such an embedding, as it provides highly similar results as the original t-SNE in a fraction of the time, making it viable for interactive exploration.

3.5.2.2. T2.b: Clustering

Abstraction. In addition to the visualization of the cells according to their similarities we also want to `derive` clusters of similar cells for quicker labeling in the following step. Here, we use the dimensionality-reduced data created in the previous step as input, to make sure derived clusters match the visually identified groups in the previous step. However, in principle other forms of clustering using the complete high-dimensional information are possible. Those would require the high-dimensional data as input. The output is the clustered data, i.e., each cell annotated with a cluster id.

ImaCytE. In our implementation, we use the Mean-Shift algorithm [38] for clustering the dimensionality reduced data. Mean-Shift clustering is density-based and therefore the resulting logical clusters closely resemble visual clusters in the embedding. Furthermore, it can extract arbitrarily shaped clusters and does not require the specification of the number

of clusters in advance. Given an existing two-dimensional embedding, it can be computed very quickly. All these properties make it a suitable technique for defining clusters in this setting. It should be noted, however, that there are potential issues the user needs to be aware of. In brief, the two-dimensional embedding might not be able to sufficiently extract all meaningful structure of the high-dimensional space, resulting in heterogeneous visual clusters. Furthermore, particularly for very large data, the optimization process of t-SNE tends to end in local minima, 'tearing' high-dimensional clusters into two or more visual clusters. For a more detailed discussion, we refer to our previous work [4]. As indicated above, in principle any kind of clustering could be employed in this step and combined with the visualization in the two-dimensional embedding and in our implementation, we allow loading of clusterings created outside of our application to replace this step if desired. In practice, however, we [4, 16] and others [14, 39] have found that clustering t-SNE embeddings of Mass Cytometry data produces biologically meaningful results and in case structure is missed our interactive workflow allows for adjustments, such as merging of similar clusters in the next step.

3.5.2.3. T2.c: Verify and Name Clustering

Abstraction. Task T2.c combines cluster verification and annotation. Based on the complete aggregated protein expression profile the user can `discover`, `verify` and `annotate` the derived clusters. Compared to its counterpart in our previous work, we add the possibility to `verify` the clustering in relation to the tissue and introduce a semantic, two-level `annotation` scheme.

The input are the clustered data-points from the previous step alongside their spatial as well as high-dimensional feature-space information. The output is a semantic label for each cell.

ImaCytE. The basis for the implementation of task T2.c again forms our implementation in Cytosplore [4, Task 2c]. As in our previous work, the expression of one protein can be used to color-code the embedding allowing the user to asses the homogeneity of the clusters with respect to the given protein. Furthermore, we use a cluster heatmap, showing the median expression of each protein per cluster allowing the biological interpretation of each cluster. Through the heatmap, the user can also directly adjust the clustering. When multiple clusters with a similar biological interpretation have been extracted they can be merged. This process is supported by sorting the clusters in the heatmap by similarity. Therefore, we compute a hierarchical clustering, based on the median expression values of each cluster and add a dendrogram to the heatmap.

We extend our original work in two ways. We introduce a semantic, two-level color-coding for the identified clusters, separating main cell types and sub-types. Making use of the spatial nature of the data, we extend our previous work with the possibility to verify all steps by visualizing the cluster result in the image view.

Since the number of cells in typical Imaging Mass Cytometry datasets is much smaller than in conventional Mass Cytometry, we do not need to divide the data as in our previous work [4, Task 1]. To retain the semantic separation of the data into main types, i.e., cell lineages, and sub-types of cells, we create a two-level color scheme illustrated in Figure 5. Based on the median expression of the derived clusters we compute a hierarchical clustering. Afterwards, a user-defined dissimilarity threshold (Figure 5b), based on the Euclidean distance in the cluster hierarchy, is used to separate the hierarchy into the main

cell types (clusters with large Euclidean distance, branching above the threshold) and sub-types (clusters with small Euclidean distance, branching below the threshold). We then assign semantic color-coding to the clusters according to the their separation. We use the D3 [40] categorical colormap to assign primary colors to the main partitions (Figure 5c, red, blue, green) and vary the saturation of those colors within a partition (Figure 5c, shades of blue) to demonstrate the correlation of those clusters. Finally, we provide the option for the user to manually assign specific colors to the major groups to comply with the phenotypes and their biological significance.

As described above, the second addition is the introduction of the spatial information for verification of the resulting clusters. Here, in particular, we color the cells in the image view according to the color-scheme described above to verify the biological significance of the derived clusters. For example, if staining issues were missed in the quality control step it can happen that several cluster with the same biological meaning are created. Figure 6 illustrates such a case. Clusters C1 and C2 show similar protein expression patterns, with generally larger values for C2. Both clusters are visible in sample 1, indicating that they might represent the same phenotype with a large dynamic range. Sample 2 only exhibits cells from cluster C1, indicating that staining efficiency in that sample was worse than for sample 1 resulting in smaller dynamic range. With this information, the user can now decide to either go back to the quality control step and further investigate sample 2 or simply merge the two clusters to indicate their biological similarity.

All steps of Task T2 are completely interactive and linked, allowing iterative refinement of the clustering. Changes in clustering automatically trigger updates of the color-coding of the tissue in the spatial view, the embedding scatter plot and the heatmap. Thereby, the user can verify clustering results immediately across several linked views. Furthermore, the user can interactively probe in the heatmap, embedding, or spatial view through linked selections. Selecting one or multiple cells or clusters in any of the three views highlights the cells in any of the other visualizations. Highlighting in the heatmap and image views is achieved by fading out the non-selected area. In the scatter plot we change the mark from circles to crosses for the selection.

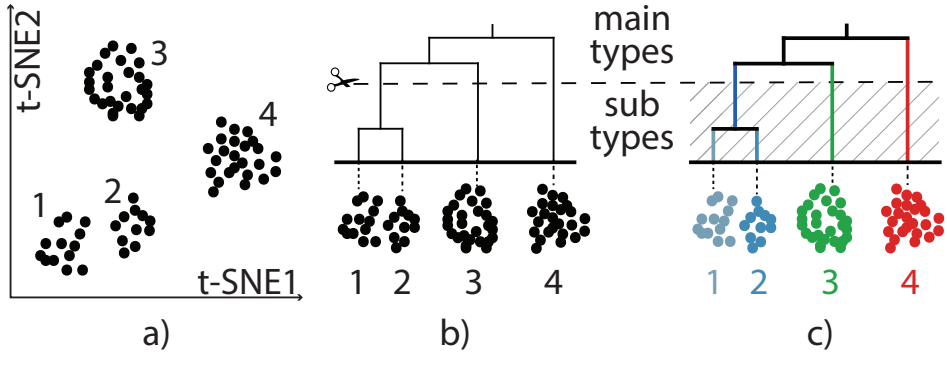

Figure 5: ***Two-Level Color-Coding***. *Clusters from the embedding, a), are hierarchically clustered, b), and a user-defined threshold is used to define main- and sub-types. Finally, color hues are assigned to the main-types and different saturation is applied to the corresponding sub-types, c).*

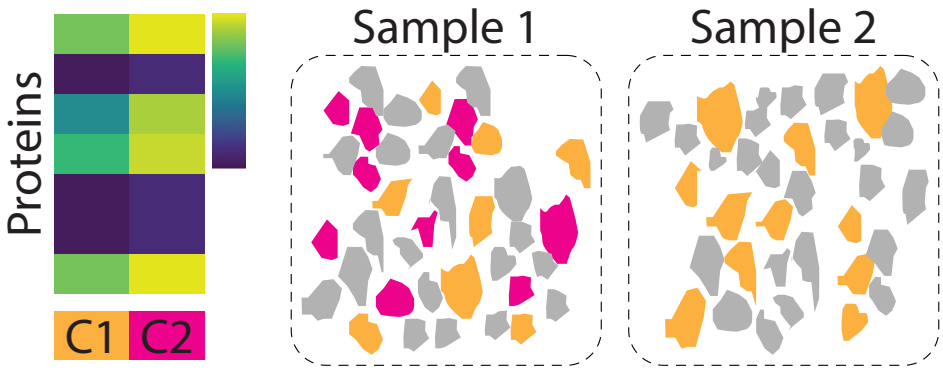

Figure 6: ***Variation in Staining Efficiency*** *becoming apparent during cluster verification. Two clusters, C1 and C2, are identified, however, their protein expression, visualized in the heatmap, indicates an offset that can be caused by insufficient staining. Visualizing the clusters in their spatial context (grey indicates other cells) reveals that one of the clusters is indeed only existent in one of the samples and is likely caused by insufficient staining.*

3.5.3. T3: CELL MICROENVIRONMENT EXPLORATION

Task T3 (Figure 7) is made possible due to the spatial resolution of the data and comprises the exploration of the cell microenvironment. In an overview first, detail on demand fashion, the task is divided into two sub-tasks. The first sub-task, T3.a Spatial Interactions Overview, is to gain an overview of which phenotypes interact with each other on a global scale. These interactions characterize the tissue function and can be helpful in order to categorize the tissue samples. However, a phenotypical subset can include thousands of cells with distinct local cell microenvironments, meaning subsets within a phenotype that have different spatial interactions from those illustrated in the overview are likely to exist. The differentiation of those subsets is of major interest to identify functionality of cells, beyond their phenotype. Sub-task T3.b, Spatial Interactions Details, describes the exploration of details of the interaction patterns on a local microenvironment scale.

For both sub-tasks it is important to link the abstract results of spatial interactions to the spatial location of the cells in the tissue. As an example our collaborators described *floating cancer cells*. Floating cancer cells are cancer cells that abandoned a tumor. These cells are prone to attacks from immune cells and as a result exhibit diverse microenvironments. Finding cells with a similarly diverse microenvironment within the tumor would have completely different biological implications. Therefore, it is important to provide the means to inspect the abstract results linked to tissue location. In summary, this example illustrates that we need to identify subsets of a phenotype with unique microenvironment characteristics, connect them to their spatial location and examine their characteristics separately, as they can be useful biomarkers for tissue functionality.

3.5.3.1. T3.A: SPATIAL INTERACTIONS OVERVIEW

Abstraction. For sub-task T3.a, the user `discovers` cells of which phenotypes interact spatially, `explores` how much they interact and `identifies` which of the interactions are significant. The input of this sub-task are all cells alongside their location in the tissue and their phenotype identified in Task T2. The output is the quantified amount of pairwise

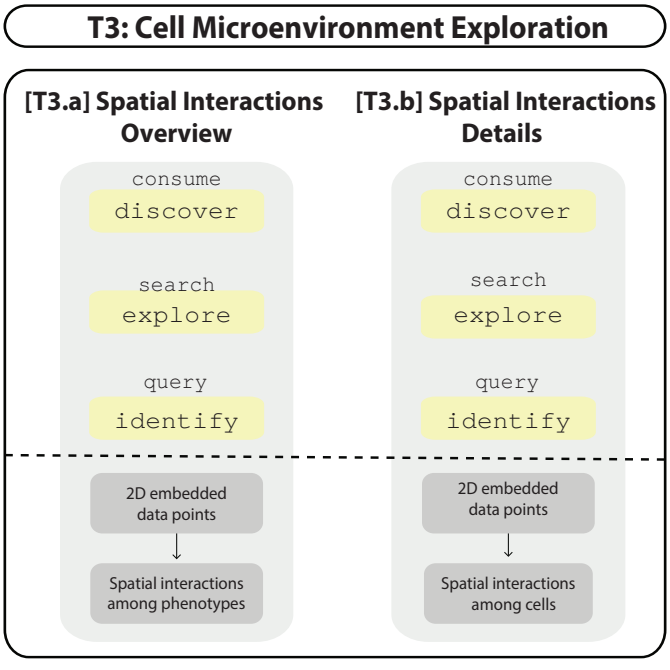

Figure 7: ***The Cell Microenvironment Exploration Task****, including two sub-tasks. The first generates an overview of the spatial interactions among phenotypes. Afterwards, spatial interaction details on the cellular level are explored.*

spatial interaction between phenotypes.

ImaCytE. Here, the goal is to provide a global overview of which phenotypes can be found in the microenvironment of cells with a specific phenotype. Therefore, for every cell we count the number of occurrences of each phenotype in its direct microenvironment and aggregate these counts over all cells of the same phenotype. The result is a directed and weighted graph, where each node represents a phenotype. The weighted link between two nodes illustrates the number of times a cell of the phenotype at the end of the link occurs in the microenvironment of a cell of the phenotype at the source of the link. In our current studies, the number of nodes n is typically in the range of $20-40$. While pruning could be used to remove links with low weights, typically all nodes are connected to all other nodes, resulting in up to n^2 links. While intuitive to read, a node link diagram would be too dense and suffer in terms of readability [41]. Therefore, we chose a heatmap (Figure 8) to visualize the resulting graph, with the source nodes on the horizontal axis and the target nodes on the vertical axis. The weight or number of interactions is then indicated using color in the cell at the intersection of the source and target node. To not interfere with the qualitative colormap used to identify the different phenotypes, we use a white-to-black colormap to indicate the weight. The values in the heatmap are normalized. The user can choose different normalization criteria for the heatmap such as per row or per column. If the user is interested firstly in the different phenotypes that are most likely to exist in the microenvironment of a specific source cell normalization per column is

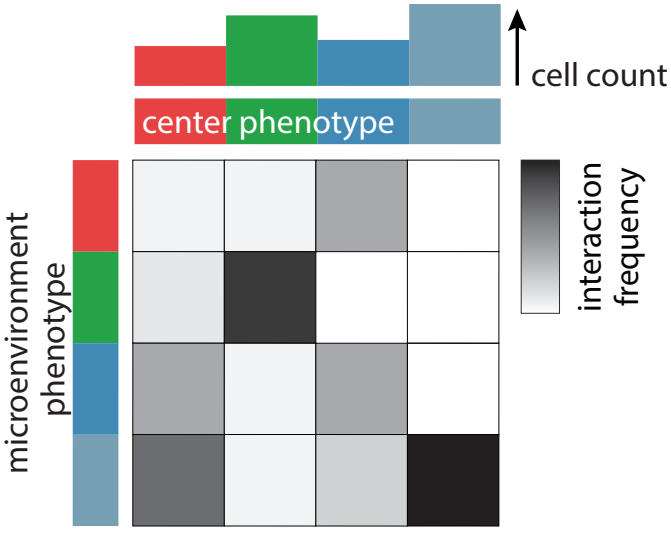

Figure 8: ***The Interaction Heatmap****, normalized per column, illustrates the relative frequency that a phenotype of interest (phenotypes in horizontal axes) interacts with a phenotypes existing in its microenvironment (phenotypes in vertical axes). A bar chart on top of the heatmap depicts the amount of cells for each phenotype.*

the most straightforward to read. Complementary, normalization per row allows to easily identify in whose microenvironment a specific phenotype most likely exists.

We label the rows and columns of the heatmap with the colors, assigned to the clusters in task T2 and sort it with the same sorting, based on hierarchical clustering, used to define the color-scheme. We indicate the frequency of occurrence of each phenotype by a bar on top of the heatmap, where height indicates the number of occurrences of the given type in linear scale. The resulting bar chart allows to put the significance of the interactions in context to occurrence of the specific phenotype. We also use the heatmap to filter the detail interactions described in the following Section 3.5.3.2. In short, selecting an interaction in the heatmap highlights the individual cells forming this interaction in the tissue and the visualization of the spatial interaction details view is filtered according to the cell phenotypes of that interaction. Selection is possible for individual interactions between two specific phenotypes (a single box in the heatmap) or whole columns or rows (all interactions including a specific phenotype) as well as combinations of the above.

3.5.3.2. T3.B: SPATIAL INTERACTIONS DETAILS

The goal of sub-task T3.a is to provide the user with a general overview of the types of cells that interact. Considering the example in Figure 8, we can see that cells in the green cluster interact with other cells from the same cluster at high frequency, but rarely with any other types. Here, the goal is to provide detailed insight into differences within local microenvironments, for example to find out whether interactions between cells of certain phenotypes attract a third type to the same environment.

Abstraction. In sub-task T3.b, the user `explores` local details of spatial interactions

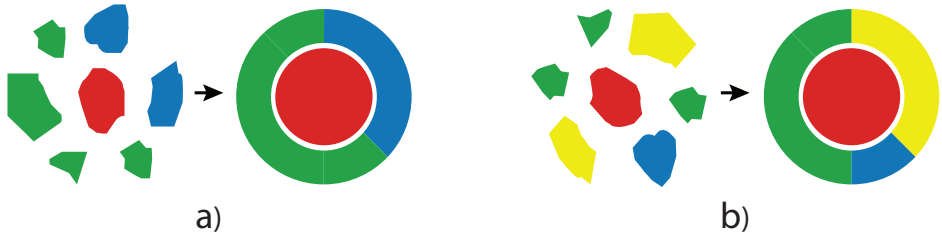

Figure 9: ***Abstract Representation of Unique Microenvironments.*** *We represent different unique microenvironments by a donut chart around a representative for the center cell. The relative frequencies of cells are preserved by the fractions in the donut chart.*

and `identifies` subsets of a specific phenotype that differ qualitatively or quantitatively with regard to their microenvironments. The input is the same as for sub-task T3.a; All cells alongside their location in the tissue and their phenotype identified in Task T2. The output is a categorization of localized microenvironments.

ImaCytE. In a typical exploration we might observe several thousand microenvironments consisting of unique combinations and quantities of different phenotypes. In order to support the detailed `exploration` and `discovery` of interesting microenvironments and `identify` their significance, we aggregate these environments to motifs. The biological importance of a microenvironment is defined by two parameters. Most important is which cell types interact, while the quantities of interacting cells are of secondary interest. Therefore, we define a motif to represent all cells of the same phenotype with identical combinations of phenotypes in their microenvironment, irrespective of their quantities.

To explore the resulting motifs, we designed a glyph based on a simple donut chart, abstracting the microenvironment. Two examples for mapping the microenvironment of a single cell to the glyph are illustrated in Figure 9. In principle, the cell of interest is indicated by a circle, colored according to the cells phenotype, in the center of the glyph, while the occurrences of different phenotypes in the cells microenvironment are summarized by a donut chart around it. We chose this basic representation for several reasons. First and foremost, the circular layout around the cell of interest mimics the actual layout of the cell and its microenvironment in the tissue and is therefore intuitive to understand. As indicated above, it is more important which cells interact, than the actual quantities. Hence, we decided that the intuitive representation outweighs potentially harder readability of the frequencies, compared to, for example, a bar chart-based design. Furthermore, we typically observe no more than three to four different phenotypes in each unique microenvironment making the donut a reasonable choice [42] for comparing frequencies.

Thus far, the described design is a direct representation for a single cell and its microenvironment. To use it for visualizing the detected motifs, we need to extend it to be able to represent multiple microenvironments consisting of the same phenotypes but at different quantities. We illustrate the process in Figure 10.

Different cells can be of largely different size and shape. Consequently, the number of cells in their microenvironment can vary significantly from none up to tens of cells. Therefore, when aggregating multiple cells, we only retain the relative frequency of different

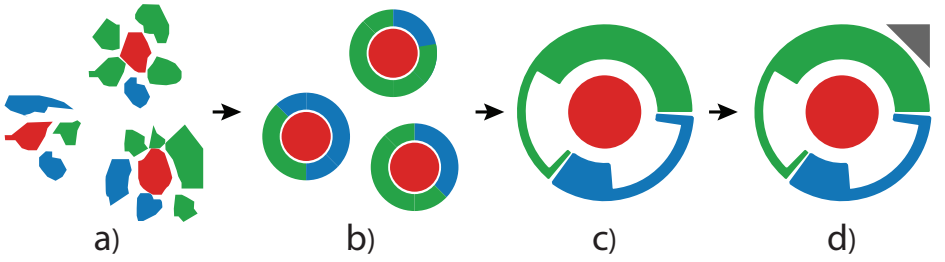

Figure 10: ***Glyph Design****, representing a motif of a unique microenvironment (only blue and green cells) for a red cell, a). b) each microenvironment can be abstracted as described in Figure 9. In c), we combine the different instances of the motif to a single glyph, showing the mean frequencies and variation. Finally, in d) we add an indicator for the significance of the motif.*

phenotypes, rather than their absolute numbers. As such, in the first step of creating the motifs and their corresponding glyphs, we aggregate all unique microenvironments of a given phenotype to their relative frequency version (Figure 10a→b). In the second step we create an aggregate glyph (Figure 10c) that uses the mean relative frequency of all microenvironments corresponding to the motif for the segments of the donut. We compute the standard deviation of the relative frequencies for each phenotype in the microenvironment, to indicate the variation. There is no standard way to visualize error or variation within donut charts. Inspired by the work of Gove and Herzog [43] who reduce the amount of paint in a box in a heatmap to indicate uncertainty, we cut out parts of the arcs in the donut to indicate the standard deviation. A small cutout, leaving a lot of paint in the segment, indicates little variation while a large cut out indicates large standard deviation. In rare cases when the standard deviation is larger than the mean value we limit the cut out to the size of the segment, leaving only the outline. To indicate the frequency occurrence of this motif, we scale the inner circle according to how often this motif appears in relation to all other motifs with the same center phenotype. Finally, we add a triangle to the top-right of the glyph (Figure 10d) to indicate significance of each motif when shown in a grid-based layout. We compute the significance of each motif using a permutation test following Hooton et al. [44]. The fill color of the triangle indicates the significance according to a white to black color-scale.

During the design phase of the glyph, we have considered other visualization designs. Typical bar chart-based design would potentially provide better comparability between motifs, however lack the intuitiveness of the circular representation, as indicated above. Other radial layouts include radar charts, or radial box-plots. While these techniques would provide very exact representations of the quantitative aspects of the microenvironments, such a detailed representation is ultimately not necessary and could introduce significant clutter, especially when laid out as small multiples.

For the exploration of the detailed cell microenvironments we lay out the glyphs in a small multiples view. We provide multiple filter options to focus on specific phenotypes and reduce the number of displayed glyphs. To remove outliers and noise, the user can define a minimum number of cells and their environments that need to be captured by a motif. Then, the user can focus onto motifs containing interactions containing specific

phenotypes by selection in the interaction heatmap, described in Section 3.5.2. Therefore, she simply clicks on a column, row, or individual boxes in the heatmap or combines multiple selections by holding the shift key. Once the motifs are filtered to a set of interest, those can be highlighted in the tissue, for example to identify whether certain interactions are specific to a type of tissue, for example, healthy or cancerous. The variation contained in each motif can also be investigated further in a secondary view. Here we only group those neighborhoods with equal qualitative and quantitative composition. I.e. this view shows motifs such as those in Figure 9 and Figure 10b without any variation and indication thereof.

3.5.4. IMPLEMENTATION

The presented prototype is the result of an iterative design process, carried out in close collaboration with domain expert partners (co-authors of this manuscript) from the Immunohematology department at Leiden University Medical Center (LUMC). Figure 1 shows the user interface of ImaCytE, implemented in MATLAB as a stand-alone application. In the repository [45] source code and binaries are available.

3.6. CASE STUDY

Here, we demonstrate the effectiveness of our prototype by an exemplary visual analysis of Imaging Mass Cytometry data. To verify that the design decisions made were not limited to the specific needs of a single collaborator we carried out this case study with another collaborating partner, who was not involved in the design process of the system. The expert is a PhD candidate at the Pathology department at LUMC and is interested in the interaction between immune and cancer cells.

For the case study, she acquired eight imaging Mass Cytometry samples from different regions of interest. Each sample approximately covers the area of 1mm^2 resulting in images in the range of 1.000×1.000 pixels. The abundance of 40 proteins was measured for each sample, resulting in 40 values per pixel in the resulting images. We preprocessed the acquired data as described in Section 3.4. The total number of segmented cells from the given samples is 23682.

For the case study we provided the expert with a detailed introduction to the software after which she was able to explore the data independently. Since, at this point, we were not aiming at quantitative performance measurements we supported her in the analysis, whenever questions about the software arose.

3.6.1. QUALITY CONTROL

In the quality control step of the analysis the expert was interested in identifying samples affected by batch effects and low signal markers. In order to verify the robustness of the measurements, she selected Vimentin and Keratin, two proteins that are expressed mutually exclusive by stromal and cancer cells, respectively, and therefore provide a predictable expression pattern. After the selection, the samples were color-coded, as illustrated for four samples in Figure 11a. The expert noticed that regions enriched for Vimentin (bright yellow color, Figure 11a, top row) were mutually exclusive from regions enriched for Keratin (Figure 11a, bottom row), consistent with her previous knowledge and indicating proper

staining for these markers. Subsequently, by exploring the different protein expressions individually for all eight tissue samples, she observed that protein 19, was moderately expressed in all analyzed samples, except for sample 7 (Figure 11b), where it was expressed with high abundance. These observations indicate that this sample is a necrotic part of the tissue, causing abnormal antibody binding. Based on this discovery, the expert removed the sample from further analysis.

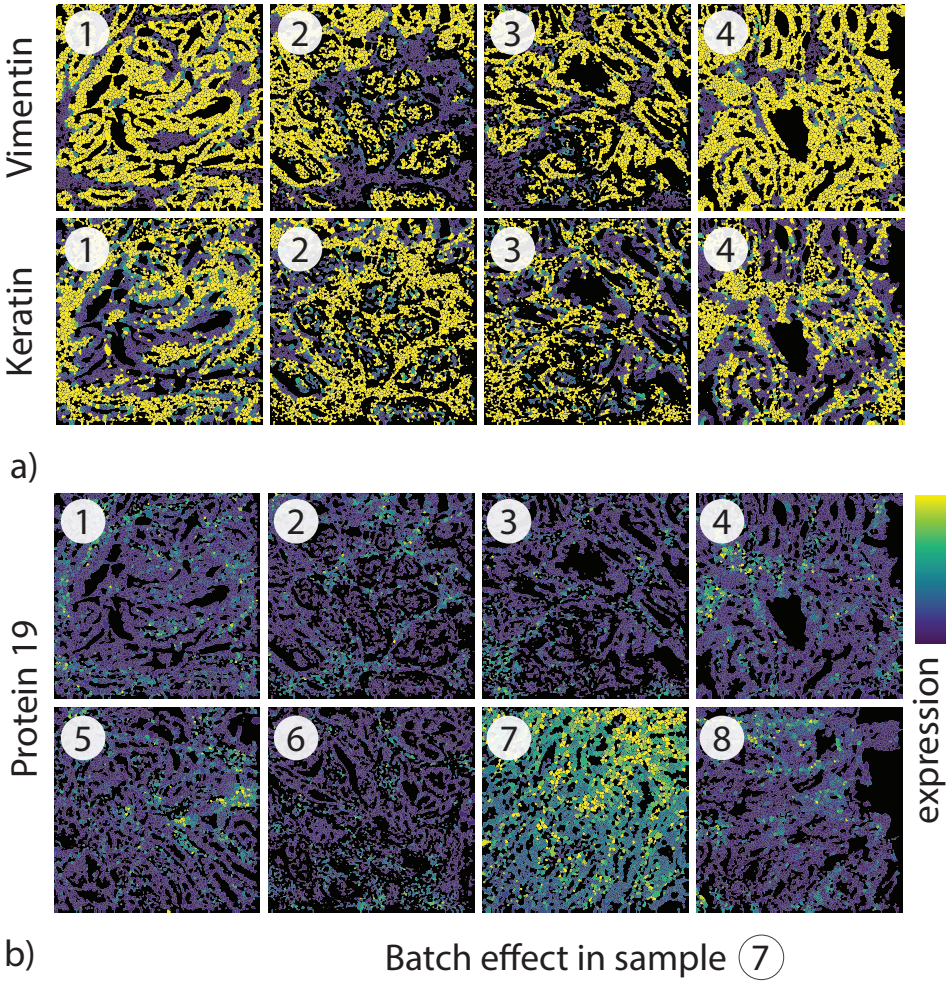

Figure 11: ***Quality control overview*** *The mutually exclusive expression profiles of Vimentin and Keratin for the first 4 samples is illustrated in a). Protein 19 is overly expressed in sample 7 indicating a potential batch effect, b). The identical circled numbers indicate identical tissue samples throughout this figure and Figure 12c.*

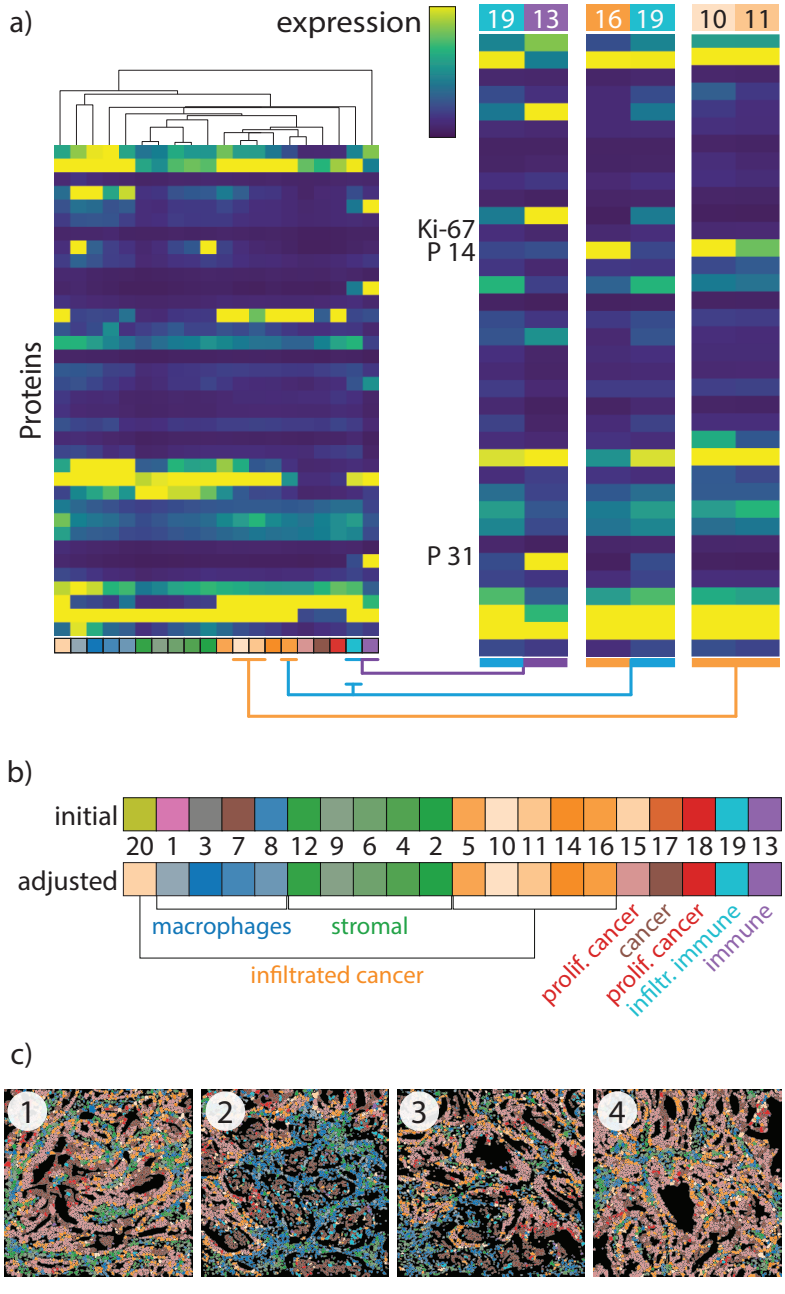

Figure 12: ***Clustering Results.*** *a) shows the heatmap with different identified clusters and zooms of several clusters for comparison. b) shows the initial, automatically created color-map and the user-adjusted final version and the corresponding phenotypes. Phenotypes are indicated in the tissue with the same color-coding, c).*

3.6.2. CELL PHENOTYPE IDENTIFICATION

The next step of the analysis is the identification and labeling of similar cell phenotypes. First, the expert selected the proteins to be used for dimensionality reduction. I.e., proteins such as protein 3 that was found overly abundant in all samples were removed from the input set. Furthermore, as the implemented A-tSNE produces results in a short time the expert experimented with different protein selections to acquire a deeper understanding of the correlation between existing phenotypes. After selecting the proteins that she would eventually use and computing the final A-tSNE embedding, she started the iterative process of cluster verification. The implemented Mean-Shift clustering is dependent on the kernel bandwidth of the density estimator. To make sure to not miss any small but important clusters, the expert adjusted the kernel bandwidth to capture the smallest clusters that were discernible on the scatter plot. The automatic color scheme generation which roughly depicts the semantics between the clusters provided the expert with an interactive illustration of all the phenotypes that exist in the samples alongside their location (Figure 12c). Considering the main phenotypes that she identified in the tissue, she started merging semantically-related clusters.

The merging procedure was carried out by identifying similar clusters, according to their corresponding protein expressions, in the heatmap and verifying overall homogeneous expression within those clusters in the embedding view. The linked selections in the heatmap and the scatter plot made it easy to to combine those two steps as selected clusters were automatically highlighted in the either view. The first step is identifying similar clusters that are potential merging targets. Here, the hierarchical clustering and corresponding dendrogram, illustrated on top of the heatmap in Figure 12a, place clusters with a small Euclidean distance according to all markers next to each other in the heatmap and serve as a first suggestion for merging. Typically, a small distance indicates potential targets for merging, however, it is important to verify the complete expression, as for some proteins a small difference can be of high significance. Such a case is illustrated by the comparison of clusters 10 and 11 (zoom, orange line in Figure 12a). Clusters 10 and 11 have very similar overall expressions, but after inspection the expert decided to keep them separate, as the small difference in the Ki-67 protein could indicate a significant biological difference. In the end, the expert kept twenty phenotypically distinct clusters shown in the final heatmap in Figure 12a. Once the clusters were final, the expert adjusted the color-scheme to her needs. The initial, automatically created color-coding, illustrated in Figure 12b, top row, identified ten meta clusters and assigned different corresponding colors from the qualitative colormap with different saturation for the different sub-types. While the color-scheme captured the main differences the expert decided to simplify the categorization. Most importantly she grouped clusters 1, 3, 7 and 8 under the umbrella of macrophage-related phenotypes, as these were not of major importance in her analysis. Furthermore, most cancer cells types were classified as a single group (orange) in the original scheme. To easily differentiate between infiltrated, proliferating, and other cancer types she separated those types into three groups (orange, red, and brown, respectively). The final scheme is shown in Figure 12b, bottom row.

With our tool the expert was able to quickly identify and merge similar clusters, define a semantically meaningful color-scheme, and visualize the spatial distribution of the identified phenotypes in the tissue (Figure 12c).

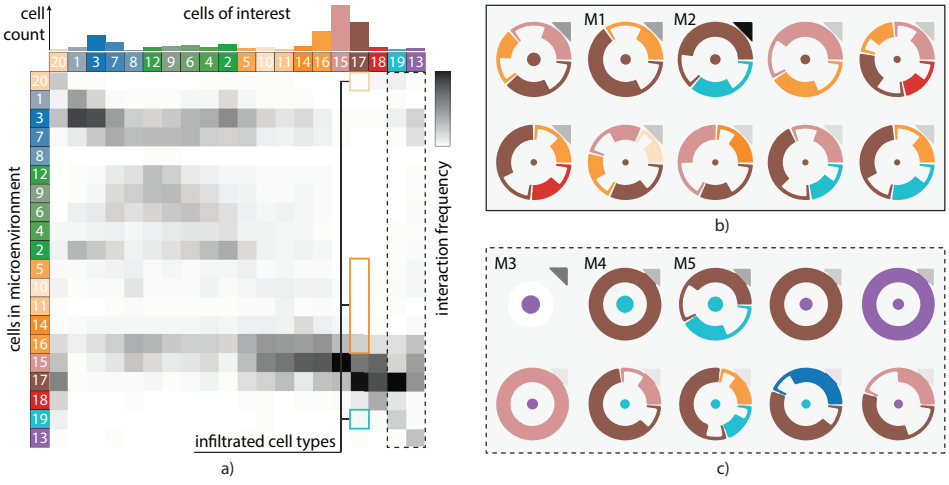

Figure 13: ***Interaction Overview and Detail Motif Glyphs.*** *a) shows the interaction heatmap, providing an overview of the interactions in the tissue. b) and c) show the ten most significant detail glyphs for the selections (solid/dashed boxes) in a).*

3.6.3. CELL MICROENVIRONMENT EXPLORATION

Having the different phenotypes identified, the expert continued with the analysis to discover spatial interactions of non-proliferating cancer cells (cluster 17, brown), as they were her main phenotype of interest. First, she inspected the interaction heatmap (Figure 13a), where she observed a high frequency of cell interactions with non-proliferating (cluster 17, brown) and proliferating cancer cells (clusters 15 and 18, red). Nonetheless, she was interested in the interactions that involved the infiltrated cancer cells (orange). More specifically, she was interested in whether infiltrated cancer cells create exclusive microenvironments with non-proliferating cancer cells, as such exclusive microenvironments could reveal a gradual deterioration of the tumor expansion. Hence, she selected the interactions between

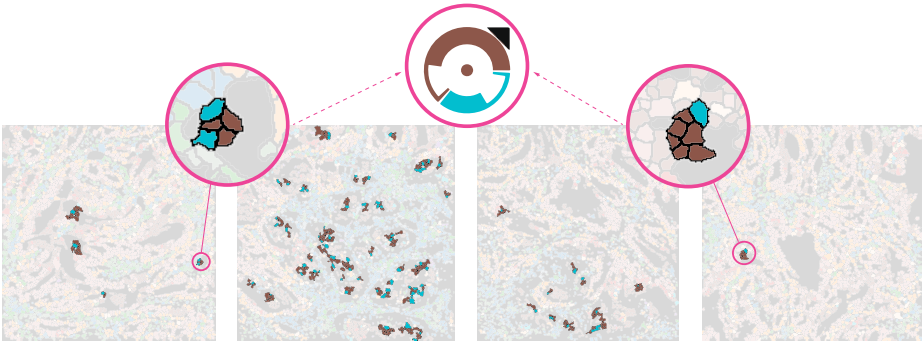

Figure 14: ***Highlighted Interaction Motifs*** *in the image view. Cellular interactions according to motif M2 (Figure 13b) are highlighted.*

cancer and infiltrated cancer and immune cells as indicated in Figure 13a, dashed line, and filtered the motifs accordingly. Afterwards, she ordered the motifs according to the frequency of their occurrence in descending order. After the reordering, she pointed out a motif with cells of cluster 19 and cluster 17 (M1, Figure 13b) and a motif with cells of cluster 16 and cluster 17 (M2, Figure 13b), as those motifs create exclusive microenvironments with infiltrated cancer and immune cells, respectively. These two motifs occur at similar rates, as can be seen by the similar sizes of the center circle. Hovering over the motifs revealed that M1 consists of 170 cells, whereas M2 consists of 168 cells. However, cluster 16 is significantly larger than cluster 19, as indicated by the difference in size of the large orange cluster 16 bar and the extremely short light-blue cluster 19 bar in the cell count plot on top of the interaction heatmap (Figure 13a). Hence, the expert hypothesized that the cells of cluster 19 create exclusive microenvironments with cancer cells with higher frequency than cluster 16 cells and that this differentiation in their functionality causes the cancer cells to stop proliferating. Going back to the protein expression heatmap and comparing clusters 16 and 19 (Figure 12a), she could identify differing expression in protein 14 for these two clusters, leading her to the hypothesis that this protein might influence the proliferation process. Finally, selecting the motifs and highlighting them in the image view, as illustrated in Figure 14, shows that the corresponding cells can be found in several tissue samples, indicating that this microenvironment is not a sample-specific artifact.

The large proportion of cells assigned to cluster 19 creating exclusive microenvironments with cancer cells stimulated the expert to explore not only the interaction of cluster 19, but also the interactions of cluster 13, corresponding to the pure version of the same immune cell type. Since, the interaction heatmap did not provide a clear difference in their interaction patterns (dashed box, Figure 13a) with other phenotypes, she start exploring the generated motifs. In order to filter only the motifs with center cells from clusters 13 and 19, she selected the columns by clicking the color-bar on top of the heatmap. The most frequent motif of cluster 13 (M3, Figure 13c) represents cells that do not have any other cells in their microenvironment (non-existing outer circle). The most significant motifs corresponding to cluster 19 (M4 and M5, Figure 13c) create exclusive microenvironments with non-proliferating cancer cells, coinciding with the observations for the cancer cell environments, particular M2, explored in the previous step. Consequently, the expert came up with the hypothesis that the main cause for the differentiation of infiltrated immune cells (cluster 19) from the pure immune cells (cluster 13) is an encounter with non-proliferating cancer cells. The comparison of clusters 13 and 19 in Figure 12b shows the significant difference in the expression of protein 31 between the two clusters. Hence, her hypothesis is that the type corresponding to cluster 19 originates from the same type as cluster 13, but when they encounter non-proliferating cancer cells they stop expressing protein 31 to create the new phenotypical subset.

3.6.4. EXPERT FEEDBACK

After finishing the case study, we collected qualitative feedback from the expert. Generally, she liked that the software *"is very straightforward and easy to use. Also, the way it allows to move back and forth between the various visualizations provides a lot of added value to the exploration"*. The semantic phenotype color-coding was *"helpful to have a rough*

estimation of the location for my phenotypes of interest and their interacting phenotypes". For the interaction analysis part of the case study, she particularly liked the interactivity of the filtering as it helped her to identify the different phenotype interactions in the tissue. The overview first, detail on demand approach, in combination with the motif visualization was insightful and *"add[s] information that [she] would otherwise lack"*. In a future release of the tool, she would like to see an extension of the quality control functionality, to provide direct adjustments, such as removal of batch effects or imputation, so that she does not have to discard problematic samples completely. Another useful addition could be the ability to validate her visual exploration findings statistically and acquire more information regarding the distribution of the proteins within each cell.

3.7. CONCLUSION AND FUTURE WORK

We presented a workflow for cell phenotype identification and microenvironment exploration of high-dimensional Imaging Mass Cytometry data and implemented the workflow in an integrated framework, called ImaCytE. ImaCytE enables the end-to-end analysis of segmented Imaging Mass Cytometry data in an interactive and data-driven manner. We introduced motifs and a glyph-based representation for the discovery of exclusive microenvironments and showed the importance of interactive analysis and linking the microenvironment exploration with the tissue images. In the presented case study, we have shown the importance of interactive quality control and the value of exploring the interactions of phenotypes as whole and subsets of phenotypes and their spatial interactions. Most importantly, we have shown that the presented workflow allows effective hypothesis generation by exploratory analysis of unknown data.

While, in principle, ImaCytE supports parallel exploration of tens of samples, our work is focused on the interactive exploration of a few samples at a time. We can imagine several extensions for future research. To identify biomarkers causing differentiation between patients, comparative approaches would be helpful. To scale to large amounts of samples, or to add new samples acquired in a clinical setting, the results of an interactive exploration could be used to train a model for automatic classification of more samples.

ACKNOWLEDGEMENTS

We would like to thank M.E. IJsselsteijn and Dr. N.F. Miranda for performing the biological experiments, B. van Lew for narrating the supplemental video, N. Li and N Guo for their valuable comments on our tool. Moreover, we would like to thank the anonymous reviewers for their constructive criticism that lead to significant improvements of the paper and the application. This work received funding through Leiden University Data Science Research Programme. B.P.F.Lelieveldt received partial funding from H2020-Marie Skłodowska-Curie Action Research and Innovation Staff Exchange (RISE) Grant 644373-PRISAR.

REFERENCES

[1] P. Mazzarello, "A unifying concept: the history of cell theory," *Nature Cell Biology*, vol. 1, no. 1, pp. E13–E15, 1999.

[2] C. Giesen, H. A. Wang, D. Schapiro, N. Zivanovic, A. Jacobs, B. Hattendorf, P. J. Schüffler, D. Grolimund, J. M. Buhmann, S. Brandt, *et al.*, "Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry," *Nature methods*, vol. 11, no. 4, p. 417, 2014.

[3] N. Crosetto, M. Bienko, and A. Van Oudenaarden, "Spatially resolved transcriptomics and beyond," *Nature Reviews Genetics*, vol. 16, no. 1, p. 57, 2015.

[4] T. Höllt, N. Pezzotti, V. van Unen, F. Koning, E. Eisemann, B. P. F. Lelieveldt, and A. Vilanova, "Cytosplore: Interactive immune cell phenotyping for large single-cell datasets," *Computer Graphics Forum (Proceedings of EuroVis)*, vol. 35, no. 3, pp. 171–180, 2016.

[5] M. Meyer, T. Munzner, A. DePace, and H. Pfister, "Multeesum: A tool for comparative spatial and temporal gene expression data," *IEEE Transactions on Visualization and Computer Graphics (InfoVis '10)*, vol. 16, no. 6, pp. 908–917, 2010.

[6] W. M. Abdelmoula, B. Balluff, S. Englert, J. Dijkstra, M. J. Reinders, A. Walch, L. A. McDonnell, and B. P. Lelieveldt, "Data-driven identification of prognostic tumor subpopulations using spatially mapped t-sne of mass spectrometry imaging data," *Proceedings of the National Academy of Sciences*, vol. 113, no. 43, pp. 12244–12249, 2016.

[7] F. Hillenkamp and J. Peter-Katalinic, *MALDI MS: a practical guide to instrumentation, methods and applications*. John Wiley & Sons, 2013.

[8] J. Salamon, X. Qian, M. Nilsson, and D. J. Lynn, "Network visualization and analysis of spatially aware gene expression data with insitunet," *Cell systems*, 2018.

[9] D. R. Bandura, V. I. Baranov, O. I. Ornatsky, A. Antonov, R. Kinach, X. Lou, S. Pavlov, S. Vorobiev, J. E. Dick, and S. D. Tanner, "Mass cytometry: technique for real time single cell multitarget immunoassay based on inductively coupled plasma time-of-flight mass spectrometry," *Analytical chemistry*, vol. 81, no. 16, pp. 6813–6822, 2009.

[10] N. Samusik, Z. Good, M. H. Spitzer, K. L. Davis, and G. P. Nolan, "Automated mapping of phenotype space with single-cell data," *Nature methods*, vol. 13, no. 6, pp. 493–496, 2016.

[11] E. R. Zunder, E. Lujan, Y. Goltsev, M. Wernig, and G. P. Nolan, "A continuous molecular roadmap to ipsc reprogramming through progression analysis of single-cell mass cytometry," *Cell Stem Cell*, vol. 16, no. 3, pp. 323–337, 2015.

[12] S. Van Gassen, B. Callebaut, M. J. Van Helden, B. N. Lambrecht, P. Demeester, T. Dhaene, and Y. Saeys, "Flowsom: Using self-organizing maps for visualization

and interpretation of cytometry data," *Cytometry Part A*, vol. 87, no. 7, pp. 636–645, 2015.

[13] J. H. Levine, E. F. Simonds, S. C. Bendall, K. L. Davis, D. A. El-ad, M. D. Tadmor, O. Litvin, H. G. Fienberg, A. Jager, E. R. Zunder, *et al.*, "Data-driven phenotypic dissection of aml reveals progenitor-like cells that correlate with prognosis," *Cell*, vol. 162, no. 1, pp. 184–197, 2015.

[14] E.-a. D. Amir, K. L. Davis, M. D. Tadmor, E. F. Simonds, J. H. Levine, S. C. Bendall, D. K. Shenfeld, S. Krishnaswamy, G. P. Nolan, and D. Pe'er, "visne enables visualization of high dimensional single-cell data and reveals phenotypic heterogeneity of leukemia," *Nature biotechnology*, vol. 31, no. 6, p. 545, 2013.

[15] L. v. d. Maaten and G. Hinton, "Visualizing data using t-sne," *Journal of machine learning research*, vol. 9, no. Nov, pp. 2579–2605, 2008.

[16] V. Unen, T. Höllt, N. Pezzotti, N. Li, M. J. Reinders, E. Eisemann, F. Koning, A. Vilanova, and B. P. Lelieveldt, "Visual analysis of mass cytometry data by hierarchical stochastic neighbour embedding reveals rare cell types," *Nature communications*, vol. 8, no. 1, p. 1740, 2017.

[17] C. Stolper, A. Perer, and D. Gotz, "Progressive visual analytics: User-driven visual exploration of in-progress analytics," *IEEE Transactions on Visualization and Computer Graphics*, vol. 20, no. 12, pp. 1653–1662, 2014.

[18] T. Mühlbacher, H. Piringer, S. Gratzl, M. Sedlmair, and M. Streit, "Opening the black box: Strategies for increased user involvement in existing algorithm implementations," *IEEE Transactions on Visualization and Computer Graphics*, vol. 20, no. 12, pp. 1643–1652, 2014.

[19] D. Schapiro, H. W. Jackson, S. Raghuraman, J. R. Fischer, V. R. T. Zanotelli, D. Schulz, C. Giesen, R. Catena, Z. Varga, and B. Bodenmiller, "histoCAT: analysis of cell phenotypes and interactions in multiplex image cytometry data," *Nat Methods*, vol. 14, no. 9, pp. 873–876, 2017.

[20] T. Ropinski, S. Oeltze, and B. Preim, "Survey of glyph-based visualization techniques for spatial multivariate medical data," *Computers & Graphics*, vol. 35, no. 2, pp. 392–401, 2011.

[21] R. Borgo, J. Kehrer, D. H. Chung, E. Maguire, R. S. Laramee, H. Hauser, M. Ward, and M. Chen, "Glyph-based visualization: Foundations, design guidelines, techniques and applications.," in *Eurographics (STARs)*, pp. 39–63, 2013.

[22] N. Gehlenborg, S. I. O'donoghue, N. S. Baliga, A. Goesmann, M. A. Hibbs, H. Kitano, O. Kohlbacher, H. Neuweger, R. Schneider, D. Tenenbaum, *et al.*, "Visualization of omics data for systems biology," *Nature methods*, vol. 7, no. 3s, p. S56, 2010.

[23] T. von Landesberger, M. Görner, R. Rehner, and T. Schreck, "A system for interactive visual analysis of large graphs using motifs in graph editing and aggregation.," in *VMV*, vol. 9, pp. 331–340, 2009.

[24] C. Dunne and B. Shneiderman, "Motif simplification: improving network visualization readability with fan, connector, and clique glyphs," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 3247–3256, ACM, 2013.

[25] J. H. Lee, E. R. Daugharthy, J. Scheiman, R. Kalhor, T. C. Ferrante, R. Terry, B. M. Turczyk, J. L. Yang, H. S. Lee, J. Aach, *et al.*, "Fluorescent in situ sequencing (fisseq) of rna for gene expression profiling in intact cells and tissues," *Nature protocols*, vol. 10, no. 3, p. 442, 2015.

[26] A. M. Femino, F. S. Fay, K. Fogarty, and R. H. Singer, "Visualization of single rna transcripts in situ," *Science*, vol. 280, no. 5363, pp. 585–590, 1998.

[27] C. Larsson, J. Koch, A. Nygren, G. Janssen, A. K. Raap, U. Landegren, and M. Nilsson, "In situ genotyping individual dna molecules by target-primed rolling-circle amplification of padlock probes," *Nature methods*, vol. 1, no. 3, p. 227, 2004.

[28] C. Sommer, C. Straehle, U. Koethe, and F. A. Hamprecht, "Ilastik: Interactive learning and segmentation toolkit," in *Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on*, pp. 230–233, IEEE, 2011.

[29] T. R. Jones, I. H. Kang, D. B. Wheeler, R. A. Lindquist, A. Papallo, D. M. Sabatini, P. Golland, and A. E. Carpenter, "Cellprofiler analyst: data exploration and analysis software for complex image-based screens," *BMC bioinformatics*, vol. 9, no. 1, p. 482, 2008.

[30] M. Brehmer and T. Munzner, "A multi-level typology of abstract visualization tasks," *IEEE Transactions on Visualization and Computer Graphics*, vol. 19, no. 12, pp. 2376–2385, 2013.

[31] M. Brehmer, M. Sedlmair, S. Ingram, and T. Munzner, "Visualizing dimensionally-reduced data: Interviews with analysts and a characterization of task sequences," in *Proceedings of the Fifth Workshop on Beyond Time and Errors: Novel Evaluation Methods for Visualization*, pp. 1–8, ACM, 2014.

[32] I. Buchwalow, V. Samoilova, W. Boecker, and M. Tiemann, "Non-specific binding of antibodies in immunohistochemistry: fallacies and facts," *Scientific reports*, vol. 1, no. 1, pp. 1–6, 2011.

[33] J. Mackinlay, "Automating the design of graphical presentations of relational information," *Acm Transactions On Graphics (Tog)*, vol. 5, no. 2, pp. 110–141, 1986.

[34] J. D. Hunter, "Matplotlib: A 2d graphics environment," *Computing In Science & Engineering*, vol. 9, no. 3, pp. 90–95, 2007.

[35] E. R. Tufte, *Envisioning information*. Graphics Press, 1990.

[36] G. C. Linderman and S. Steinerberger, "Clustering with t-sne, provably," *CoRR*, vol. abs/1706.02582, 2017.

[37] N. Pezzotti, B. P. Lelieveldt, L. van der Maaten, T. Höllt, E. Eisemann, and A. Vilanova, "Approximated and user steerable tsne for progressive visual analytics," *IEEE transactions on visualization and computer graphics*, vol. 23, no. 7, pp. 1739–1752, 2017.

[38] Y. Cheng, "Mean shift, mode seeking, and clustering," *IEEE transactions on pattern analysis and machine intelligence*, vol. 17, no. 8, pp. 790–799, 1995.

[39] K. Shekhar, P. Brodin, M. M. Davis, and A. K. Chakraborty, "Automatic classification of cellular expression by nonlinear stochastic embedding (ACCENSE)," *Proceedings of the National Academy of Sciences*, vol. 111, no. 1, pp. 202–207, 2014.

[40] M. Bostock, V. Ogievetsky, and J. Heer, "D^3 data-driven documents," *IEEE Transactions on Visualization & Computer Graphics*, no. 12, pp. 2301–2309, 2011.

[41] M. Ghoniem, J.-D. Fekete, and P. Castagliola, "On the readability of graphs using node-link and matrix-based representations: a controlled experiment and statistical analysis," *Information Visualization*, vol. 4, no. 2, pp. 114–135, 2005.

[42] I. Spence and S. Lewandowsky, "Displaying proportions and percentages," *Applied Cognitive Psychology*, vol. 5, no. 1, pp. 61–77, 1991.

[43] R. Gove and B. Herzog, "Visualizing uncertain critical paths in schedule management." Poster at the IEEE Conference on Visualization (VIS), 2013.

[44] J. W. Hooton, "Randomization tests: statistics for experimenters," *Computer methods and programs in biomedicine*, vol. 35, no. 1, pp. 43–51, 1991.

[45] A. Somarakis and T. Höllt, "ImaCytE open source software." https://www.doi.org/10.5281/zenodo.3345951, 2019. [Online; Accessed: 2021-07-14].

4

Cohort Comparison

This chapter was adapted from:

A. Somarakis, M. E. Ijsselsteijn, S. J. Luk, B. Kenkhuis, N. F. C. C. de Miranda, B. P. F. Lelieveldt, and T. Höllt, "Visual cohort comparison for spatial single-cell omics-data," *IEEE Transactions on Visualization and Computer Graphics*, vol. 27, no. 2, pp. 733–743, 2021.

4.1. ABSTRACT

Spatially-resolved omics-data enable researchers to precisely distinguish cell types in tissue and explore their spatial interactions, enabling deep understanding of tissue functionality. To understand what causes or deteriorates a disease and identify related biomarkers, clinical researchers regularly perform large-scale cohort studies, requiring the comparison of such data at cellular level. In such studies, with little *a-priori* knowledge of what to expect in the data, explorative data analysis is a necessity. Here, we present an interactive visual analysis workflow for the comparison of cohorts of spatially-resolved omics-data. Our workflow allows the comparative analysis of two cohorts based on multiple levels-of-detail, from simple abundance of contained cell types over complex co-localization patterns to individual comparison of complete tissue images. As a result, the workflow enables the identification of cohort-differentiating features, as well as outlier samples at any stage of the workflow. During the development of the workflow, we continuously consulted with domain experts. To show the effectiveness of the workflow, we conducted multiple case studies with domain experts from different application areas and with different data modalities.

4.2. INTRODUCTION

Omics-data describe biochemical properties, such as genomics, transcriptomics, proteomics, or metabolomics of biological systems [1], such as cells. In recent years, high-resolution spatial measurements of such systems have become available. State of the art spatially-resolved omics modalities [2–6] enable the precise characterization of cellular populations in tissue, enabling the discovery and identification of novel cell types[7] in large cohorts of samples. Information about the cell type, in combination with the specific location of each cell creates many heterogeneous multi-cellular patterns.

With the identification of these multi-cellular patterns, a crucial question arises; are such patterns correlated with clinical information, such as survival rate? Current research findings [8–10] support the clinical importance of analysing spatial multi-cellular interactions. Hence, the development of workflows for the systematic comparison of cohorts consisting of spatially-resolved omics-data with specific clinical characteristics is essential for the understanding of tissue functionality.

In the majority of life-science studies, the comparison of cohorts of samples is based on statistical comparison of predefined finite number of elements [11–14]. However, traditional statistical approaches, based on prior knowledge pose the risk of missing unexpected correlations and cannot capture the vast combinatorial space [15] of spatial configurations for all different cell types. Moreover, they depend on high quality input which often cannot be guaranteed with single-cell omics-data due to uncertainty in cell segmentation and cell type identification. Comparative visualization [16] can provide useful insights into the differentiating factors of two cohorts and enables the interactive, data-driven exploration of the vast combinatorial space while simultaneously investigating the biological relevance and plausibility of findings with regard to the preprocessing.

Here, we extended our previous work focused on the identification and exploration of multi-cellular spatial interactions in single-cell omics-data [17] to enable interactive comparison of cohorts of such data. The main goals are to identify the characteristics that

differentiate a cohort, explore the cohorts' heterogeneity and relate these characteristics directly to the tissue. In some cases, just the comparison of the cell types abundance is adequate to differentiate cohorts. In other cases, a detailed comparison of contained cells and their specific neighborhoods, i.e. microenvironments is needed.

We propose an interactive, data-driven cohort comparison workflow. More specifically the main contributions of this paper are:

1. A workflow for the comparison of cohorts of spatially-resolved single-cell omics-data, specifically addressing the following tasks

   T1 compare cohorts based on the abundance of different cell types,

   T2 compare cohorts based on multi-cellular microenvironments,

   T3 detect outliers within each cohort, and

   T4 relate findings to their spatial position.

2. A protoype implementation of the described workflow

The remainder of this paper is structured as follows. We present related work in Section 4.3, followed by a brief description of target users, input data and tasks in Section 4.4. In Section 4.5, we describe the rationale behind our visual design and implementation in our prototype. We present a set of case studies and user feedback in Section 4.6. Finally, we discuss the limitations of our work and conclude in Section 4.7.

4.3. Related Work

The visual analytics community spent considerable effort on approaches for the exploration of cohorts of medical data combining spatial and non-spatial features. Preim et al. [18] provide an overview of image-centric approaches [19–21] focused on the exploration of large imaging cohorts and derived attributes. For the data analysis, these approaches share linking of attribute views with image views to provide context, visual queries for direct feedback, and interactive definition of groups of attributes. They typically deal with traditional medical imaging databases, such as those acquired by computed tomography (CT) or magnetic resonance imaging (MRI).

Dealing with microscopic images, Screenit [22] offers a system of linked views, similar to our system, to explore the drug screening results of cell cultures at multiple levels of detail. However, only recently, spatially-resolved omics-data [2–4] have become a standard tool for the exploration of tissue structure at the cellular level. Consequently, only few visual analysis tools exist that address the specific needs of such data. Facetto [23] is a scalable framework that allows hierarchical cell type identification in large multiplexed images. histoCAT [24] enables the identification of cell types and the significant pairwise spatial interactions between them. CytoMAP [25] offers an extensive toolbox for the exploration of tissue structure based on the analysis of spatial interactions. In our previous work on ImaCytE [17], we propose an interactive exploratory pipeline for cell type identification and neighborhood analysis in spatial single-cell data. Minerva [26] extends such exploration concepts with storytelling tools, to support communication and sharing of results. All of the above focus on the identification or exploration of cell types or significant multi-cellular

interactions in a single cohort of spatial single-cell data. Here, we use some of the concepts introduced in these works and extend them to introduce the first workflow for comparative analysis of two cohorts of such data, based on the abundance of cell types, as well as colocation patterns.

Based on a survey on existing comparative visualization tools [27], Gleicher et al. define a taxonomy that divides comparative visualization into juxtaposition (side-by-side placement), superposition (layering), and explicit encoding. A large body of work on comparative visualization for individual images exist. For example, Blaas et al. [19] combine superposition with explicit coding of the differences using complementary colors for the comparands, which cancels out in regions without differences. We use the same technique in some of our charts. Lindemann et al. [28], Maries et al. [29] and Ma et al. [30] utilize juxtaposition in an interactive comparative visualization pipeline for one-to-one comparison of segmentation results of brain imaging data. Juxtaposition for the comparison of images is also utilized in our work.

Schmidt et al. [31] facilitate the comparison of images with small differences within an ensemble. Raidou et al. [32] compare volume data and corresponding segmentations of bladders to explore the results of longitudinal radiotherapy treatment studies. Both works focus on all-to-all comparison of (3D) images in a single group, compared to the between-cohort comparison presented in this work. Basole et al. [33] as well as Wagner et al. [34] propose pipelines for the comparison of two cohorts. In their comparison workflow they use the same visual enocdings in order to compare the cohorts as a whole and simultaneously provide information for the intra-cohort heterogeneity, similar to the visual encodings we utilize in our system. Both approaches are limited to non-spatial healthcare data, though. Zhang et al. [35] present a visual analytics approach to compare two cohorts of diffusion tensor images. While we took some inspiration from their work, such as using complementary colors for the two cohorts that cancel each other out when overlapping, ultimately, the solutions described in their work are specific to tensor data and not easily transferrable to the spatial single-cell data described here.

4.4. ABSTRACTION

Recent developments in the spatially-resolved omics field manifest a wide variety of available modalities [3, 5, 36, 37]. These technologies measure transcriptomics or proteomics information at sub-cellular resolution, resulting in high-resolution image data with tens to thousands of values per pixel. Since researchers are interested in this information per cell, rather than per pixel, these images are typically pre-processed by segmenting individual cells and aggregating the values of the segmented pixels. Based on this aggregated information and potentially further features like morphology, the function and type of the segmented cells can be identified [24]. Both, cell segmentation [24, 38], as well as cell type identification [17, 23–25] in this kind of data is an active research topic. Large variations in cellular morphology and different quality of marker staining, among others, can lead to a considerable amount of uncertainty in the result of these preprocessing steps, making the validation, for example by referencing the actual images, during comparison imperative.

4.4.1. TARGET USERS AND GOALS

Our proposed workflow is targeted at clinical researchers who want to analyze their own data, for example to do an initial exploration of the data to form hypotheses. Typical goals when doing comparative analysis of two cohorts of spatial single-cell data could be the identification of cell types that are abundant in one cohort but not the other or cell co-localization patterns that are correlated with one of the cohorts. Such correlations or biomarkers [39] can be used for prognosis, monitoring or therapy of disease. While scripting in python or R is becoming more common in the domain, all our collaborators prefer visual exploration through GUI interfaces. Our proposed workflow is the first such visual exploration system that supports the comparative analysis of two cohorts of spatial single-cell data.

4.4.2. INPUT DATA

The overarching goal of our workflow is the comparison of two cohorts of spatially-resolved omics data as briefly introduced above. A single cohort consists of a set of samples, i.e., segmented and classified images as described above. Depending on the goal of the study, the samples consist of multiple images from a single subject or an arbitrary number of samples from multiple subjects. Typically, the two cohorts describe different populations, for instance, cancer patients who respond well to treatment in one cohort and those who respond worse in the second. A typical cohort consists of tens to hundreds of images, each consisting of thousands of segmented cells.

In a typical study, tens to hundreds of different cell types will be identified. The granularity depends on the goal of the study, as well as the data modality. For example, the Vectra imaging system [40] measures only a few different proteins (i.e. 4 in the case study in Section 4.6.3). Assuming differentiation into only low and high abundance, this results in an upper limit of 2^4 16 differentiable cell types. Other systems, such as Imaging Mass Cytometry, allow the measurement of up to 40 proteins, such that the number of cell types is limited rather by which types are of interest for the given study. A broad study would capture in the order of a hundred different cell types.

For each sample, we store the segmentation mask including a cell type label, i.e. class, for each segmented cell. Based on the cell segmentation mask, we derive the microenvironment for each cell. The microenvironment consists of the cell types and their abundance in the neighborhood of the given cell. We store the corresponding information per cell as a list of all cells that are contained in the microenvironment. The microenvironment of a cell varies according to the resolution of the modality and the type of sample. For example, in a tumor crowded with compact cells we would consider cells belonging to the microenvironment in a smaller distance, compared to brain tissue, where interacting cells can be further apart. Therefore, the distance defining the microenvironment of a cell is specified by the user. Typically, the microenvironment of a cell consists of no more than some tens of cells.

4.4.3. IDENTIFIED TASKS

In the following, we describe a set of tasks that we have identified in close collaboration with our domain expert partners from the pathology department at LUMC (co-authors of this manuscript). In general, we compare the two cohorts, based on the contained samples.

The first step of the workflow is comparing the cohorts according to the abundance of different cell types per sample (**T1**). This allows a simple differentiation of the cohorts based on the contained cells. In the second step, we further want to identify patterns in the cells' microenvironments that differentiate the cohorts. In **T2**, we compare cohorts based on multi-cellular microenvironments. Throughout the process we support visual detection of outliers within each cohort (**T3**), according to the abundance of contained cells and their microenvironments, and relate any findings to their spatial position (**T4**).

In the following, we describe and abstract **T1**-**T4** in more detail using Brehmer and Munzners task typology [41]. For references to this typology, we use a `mono-spaced` font.

T1 Cohort comparison based on the abundance of different cell types and combinations thereof in cohort samples. The relative abundance of a cell type in the samples forming a cohort and how much a specific subject deviates from the distribution within the cohort are important clinical biomarkers. As cell types can be of different granularity, it should also be possible to compare the cohorts, based on combinations of cell types. A trivial example is differentiating a cohort of cancer patients and a cohort of healthy subjects by comparing the abundance of tumor cells in the contained samples, where "tumor cells" can be a single cell type, or a combination of cell types according to a more fine grained definition. In this task **T1**, the user `compares` the two cohorts based on the abundance of different cell types within samples forming the cohort `discovering` and `locating` the cell type(s) that differentiate the two cohorts. The input for **T1** is the abundance of each cell type for each sample that we `summarize` as distributions over all samples in one cohort. The output is a list of cell types that differentiate the two cohorts.

T2 Cohort comparison based on multi-cellular microenvironments. The goal of **T2** is to compare the two cohorts according to the spatial co-localization patterns of each sample, as the comparison only based on cell type abundance is not enough to assess tissue functionality. Domain researchers hypothesize that tissue functionality also depends on the cell's interactions with other cells. While co-localization does not automatically lead to such interactions, it is a pre-condition. We facilitate the identification of such spatial features by breaking this task down into a high-level comparison, based on how often any two cell types are spatially co-located (**T2.a**), and a detail comparison where complex user-defined microenvironments can be explored (**T2.b**). In task **T2.a**, the user `discovers` combinations of two cell types that are most differentiating between the two cohorts. The input for this task is the abundance of each combination of two cell types in a microenvironment within the cohort sample. The output is a combination of two cell types to be used for further `exploration`. In task **T2.b**, the user further `explores` and `compares` the two cohorts based on more complex microenvironment compositions. Therefore, the user `produces` these more complex microenvironments by combining different cell types, typically starting with the combination found in **T2.a**. The input for **T2.b** is the complete set of cell microenvironments, optionally filtered to those including the combination of interest `discovered` in **T2.a**. The output is a set of detailed microenvironments differentiating the two cohorts.

T3 **Outlier detection within each cohort.** Detecting outliers within a cohort can provide additional important clinical information. For example subjects with different stages of a disease in the same cohort might exhibit different cell profiles [42]. Therefore, **T3** consists of `identifying` and `locating` outlying samples and their corresponding features identified in **T1** and **T2**. The input to this task is the abundance of cells and their microenvironments, as `identified` in **T1** and **T2**. The output is a list of outlying samples.

T4 **Relate findings to their spatial position.** As described above, **T1**-**T3** can be carried out based on cell abundance and microenvironment descriptions per sample, without consulting the actual imaging data. However, to verify individual findings we inspect the cells and their neighborhoods in their tissue context. Therefore, **T4** relates any findings to their spatial position. The analyst `locates` the structure of interest in their spatial location and `identifies` issues that were not apparent in the abstract representation. The input to **T4** are the segmented images and a structure of interest found with **T1**-**T3**, the output is a verified or rejected finding from **T1**-**T3**.

4.5. WORKFLOW

We designed a workflow to support the four tasks, identified and described in Section 4.4.3 and implemented it in a multiple-linked-views system, shown in Figure 1. The system is divided in three main blocks, where the left (Figure 1a) and right (Figure 1c) blocks support **T1** and **T2**, respectively by comparing the cohorts based on their cell type abundance and spatial interactions. **T4** relies on the inspection of tissue samples and supports **T1**-**T3**.

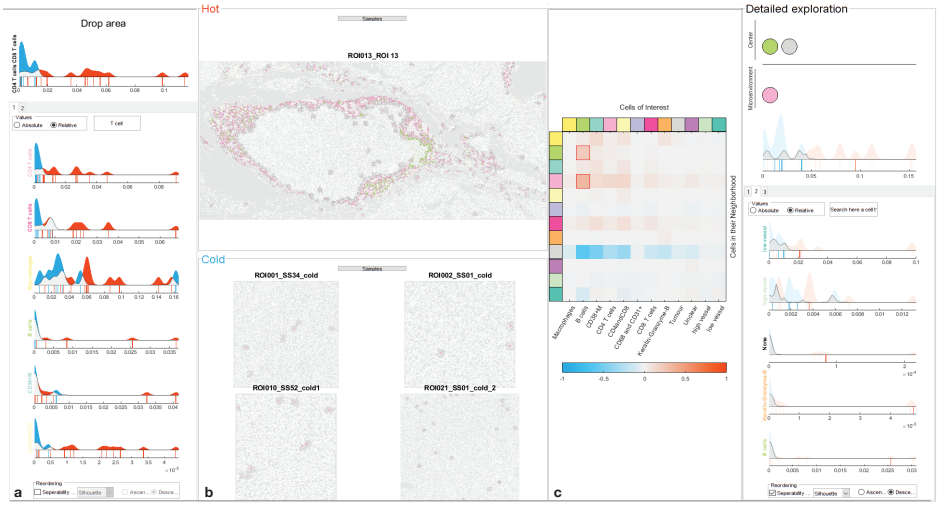

Figure 1: ***Screenshot of our integrated system*** *including the view for the comparison based on the cell abundance using raincloud plots (a), the tissue view, showing selected samples of the two cohorts (b), and the multi-cellular microenvironment comparison view using a difference heatmap and raincloud plots (c).*

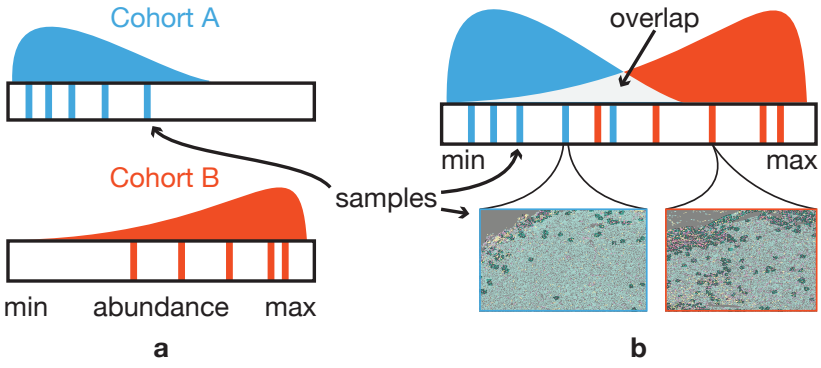

Figure 2: ***Comparison of two cohorts based on a cell type abundance.*** *(a) Individual raincloud plots for two cohorts showing the distribution (cloud) of samples (rain drops) according to the abundance of a contained cell type. (b) Superposition makes the difference visible by the large amount of color and small light-gray overlap area in the area chart.*

Therefore, we show the corresponding images between the views (Figure 1b) for **T1** and **T2** to support the user in directly making the connection for structures identified in any of the tasks to their spatial position. All views allow filtering the data to support visual outlier detection (**T3**).

4.5.1. Comparison Based on Cell Type Abundance

In the first step, we are interested to compare two cohorts according to the abundance of the different existing cell types in each of the contained samples (**T1**) and visually detect possible outliers in each of the cohorts (**T3**). Therefore, we first compute the number of cells of each type within each sample and then visualize the distribution of samples within both cohorts according to this value by superposing two simplified versions of raincloud plots [43]. This plot consists of a density (estimated using a kernel density estimate) plot showing the distribution of samples (the *cloud*) above a one-dimensional scatterplot with vertical lines as marks for the individual samples (*rain*-drops). This combination has proven very effective for our goals in **T1**-**T3**. The superposition of the density plots has shown to be very effective for the comparison of two distributions [44]. Both, the density plot [45] and the one-dimensional scatterplot [46], support visual detection of outliers. Furthermore, individual samples can be efficiently selected in the scatterplot for filtering. Additionally, for easier comparison between samples of different sizes, we enable the user to select whether the x-axis should represent the number of cells either as absolute values, or relative to the number of cells in that sample. As our primary goal is the comparison of the two cohorts, rather than the shape of individual plots, we want to emphasize the differences, rather than the commonalities. Therefore, following the same principle as Blaas et. al. [19], we use complementary colors for the two cohorts, i.e. blue and orange and blend the PDFs additively to receive a neutral light-gray in the overlapping areas as shown in Figure 2b.

The resulting raincloud plot allows the comparison of the composition of the two cohorts, according to the abundance of a single cell type within the contained samples. To allow the inspection of these distributions for all cell types, we use a small multiples approach [47, Chapter 4] and show the raincloud plots for several cell types in the same

view (Figure 3).

As indicated in Section 4.4.2, some studies can contain up to 100 different cell types. Finding a specific type of interest or the types that are the most differentiating for the two cohorts manually is not feasible in such a case. Therefore, we provide the possibility to sort the plots according to how well the corresponding distributions of the two cohorts separate, by default using the Silhouette metric [48], as it is invariant to the range of the input data. For advanced users we provide a set of other metrics, such as Dunn's index [49] which is efficient for compact and well separated clusters. In addition, we provide filtering by means of a textual search box (Figure 3a), based on the cell labels in the input data. Typing, for example *tumor* in this box will bring plots with the term *tumor* in their provided label to the top of the view (Figure 3b).

In some cases, the analyst might also be interested in aggregating the information on several cell types. For example, when several different cancer cell sub-types were identified in the original classification, but the analyst is only interested in how the cancer cells are distributed as a whole. To that end, we enabled the user to combine cell types, by gradually dragging and dropping the corresponding plots into a drop area on top of the

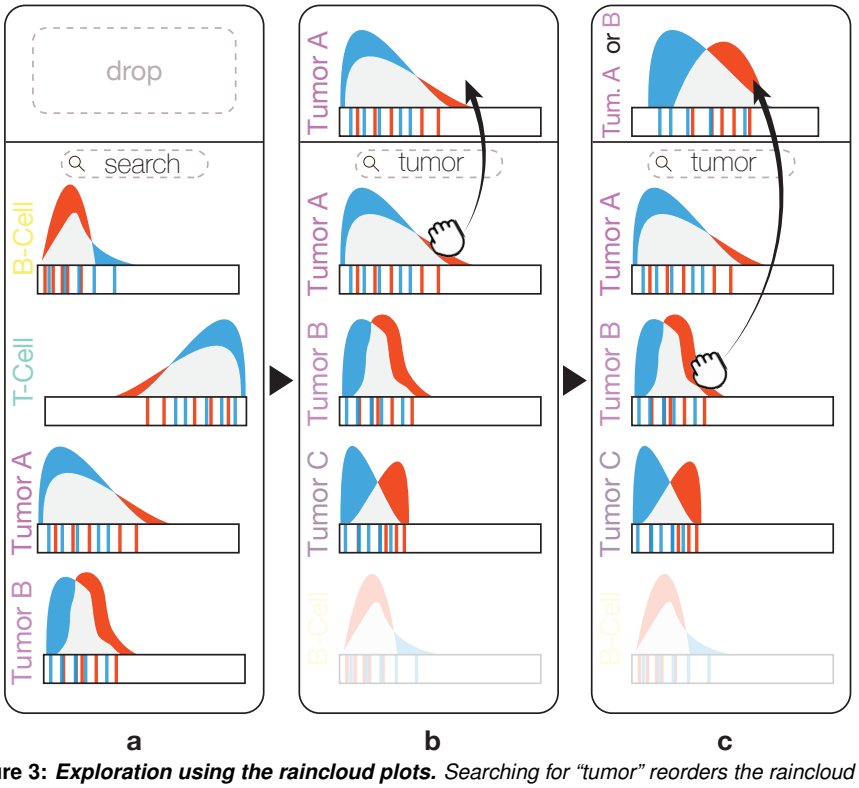

Figure 3: ***Exploration using the raincloud plots.*** *Searching for "tumor" reorders the raincloud plots by placing the plots corresponding to cell types containing the term 'tumor" in their label on top of the list (b). Dragging a raincloud plot and dropping it in the drop area (b,c), creates a new raincloud plot depicting the abundance of the cell types represented from the accumulated dropped raincloud plots.*

view (Figure 3b,c). The abundances of the dropped cell types are then aggregated as if they were a single cell type and a new distribution is created on-the-fly.

All views in our system are linked and allow cross-selection. For example, selection one or more lines in a raincloud plot filters the tissue view (Figure 1b) to show only the corresponding samples, with the cell type corresponding to the raincloud plot emphasized (**T4**). Further, these samples are also highlighted in the other raincloud plots, for example to verify whether a sample that is an outlier for one cell type also shows different behavior for other types (**T3**). To ensure that outliers in one cohort are not occluded by samples of the other cohort, the user can select to fade out one of the cohorts (**T3**).

4.5.2. COMPARISON BASED ON CELLULAR MICROENVIRONMENTS

The comparison of the cohorts based on their spatial interactions patterns, as indicated in task **T2**, is performed in two steps. The first step is to gain a global overview and compare the cohorts based on pairwise co-occurrences of cell types (**T2.a**). In the second step, the analyst can go into detail, explore and built specific, detailed microenvironments, consisting of an arbitrary number of cell types, and compare the distribution of these microenvironments among the two cohorts (**T2.b**). Throughout this process, we allow locating the identified microenvironments with the actual tissue images (**T4**) and in the second step, samples that are outliers in their cohort, according to the created microenvironment can be identified (**T3**).

4.5.2.1. PAIRWISE OVERVIEW

Following ImaCytE [17], we define the microenvironment of a cell, based on a user-defined distance as explained in Section 4.4.2.

We then compute the frequency for each cell type to occur in each other cell type's microenvironment throughout the cohort. For a detailed description we refer to our previous work [17, Section 4.3].

The result of this process is a directed and weighted graph, where each node represents a cell type and the link between two nodes defines the frequency of the target node appearing in the microenvironment of the source node. In ImaCytE, we visualize this frequency graph as a heatmap. Here, instead of showing the frequencies F, we compute the signed differences D in frequency between the two cohorts C_A and C_B. $D_t(C_A, C_B)$ $F(C_A) - F(C_B)$. We encode D using color based on the same heatmap layout, illustrated in Figure 4. The vertical axis shows the cell type of interest and the horizontal axis the cell types in the microenvironments. A large positive value indicates that the combination exists predominantly in Cohort A, while a large negative value means the combination predominantly exists in Cohort B. Based on this, we define a simple color map using the same colors previously assigned to the two cohorts and map the maximum absolute value $max(|D_t|)$ to the color assigned to Cohort A (i.e. blue) and $-max(|D_t|)$ to the color assigned to Cohort B (i.e. orange). Using the same concept of blending between the two colors, described in Section 4.5.1, the middle of this colormap, corresponding to D_t 0, will be a neutral light-grey, indicating both cohorts exhibit similar abundance of the given combination (compare Figure 4).

During one of the case studies (Section 4.6.1), it became clear that using the relative frequencies, used in ImaCytE [17] and the required normalization biased the heatmap

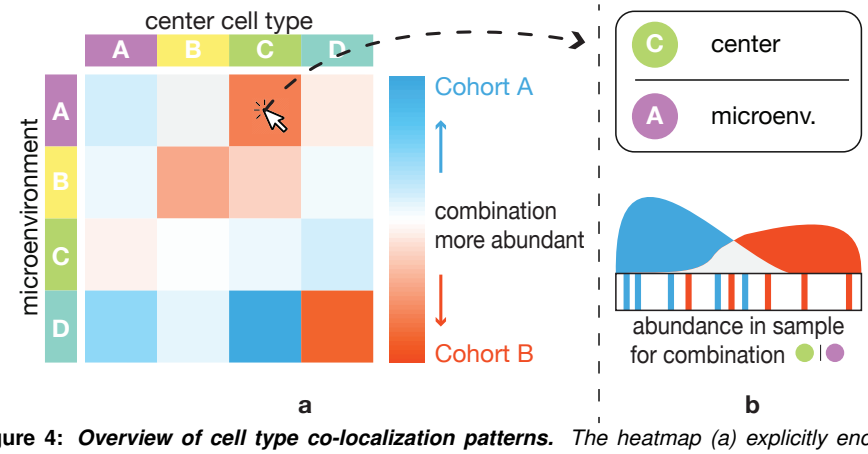

Figure 4: ***Overview of cell type co-localization patterns.*** *The heatmap (a) explicitly encodes differences in the abundance of pairwise combinations of cell types in the two cohorts. Clicking on one of the combinations sets this combination in the detail view (b), showing the distribution of samples according to the abundance of this combination.*

towards differences in small cell populations. To counter this issue, we provide the option to compute the heatmap using the separability metrics, also used for sorting the raincloud plots (Section 4.5.1). As these metrics only provide information on how different the cohorts are, we compute the mean abundance of the given cell type combination for all samples in a cohort and use the sign of the two cohort's difference in combination with the separability metric.

The resulting heatmap effectively shows cell type combinations that differentiate the two cohorts and for which cohort each combination is predominant. The analyst can now further explore individual combinations by clicking the corresponding box in the heatmap. Thereby, the corresponding combination is selected and highlighted in the tissue view (**T3**) and the microenvironment combination tool (Section 4.5.2.2) is pre-populated with the given combination (Figure 5a) for further analysis.

4.5.2.2. DETAIL MICROENVIRONMENTS

Starting with the overview of pairwise co-localization patterns, identified with the heatmap visualization, the analyst can now in detail explore complex microenvironment structures, based on any cell type combination and link those to individual samples along their position in the distribution of the corresponding cohort.

In ImaCytE [17], we used a simple glyph to enable the visual exploration of all the existing unique microenvironments in a sample. Here, the focus is on comparing two cohorts with regard to specific microenvironments, that potentially have already been identified as interesting in a previous analysis of the individual cohorts. Therefore, instead of showing all the existing unique microenvironments, the user can compare the two cohorts based on a specific pattern of spatial interactions. To enable the user to interactively define such a pattern, we utilize an interactive visual query system [50], similar to the one presented in Polaris [51] and further explained by Heer et al. [52]. The comparison of the two cohorts then happens with the same raincloud plots introduced in Section 4.5.1 but instead of the abundance of a single cell type the plot now displays the abundance of the queried microenvironment.

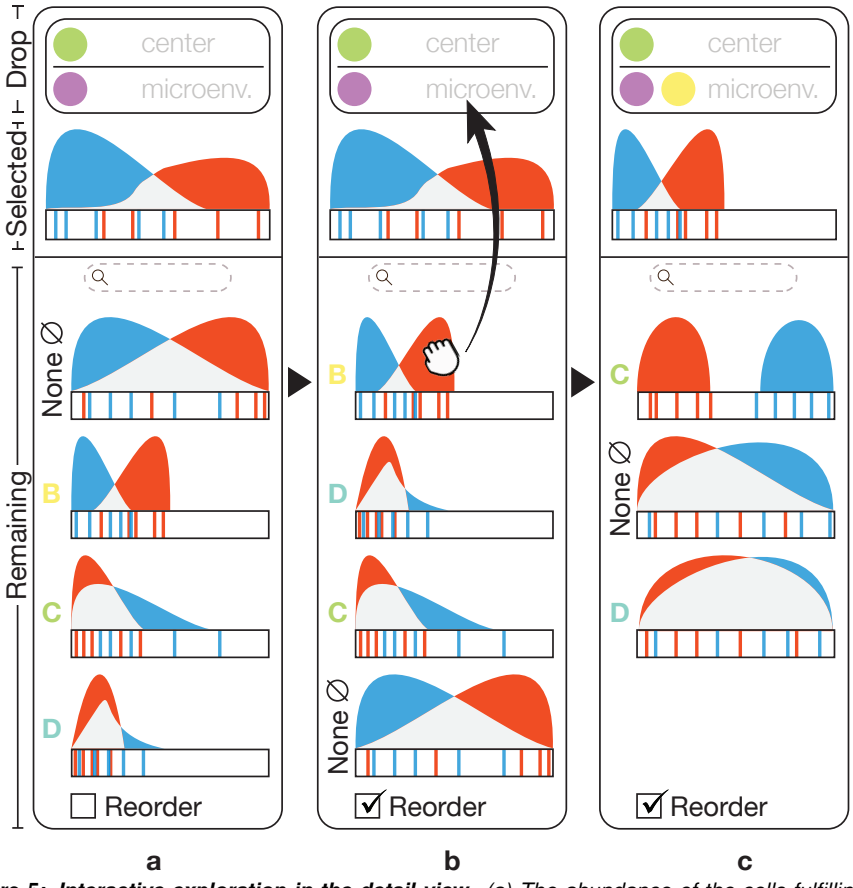

Figure 5: ***Interactive exploration in the detail view.*** *(a) The abundance of the cells fulfilling the cell type pattern in the* Drop *area is illustrated in the* Selected *raincloud plot. (b) The raincloud plots are reordered in the* Remaining *area according to their differentiating ability, the user drags the first raincloud plot and drops it in the* Drop *area. (c) The dropped raincloud plot replaces the previous one. Also, the* Drop *area and the* Remaining *plots are updated for further exploration.*

In practice, the analyst would typically start with a combination of two cell types picked from the heatmap. This simple microenvironment is illustrated on top of the detail view as illustrated in Figure 5a, where it is divided into the cell type of interest in the center of the microenvironment (i.e., cell type A, green circle, Figure 5a) and the microenvironment (i.e., cell type B, purple circle, Figure 5a). For the remainder of the paper we will denote microenvironments as ● | ●, where the circle(s) to the left of the vertical line represents the center cells combined with *or* type and the circle(s) to the right the microenvironment combined with *and*. I.e., a cell from either of the types left of the line must appear in the center and all the types to the right must appear in the surrounding of this cell. Below this (*Selected*, Figure 5a) we show the raincloud plot corresponding to the abundance of all microenvironments with at least the selected combination of cell types. Finally, further below (*Remaining*, Figure 5a) we depict the raincloud plots corresponding to the combination of the defined microenvironment plus any of the remaining cell types (here

●|●∅, ●|●●, ●|●●●, ●|●●●). The example in Figure 5a starts with *None* (indicated as ∅). At first glance, it might seem surprising that the corresponding raincloud plot is different from the initial plot above it. *None*, here means that no other additional cell type must exist in the microenvironment, whereas the initial plot shows all microenvironments that at least contain the given types. We denote this as ●|●∅. Below the *None* plot the remaining combinations are shown with the resulting raincloud plots. As described in Section 4.5.1, these plots can be ordered according to how strongly the corresponding microenvironment separates the two cohorts. Figure 5b illustrates the example after reordering. With this information the analyst can now continue exploring the microenvironments, for example by dragging the plot corresponding to cell type B (yellow) to the drop area, creating ●|●●, (Figure 5b). As the original plot already corresponded to the new microenvironment, we can now simply replace the "Selected" plot with the dragged plot (Figure 5c). The remaining raincloud plots (●|●●●, ●|●●∅, ●|●●●) are re-computed on-the-fly and shown below. Following this procedure, the user can progressively explore all interesting cell type combinations and evaluate their ability to discriminate the two cohorts and as such their potential as biomarkers.

As described in Section 4.5.1, the raincloud plots make it easy to identify samples that are outliers in their corresponding cohort (**T3**). Further, we provide the same linking and brushing features for selecting samples, as described in Section 4.5.1, to link the microenvironment patterns to the tissue view (**T4**).

4.5.3. TISSUE VIEW

In Section 4.4.2 we have described the importance of enabling the linking of any finding to its spatial location (**T4**). Therefore, we provide the tissue view (Figure 6), which shows the original segmented images and, linked to the other views, allows the inspection of selected cell types or microenvironments in the corresponding samples and their spatial context. The tissue view shows the images using color-coding for the different cell types. As we only consider the labeled segmentations as input (Section 4.4.2), we use a categorical colormap to assign a color to each label and thus cell type. We have chosen the qualitative *12 class Set 3* from colorbrewer [53] and have excluded blue and orange hues to avoid interference with the cohort colors. Colors are initially assigned based on the order of the cell type labels, but we allow the user to assign them manually by clicking on a cell type label. As typical studies have more cell types than the available ten colors, they can assign the same color to semantically grouped types. We then automatically adjust the saturation of hues that were selected multiple times to enable differentiation. While not described in detail in previous sections, this color scheme is used throughout the application to represent the different cell types and allow for easy mental linking between views. We have previously used a similar color scheme in ImaCytE [17]. To enable comparison between the cohorts, we divide the tissue view into two parts, one for each cohort. The name and color corresponding to the cohorts is shown on top of each view (Figure 6).

As described before, all views are linked. Therefore, the tissue view can be filtered to only show samples selected in other views. Further, selecting cell types or microenvironments in other views highlights them in the images by fading non-selected structures out, resulting in a light-grey for all unselected areas (Figure 6). Moreover, the tissue view supports zooming and panning across tissue samples to further assist the exploration of the

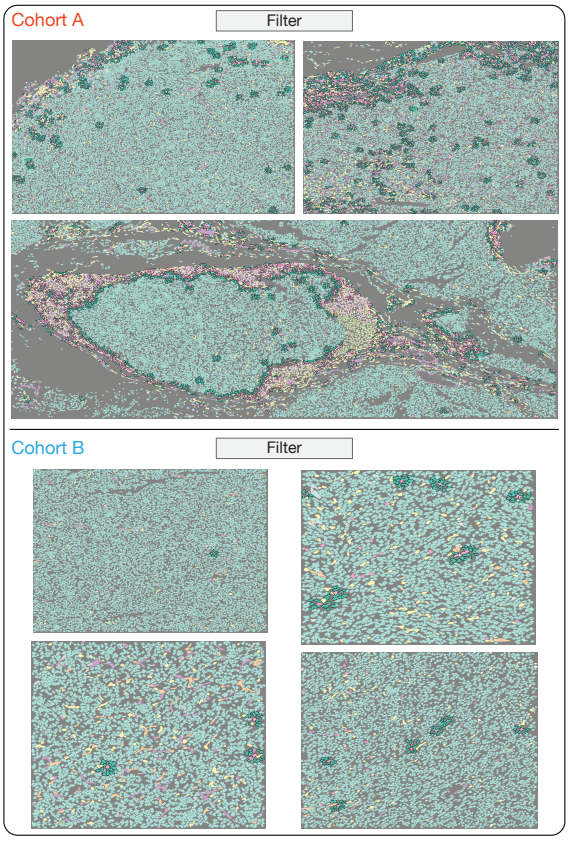

Figure 6: ***Tissue view,*** *highlighting a spatial interaction fading out the non-selected tissue structures. In the tissue samples of Cohort A, the spatial interactions form a compact structure, whereas the spatial interaction of Cohort B tissue samples are distributed all over the samples.*

(highlighted) tissue areas.

4.5.4. IMPLEMENTATION

As described in Section 4.4.1, our target users are clinical researchers with little programming experience. Therefore, we implemented the described workflow in a stand-alone GUI application. The application is implemented in MATLAB, as it allowed us to quickly build a stand-alone prototype. Source code and binaries are available on GitHub [54].

4.6. VALIDATION

In order to show the effectiveness of our workflow, we conducted three case studies with collaborators (P1-P3) at Leiden University Medical Center. P1 was also our main contact during the development of the workflow. After conducting the case studies and collecting feedback, we invited the collaborators to participate in the write up of the case-studies, and hence they are all co-authors of this manuscript. All collaborators acquired their own data with varying biological goals, using two different modalities as indicated in Table 4.1. For

Table 4.1: ***Summary of the case study characteristics.***

	Case Study	Modality	Samples in Cohort		Cell Types
			1	2	
P2	Sarcoma	IMC [3]	13	7	12
P1	Tumor	IMC [3]	19	28	60
P3	Alzheimer's	Vectra [40]	12	9	16

the case studies, we gave participants a hands-on introduction and answered any questions regarding the tool. After that, we observed the participants performing their analysis independently and reproduced their workflows for presentation in Sections. 4.6.1-4.6.3. As described in Section 4.5, for all the case studies the segmentation masks and the cell type identification had been performed as a pre-processing step by the participants. An overview of the study parameters with regard to imaging modality, numbers of samples, and numbers of included cell types is given in Table 4.1. As can be seen, the studies cover three different application areas, contain data from two different modalities, between 20 and 47 samples, and between 12 and 60 cell types. Finally, we asked the participants, as well as a fourth user of the software (P4, not a co-author of this manuscript), to fill out a short questionnaire (available in the supplemental material) via google forms [55]. The questionnaire consists of the ten standard System Usability Scale (SUS) statements [56], an additional nine statements specific to our tool, answered on a 5-point Likert scale, and five questions for open feedback. The individual plots presented in the case study have been exported directly from our tool and laid out with adjusted labels and annotations for the printout.

4.6.1. CASE STUDY I: SYNOVIAL SARCOMA (P2)

Synovial sarcoma is a rare form of cancer. During the immune response, T-cells infiltrate the sarcomas. Previous work has shown that synovial sarcomas can have areas with abundant T-cell infiltration (*hot areas*) and areas with very little T-cell infiltration (*cold areas*), in the same tumor[57]. The goal of this case study was to explore differences in the immune cell composition between these two types of areas. A total of 20 areas from 7 different tumors were imaged, of which 7 were cold (*Cold Cohort*, blue) and 13 were hot (*Hot Cohort*, orange). The size of the samples varied, with the number of cells in each image ranging from $2,678$ to $23,774$ cells. In the pre-processing step, cells were segmented and 12 different cell types were identified, based on the original data. While the number of cell types is relatively low, they cover a large range of available types, with rather coarse specificity.

4.6.1.1. CELL TYPE ABUNDANCE

In the first step of the analysis the expert was mostly interested in identifying cell type(s) that differentiate the cohorts, matching **T1** of our task analysis. Given the large variation in the number of cells per sample, he used the relative cell type abundance for comparison. First, he wanted to explore the uniformity of each cohort. As indicated above, the samples

were sorted into the two cohorts based on the infiltration of *T-cells* in the contained tumor tissue. Consequently the T-cells should exist predominantly in the Hot Cohort. As a first step, the expert wanted to verify this using the system. As there are two different types of T-cells in the dataset (*CD4* ● and *CD8 T-cells* ●) he first queried for these two cell types and created a combined raincloud plot by dragging the CD4 T-cell and CD8 T-cell plots to the combined drop area (subsection 4.5.1). The resulting combined plot (Figure 7a) confirmed that T-cells were largely non-existent in all seven samples of the Cold Cohort (blue peak close to 0, Figure 7a) but more widely distributed in the Hot Cohort (even spread of the orange distribution, Figure 7a). After navigating among the plots, he discovered the raincloud plot for *B-cells* ● (Figure 7b). This plot caught the expert's interest. Even though most samples from both cohorts hardly contain any B-cells, there are a few samples in the Hot Cohort that contain some B-cells, indicated by the orange lines to the right of the plot in Figure 7b. Given the generally low values, approximately 3 percent, even for the sample with the largest abundance, the expert decided to not further investigate these samples at this point and proceeded with other cell types. Therefore, he ordered the raincloud plots according to the Dunn's index [49]. The first plot illustrating *macrophages* ● showed a pattern similar to the T-cells (Figure 7c). Strikingly, there is an outlier (**T3**) clearly visible in the plot (highlight in Figure 7c). The corresponding sample from the Cold Cohort consists of over 16% macrophages, compared to no more than 5% for all other samples of the same cohort. Selecting the corresponding line in the plot also revealed that this sample has the highest abundance of T-cells in this cohort (though only at around 1% of cells in this sample).

At this point, the expert was curious whether the microenvironments of the macrophages

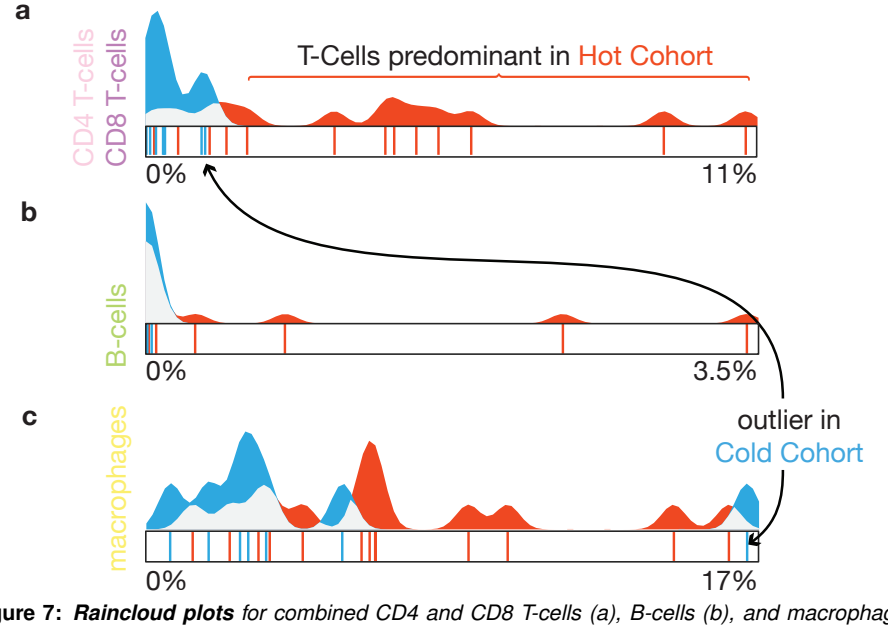

Figure 7: ***Raincloud plots*** *for combined CD4 and CD8 T-cells (a), B-cells (b), and macrophages (c). An outlier for macrophages in the cold cohort is clearly visible in (c). Selecting it showed it also contained slightly more T-cells than other samples in the cold cohort (a).*

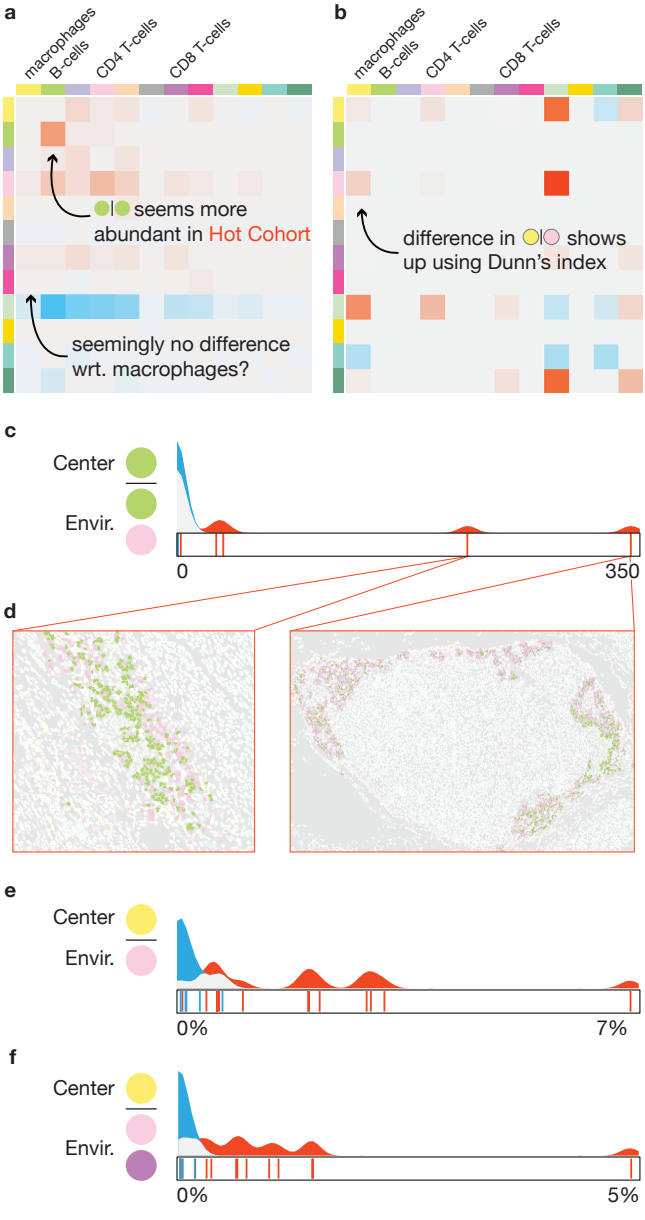

Figure 8: ***Multi-cellular microenvironment cohort comparison.*** *(a) A heatmap depicting the difference of the amount of pairwise spatial interaction between two cohorts normalized according to the abundance of each cell type. (b) A heatmap depicting the Dunn index for the samples of each cohort for each pairwise co-localization pattern. (c) The amount of B-cells having in their microenvironment B-cells and CD4 T-cells, depicting that the occured differentiation in (a) was due to the two outlier samples, which exist in a tertiary lymphoid structure, an interesting biological structure (d). The amount of macrophages having in their microenvironment CD4 T-cells (e) and CD8 T-cells (f).*

and B-cells could provide further clues on differentiating factors between and within the cohorts.

4.6.1.2. MICORENVIRONMENTS

The exploration of the differences between the two cohorts, with regard to the contained microenvironments (**T2**) starts with the overview provided by the difference heatmap (Figure 8a). The difference heatmap (Figure 8a) indicated that combinations of B-cells and B-cells ● | ● and B-cells and T-cells ● | ● were more prevalent in the Hot Cohort (highlighted orange boxes). With this information, the expert created the combined mircoenvironment ● | ● ● using the drag and drop interface. The corresponding raincloud plot showed two clear outliers in the Hot Cohort showing a larger abundance of this combination (Figure 8c). Using the linked tissue view, the expert could highlight the microenvironments in the corresponding samples (Figure 8d). The expert observed that the highlighted microenvironments were mostly present in so-called tertiary lymphoid structures [57]. While not directly relevant for the cohort comparison, he noted the two outlier samples for later detailed inspection in his standard workflow.

In the previous step, the expert had also identified macrophages ● for further exploration. Curiously, the heatmap did not show any strong differences between the two cohorts with regard to the microenvironments of this cell type. After the case study, we analyzed the data and came to the conclusion that the normalization applied to create the heatmap (subsubsection 4.5.2.1) strongly biased the heatmap in favor of small cell populations such as the B-cells in this study (subsubsection 4.6.1.1). As a result, we added the option to use the same cluster separation metrics used for sorting the raincloud plots according to their power to separate the cohorts for the heatmap as described in subsubsection 4.5.2.1. Figure 8b shows the heatmap using the Dunn's index as an example. Here, the ● | ● microenvironment is more clearly visible, while the small values of the B-Cell microenvironments are suppressed. The expert selected the corresponding box from the heatmap and examined the distribution of the samples for each cohort in the detail view. The blue area around zero (Figure 8e) indicated the absence of ● | ● microenvironment in the Cold Cohort, verifying the heatmap findings. Then, the expert having already identified the correlation among CD8 T-cells and macrophages navigated among the plots of the "Remaining" area of the detail view and located the CD8 T-cell raincloud plot. The addition of CD8 T-cells in the microenvironment of macrophages ● | ● ● further differentiated the two cohorts, shown by the restriction of the blue area to almost zero (Figure 8f). Even the strong outlier in the Cold Cohort that contained the largest amount macrophages of all samples did not show any significant co-localization of macrophages and T-cells. On the other hand, several samples in the Hot Cohort showed significant amounts of both combinations. Therefore, the expert concluded that both T-cell sub-types seems to better differentiate the hot and cold tumor areas, than their one-to-one spatial interaction or even their abundances.

4.6.2. CASE STUDY II: TUMOR METASTASIS (P1)

In this case study, the expert wanted to explore the differences in the cellular microenvironments of tumors with different clinical characteristics. In particular, she had acquired a data set, consisting of a total of 47 images taken from different tumor samples. Based on other clinical parameters she divided the set in two cohorts. The first one contains

19 images of non-metastatic tumors (*Non-Metastatic Cohort*, orange), the second 28 images of metastatic tumors (*Metastatic Cohort*, blue). She had segmented the images in a pre-processing step and identified 60 different cell types, among a total of 393,727 cells.

4.6.2.1. CELL TYPE ABUNDANCE

First, the expert was interested to discover cell type(s) which exist predominantly in one of the cohorts. Given the large amount of cell types, she ordered the raincloud plots according to the Silhouette metric in descending order, to assist her exploration. The first few plots consisted mostly of different subsets of T-cells, which had been defined in great detail in the preprocessing step. All of the corresponding plots showed a similar pattern of very small abundances for the Metastatic Cohort, indicated by a large blue peak to the left of the plot and a varying, but generally larger abundance in the Non-Metastatic Cohort. Searching for all cell types containing "T-cell" in their label showed a similar pattern for all of the remaining types (Figure 9a). This pattern is not completely surprising, as T-cells are a major factor in the immune response to cancer. For further exploration, in particular the relation of the identified T-cells to cancer cells, the expert aggregated all T-cell subsets using the drag and drop interface. The resulting raincloud plot (Figure 9b) confirmed that the T-cells clearly differentiate the two cohorts. There were, however, three samples from the Metastatic Cohort visible (blue lines, labeled A,B,C in Figure 9b) that showed a somewhat increased abundance compared to the remaining samples in that cohort. Next, the expert was interested, whether the increased amount of T-cells in the Non-Metastatic Cohort would correlate to differences in contained tumor cells. The expert searched for "tumor", to bring up the raincloud plot, corresponding to *Proliferating Tumor Cells*. However, as shown in Figure 9c, no clear separation between the two cohorts can be made, based on these cells. Finally, selecting the three outliers samples (A,B,C) in the T-cell plot did not show a specific differentiation with regard to the tumor cells.

4.6.2.2. MICORENVIRONMENTS

The last findings of subsubsection 4.6.2.1 intrigued the interest of the expert to further explore whether the tumor cells are present in the same amounts also in the microenvironment of T-cells. She quickly combined T-cells and proliferating tumor cells to a microenvironment | to bring up the corresponding raincloud plot (Figure 9d) in the detail view. The plot shows a clear differentiation among the two cohorts. In fact, this combination differentiates the two cohorts even stronger than only the T-cells. Even for the samples (Samples A,B,C) that showed increased abundance in T-cells, compared to the rest of the Metastatic Cohort, there was only a very small abundance of the | microenvironment. This strongly indicates that tumor cells exist in the microenvironment of T-cells in the Non-Metastatic Cohort, whereas in the Metastatic Cohort there is no spatial interaction between tumor and T-cells regardless their abundance. This lead the expert to hypothesize that the co-localization between the tumor and T-cells needs to be taken into account in tumor analysis, rather than the abundance of T-cells alone.

4.6.3. CASE STUDY III: ALZHEIMER'S DISEASE (P3)

The accumulation of *amyloid plaques* in the brain is an important characteristic of Alzheimer disease. These amyloid plaques are infiltrated by microglial cells, the resident immune cells of the brain. In this final case study, the expert wanted to verify the

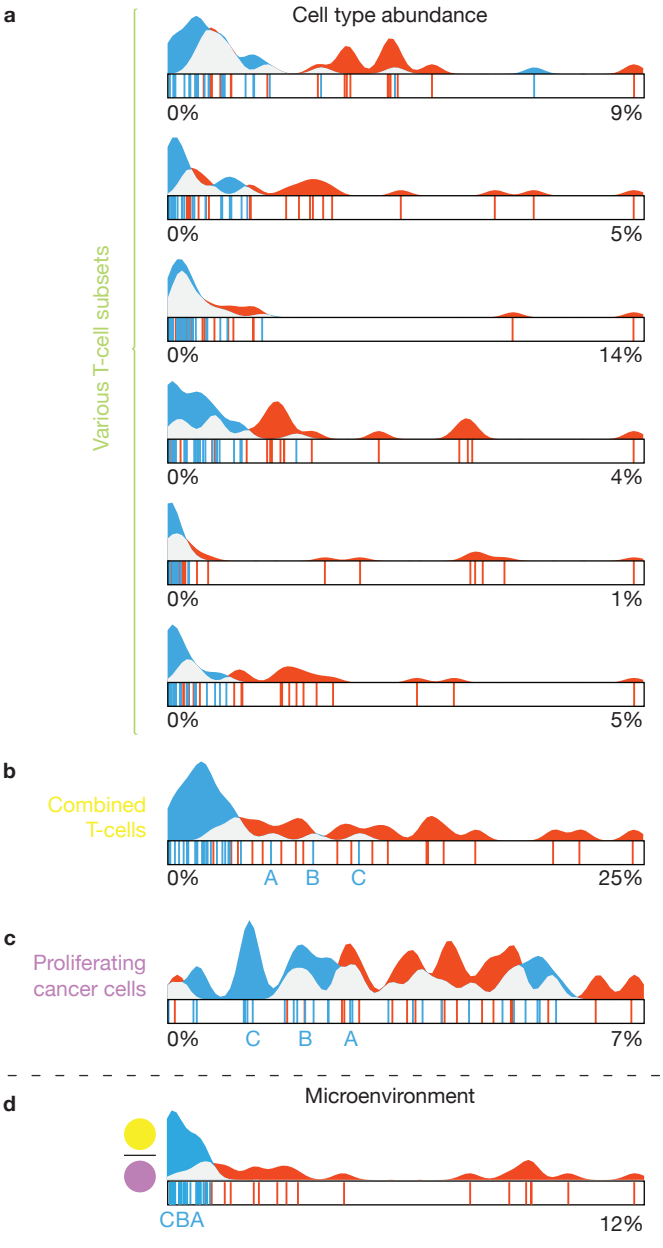

Figure 9: ***Raincloud plots*** *for various T-cell subsets (a), the aggregated plot combined from those subsets (b), and proliferating cancer cells (c). (d) shows the amount of the aggregated T-cells with proliferating cancer cells in their microenvironment. Even though the samples A-C, of the Metastatic Cohort, had a significant amount of T-cells and proliferating cancer cells (b,c) they did not spatially interact (d).*

hypothesis that the microglia cells close to and potentially attacking amyloid plaques are different from the microglia cells in healthy individuals.

The data used in this case study are somewhat different from the first two cases. The number of samples is comparable. Here, each sample represents one subject, for a total of 12 patients in the Alzheimer's Cohort (orange) and 9 healthy subjects in the Control Cohort (blue). However, each subject is described by up to 150 images, acquired with the Vectra 3.0 [40] machinery. 16 different cell types were identified and segmented in the pre-processing step. The identified cell types consist mostly of different subsets of microglia cells and as a result, the segmentation of the images is rather sparse, containing only in the order of 25 cells per image, plus the separately segmented amyloid plaques. As such, the individual images were not as important in this study as in the previous two and the data set only contained aggregated information of cell type abundance and microenvironments for all images per subject.

4.6.3.1. DATA ANALYSIS

As the experts goal was to verify a specific hypothesis, the data analysis in this study was much more targeted, compared to the rather explorative nature of the previous case studies. First, he brought up the raincloud plots corresponding to two microglia subtypes with contradictory patterns (Figure 10a,b). As can be seen in the plots *Subtype 1* ● was prevalent in the Control Cohort (blue), whereas Subtype 2 ● was mostly found in samples of the Alzheimer's Cohort (orange) but there was still some overlap between the samples from the two cohorts. This differentiation was already an indicator to verify the original hypothesis of the expert. Going back to the original data, the expert noted that the microglia Subtype 2

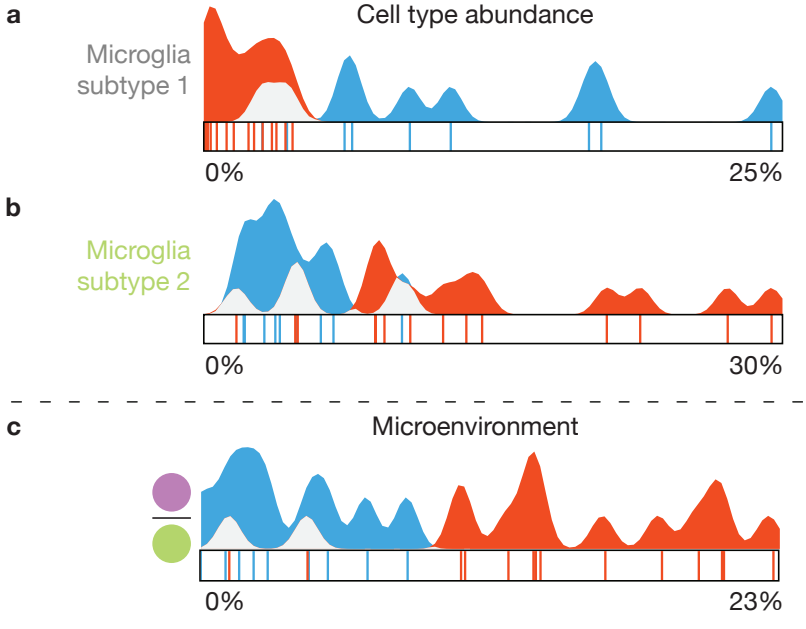

Figure 10: ***Raincloud plots*** *for microglia subtypes (a,b), and amoyloid plaques with microglia subtype 2 in their microenvironment (c).*

did not express two proteins that were expressed by Subtype 1 and hypothesized that these proteins might be suppressed when in the vicinity of the amyloid plaques ● in Alzheimer's disease patients. Consequently, he brought up the raincloud plot of the corresponding microenvironment ●|● (Figure 10c). Here, the distinction between the two cohorts is even clearer, with only two samples from the Alzheimer's Cohort in the range of the Control Cohort. The distribution further indicates that Subtype 2 seems to co-localize with amyloid plaques, supporting the generated hypothesis.

4.6.4. FEEDBACK

After the case studies, we collected feedback from the participants using a short questionnaire (available in the supplemental material) via google forms [55]. The questionnaire consists of the ten standard System Usability Scale (SUS) statements [56] (Q1–Q10), an additional nine statements specific to our tool (Q11–Q19), answered on a 5-point Likert scale, and five questions for open feedback. After the case studies, a fourth collaborator started working with the tool. After she got acquainted with it, we asked her to fill out the same questionnaire.

The average SUS-score, based on all four questionnaires was 76.25 with a standard deviation of 3.23 resulting in a *good* rating [58]. In the following we briefly summarize the feedback of the custom block of the questionnaire (Q11–Q19), for the complete set of responses we refer to the supplemental material. An overview of the responses is provided in Table 4.2. The custom part of the questionnaire is divided into three blocks. The first block (Q11–Q14) corresponds to the identified tasks (subsection 4.4.3). The second block (Q15–Q18) targets the interaction with the raincloud-based views in the cell abundance and microenvironment exploration. Finally, in the third block, we ask about general feedback.

With statements Q11–Q14 we queried whether **T1**–**T4** (subsection 4.4.3) could be carried out efficiently. (Q11; *The tool allows me to efficiently compare two cohorts, according to the abundance of contained cell types per sample* relates to **T1**, Q12 to **T2**, and so on). Generally, responses were clearly positive with strongly agree (++) or agree (+) with the exception of a neutral (○) response to Q11 and Q12, each. From the open feedback (Q20: *What functionality was missing to fully accomplish all goals?*) we could gather that participants would like to be able to "*correct[ion] cell abundance*" with regard to the amount of cells from user-defined area. Further, "*statistical testing of differences found between cohorts*" was requested, related to **T1** and **T2**.

In Q15–Q18 we were interested whether the raincloud plots were helpful to compare the distributions (Q15, **T1-T2**) and to find outliers (Q16, **T3**) as well as whether the drag and drop interaction made it easy to combine cell types (Q17) and build microenvironments (Q18). Q15–Q17 were overwhelmingly positive, with Q18 getting neutral responses by majority. The different response to Q17 and Q18 is rather unclear to us, as the interaction for combining cell types and building the detailed microenvironments is essentially the same. Unfortunately, there is also no further feedback on this in the open part of the questionnaire.

In the open feedback we can see that Participant 3 was missing "*Within subject distribution of cell types/clusters.*" As described in Figure 4.6.3, we had aggregated the very large amount of images in this study to a single dataset per subject. It might be interesting to provide a hierarchical approach in the future, that allows drilling into these subjects.

Table 4.2: ***Summary of participants' answers*** *to statements of our questionnaire on a 5-point Likert scale from very positive (++) to negative (-). No very negative (–) responses were given.*

Q	11	12	13	14	15	16	17	18	19
++	●●		●●	●	●●	●	●		
+	●	●●●	●●	●●●	●●	●●	●●●	●	●●
○	●	●				●		●●●	
-									●●

Participant 4 mentioned "*the option to compare 3 cohorts*" as a missing feature in the open feedback. While we focus on the comparison between two cohorts this is a possible future extension.

Finally, in the open feedback the "*possibility to detect outliers (and directly identify the subject*" (**T3**) was specifically mentioned as a positive aspect. The link between the abstract views and the actual images (**T4**) was highlighted by one participant: "*The rainbowplots are really cool, especially because you can go up and down to the images again.*" Particularly positive was a comment by Participant 1, that "*with the tool I already discovered a very nice thing in my existing data!*".

4.7. Discussion and Conclusion

We presented a workflow for the interactive visual comparison of two cohorts comprising single-cell omics-data, based on the cell abundance and their cell microenvironments.

The presented case studies contained up to 47 samples and up to nearly 400.000 cells. Our sorting and filtering options allow effective exploration of datasets of such sizes, however, increasing numbers to hundreds of samples will pose new challenges. In the Alzheimer's disease case study we accommodated a much larger original dataset (3286 images) by aggregating the information per patient and imaged region to a single larger image, resulting in the dataset described in Figure 4.6.3. Extending this to a hierarchical approach, facilitating the exploration of such aggregated regions and then individual images within a region might be a worthwhile extension.

At this point, our workflow is focused on two-dimensional images, as our partners currently only acquire such data. However, image stacks or volumetric measurements are becoming more readily available. Assuming a three-dimensional definition of microenvironments, the views based on abstract information, such as the raincloud plots and heatmap, would readily adapt to such data. Extensions to the spatial view, for example by volume rendering, would be necessary to inspect findings in the tissue context.

We have implemented the drag and drop interface to create simple center-neighborhood microenvironments. Nevertheless, the approach would support more advanced microenvironments through more drop targets, intuitively. For example, the neighborhood could be divided into multiple segments to allow a microenvrionment definition that has cell type A to the left and cell type B to the right of the center cell. A more traditional user interface, such as checkboxes, to assign cell types to each of those segments would be less flexible and quickly require a large amount of additional user interface elements.

Our workflow is designed to compare two clearly defined separate cohorts such as control vs. disease. Extending it to support more cohorts, or including more continuous features such as age or trial dose are open questions that certainly warrant future research.

ACKNOWLEDGEMENTS

This work received funding through Leiden University Data Science Research Programme. B.P.F.Lelieveldt received partial funding from H2020-Marie Skodowska-Curie Action Research and Innovation Staff Exchange (RISE) Grant 644373-PRISAR. N.F.C.C.de Miranda has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No. 852832)

REFERENCES

[1] A. Conesa and S. Beck, "Making multi-omics data accessible to researchers," *Scientific data*, vol. 6, no. 1, pp. 1–4, 2019.

[2] R. Ke, M. Mignardi, A. Pacureanu, J. Svedlund, J. Botling, C. Wählby, and M. Nilsson, "In situ sequencing for rna analysis in preserved tissue and cells," *Nature Methods*, vol. 10, no. 9, pp. 857–860, 2013.

[3] C. Giesen, H. A. Wang, D. Schapiro, N. Zivanovic, A. Jacobs, B. Hattendorf, P. J. Schüffler, D. Grolimund, J. M. Buhmann, S. Brandt, Z. Varga, P. J. Wild, D. Günther, and B. Bodenmiller, "Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry," *Nature Methods*, vol. 11, no. 4, pp. 417–422, 2014.

[4] N. Crosetto, M. Bienko, and A. Van Oudenaarden, "Spatially resolved transcriptomics and beyond," *Nature Reviews Genetics*, vol. 16, no. 1, pp. 57–66, 2015.

[5] J. H. Lee, E. R. Daugharthy, J. Scheiman, R. Kalhor, T. C. Ferrante, R. Terry, B. M. Turczyk, J. L. Yang, H. S. Lee, J. Aach, K. Zhang, and G. M. Church, "Fluorescent in situ sequencing (FISSEQ) of rna for gene expression profiling in intact cells and tissues," *Nature Protocols*, vol. 10, no. 3, pp. 442–458, 2015.

[6] Y. Goltsev, N. Samusik, J. Kennedy-Darling, S. Bhate, M. Hale, G. Vazquez, S. Black, and G. P. Nolan, "Deep profiling of mouse splenic architecture with codex multiplexed imaging," *Cell*, vol. 174, no. 4, pp. 968–981, 2018.

[7] V. Van Unen, T. Höllt, N. Pezzotti, N. Li, M. J. Reinders, E. Eisemann, F. Koning, A. Vilanova, and B. P. Lelieveldt, "Visual analysis of mass cytometry data by hierarchical stochastic neighbour embedding reveals rare cell types," *Nature Communications*, vol. 8, no. 1, pp. 1–10, 2017.

[8] H. R. Ali, H. W. Jackson, V. R. T. Zanotelli, E. Danenberg, J. R. Fischer, H. Bardwell, E. Provenzano, O. M. Rueda, S.-F. Chin, S. Aparicio, C. Caldas, and B. Bodenmiller, "Imaging mass cytometry and multiplatform genomics define the phenogenomic landscape of breast cancer," *Nature Cancer*, vol. 1, no. 2, pp. 163–175, 2020.

[9] L. Keren and M. Angelo, "Mapping cell phenotypes in breast cancer," *Nature Cancer*, vol. 1, no. 2, pp. 156–157, 2020.

[10] H. W. Jackson, J. R. Fischer, V. R. Zanotelli, H. R. Ali, R. Mechera, S. D. Soysal, H. Moch, S. Muenst, Z. Varga, W. P. Weber, and B. Bodenmiller, "The single-cell pathology landscape of breast cancer," *Nature*, vol. 578, no. 7796, pp. 615–620, 2020.

[11] C. J. Newschaffer, K. Otani, M. K. McDonald, and L. T. Penberthy, "Causes of death in elderly prostate cancer patients and in a comparison nonprostate cancer cohort," *Journal of the National Cancer Institute*, vol. 92, no. 8, pp. 613–621, 2000.

[12] Y. Yuan, H. Failmezger, O. M. Rueda, H. Raza Ali, S. Gräf, S. F. Chin, R. F. Schwarz, C. Curtis, M. J. Dunning, H. Bardwell, N. Johnson, S. Doyle, G. Turashvili, E. Provenzano, S. Aparicio, C. Caldas, and F. Markowetz, "Quantitative image

analysis of cellular heterogeneity in breast tumors complements genomic profiling," *Science Translational Medicine*, vol. 4, no. 157, pp. 157ra143–157ra143, 2012.

[13] A. Nagaishi, M. Takagi, A. Umemura, M. Tanaka, Y. Kitagawa, M. Matsui, M. Nishizawa, K. Sakimura, and K. Tanaka, "Clinical features of neuromyelitis optica in a large japanese cohort: Comparison between phenotypes," *Journal of Neurology, Neurosurgery and Psychiatry*, vol. 82, no. 12, pp. 1360–1364, 2011.

[14] C. Robert, A. Ribas, J. D. Wolchok, F. S. Hodi, O. Hamid, R. Kefford, J. S. Weber, A. M. Joshua, W.-J. Hwu, T. C. Gangadhar, A. Patnaik, R. Dronca, H. Zarour, R. W. Joseph, P. Boasberg, B. Chmielowski, C. Mateus, M. A. Postow, K. Gergich, J. Elassaiss-Schaap, X. N. Li, R. Iannone, S. W. Ebbinghaus, S. P. Kang, and A. Daud, "Anti-programmed-death-receptor-1 treatment with pembrolizumab in ipilimumab-refractory advanced melanoma: A randomised dose-comparison cohort of a phase 1 trial," *The Lancet*, vol. 384, no. 9948, pp. 1109–1117, 2014.

[15] L. Cibulski and B. Preim, "Visual analytics support for analysis of cohort study data: Requirements and concepts," tech. rep., Otto-Von-Guericke University Magdeburg, 2016.

[16] H.-G. Pagendarm and F. H. Post, "Comparative visualization: Approaches and examples," in *Visualization in Scientific Computing* (M. Göbel, H. Müller, and B. Urban, eds.), Springer, 1995.

[17] A. Somarakis, V. van Unen, F. Koning, B. P. Lelieveldt, and T. Höllt, "ImaCytE: visual exploration of cellular microenvironments for imaging mass cytometry data," *IEEE Transactions on Visualization and Computer Graphics*, vol. 27, no. 1, pp. 98–110, 2021.

[18] B. Preim, P. Klemm, H. Hauser, K. Hegenscheid, S. Oeltze, K. Toennies, and H. Völzke, "Visual analytics of image-centric cohort studies in epidemiology," in *Visualization in Medicine and Life Sciences III. Mathematics and Visualization.* (L. Linsen, B. Hamann, and H. C. Hege, eds.), pp. 221–248, Springer, 2016.

[19] O. Dzyubachyk, J. Blaas, C. P. Botha, M. Staring, M. Reijnierse, J. L. Bloem, R. J. Van Der Geest, and B. P. Lelieveldt, "Comparative exploration of whole-body MR through locally rigid transforms," *International Journal of Computer Assisted Radiology and Surgery*, vol. 8, no. 4, pp. 635–647, 2013.

[20] M. D. Steenwijk, J. Milles, M. Buchem, J. Reiber, and C. P. Botha, "Integrated visual analysis for heterogeneous datasets in cohort studies," in *Proceedings of the Workshop on Visual Analytics in Healthcare (VAHC)*, 2010.

[21] Z. Zhang, D. Gotz, and A. Perer, "Interactive visual patient cohort analysis," in *Proceedings of the Workshop on Visual Analytics in Healthcare (VAHC)*, 2012.

[22] K. Dinkla, H. Strobelt, B. Genest, S. Reiling, M. Borowsky, and H. Pfister, "Screenit: Visual analysis of cellular screens," *IEEE Transactions on Visualization and Computer Graphics*, vol. PP, no. 99, pp. 1–1, 2017.

[23] R. Krueger, J. Beyer, W. D. Jang, N. W. Kim, A. Sokolov, P. K. Sorger, and H. Pfister, "Facetto: Combining unsupervised and supervised learning for hierarchical phenotype analysis in multi-channel image data," *IEEE Transactions on Visualization and Computer Graphics*, vol. 26, no. 1, pp. 227–237, 2020.

[24] D. Schapiro, H. W. Jackson, S. Raghuraman, J. R. Fischer, V. R. T. Zanotelli, D. Schulz, C. Giesen, R. Catena, Z. Varga, and B. Bodenmiller, "histoCAT: analysis of cell phenotypes and interactions in multiplex image cytometry data," *Nat Methods*, vol. 14, no. 9, pp. 873–876, 2017.

[25] C. R. Stoltzfus, J. Filipek, B. H. Gern, B. E. Olin, J. M. Leal, Y. Wu, M. R. Lyons-Cohen, J. Y. Huang, C. L. Paz-Stoltzfus, C. R. Plumlee, *et al.*, "CytoMAP: A spatial analysis toolbox reveals features of myeloid cell organization in lymphoid tissues," *Cell reports*, vol. 31, no. 3, p. 107523, 2020.

[26] R. Rashid, Y.-A. Chen, J. Hoffer, J. L. Muhlich, J.-R. Lin, R. Krueger, H. Pfister, R. Mitchell, S. Santagata, and P. K. Sorger, "Interpretative guides for interacting with tissue atlas and digital pathology data using the minerva browser," *bioRxiv*, 2020.

[27] M. Gleicher, D. Albers, R. Walker, I. Jusufi, C. D. Hansen, and J. C. Roberts, "Visual comparison for information visualization," *Information Visualization*, vol. 10, no. 4, 2011.

[28] F. Lindemann, K. Laukamp, A. H. Jacobs, and K. Hinrichs, "Interactive comparative visualization of multimodal brain tumor segmentation data," in *Proceedings of Vision, Modeling & Visualization (VMV)*, 2013.

[29] A. Maries, N. Mays, M. Hunt, K. F. Wong, W. Layton, R. Boudreau, C. Rosano, and G. E. Marai, "GRACE: a visual comparison framework for integrated spatial and non-spatial geriatric data," *IEEE Transactions on Visualization and Computer Graphics*, vol. 19, no. 12, pp. 2916–2925, 2013.

[30] C. Ma, F. Pellolio, D. A. Llano, K. A. Stebbings, R. V. Kenyon, and G. E. Marai, "RemBrain: exploring dynamic biospatial networks with mosaic matrices and mirror glyphs," *Journal of Imaging Science and Technology R*, vol. 61, no. 6, pp. 0–1, 2017.

[31] J. Schmidt, M. E. Gröller, and S. Bruckner, "VAICo: visual analysis for image comparison," *IEEE Transactions on Visualization and Computer Graphics*, vol. 19, no. 12, pp. 2090–2099, 2013.

[32] R. Raidou, O. Casares-Magaz, A. Amirkhanov, V. Moiseenko, L. Muren, J. Einck, A. Vilanova, and M. Gröller, "Bladder Runner : Visual analytics for the exploration of rt-induced bladder toxicity in a cohort study," *Computer Graphics Forum*, vol. 37, no. 3, pp. 205–216, 2018.

[33] R. C. Basole, H. Park, M. Gupta, M. L. Braunstein, D. H. Chau, M. Thompson, V. Kumar, R. Pienta, and M. Kahng, "A visual analytics approach to understanding care process variation and conformance," in *Proceedings of the Workshop on Visual Analytics in Healthcare (VAHC)*, 2015.

[34] M. Wagner, D. Slijepcevic, B. Horsak, A. Rind, M. Zeppelzauer, and W. Aigner, "KAVAGait: knowledge-assisted visual analytics for clinical gait analysis," *IEEE Transactions on Visualization and Computer Graphics*, vol. 25, no. 3, pp. 1528–1542, 2019.

[35] C. Zhang, T. Höllt, M. W. A. Caan, E. Eisemann, and A. Vilanova, "Comparative visualization for diffusion tensor imaging group study at multiple levels of detail," in *Proceedings of Visual Computing for Biology and Medicine (VCBM)*, pp. 53–62, 2017.

[36] A. M. Femino, F. S. Fay, K. Fogarty, and R. H. Singer, "Visualization of single rna transcripts in situ," *Science*, vol. 280, no. 5363, pp. 585–590, 1998.

[37] L. Keren, M. Bosse, S. Thompson, T. Risom, K. Vijayaragavan, E. McCaffrey, D. Marquez, R. Angoshtari, N. F. Greenwald, H. Fienberg, J. Wang, N. Kambham, D. Kirkwood, G. Nolan, T. J. Montine, S. J. Galli, R. West, S. C. Bendall, and M. Angelo, "MIBI-TOF: a multiplexed imaging platform relates cellular phenotypes and tissue structure," *Science Advances*, vol. 5, no. 10, p. eaax5851, 2019.

[38] P. J. Schüffler, D. Schapiro, C. Giesen, H. A. Wang, B. Bodenmiller, and J. M. Buhmann, "Automatic single cell segmentation on highly multiplexed tissue images," *Cytometry Part A*, vol. 87, no. 10, pp. 936–942, 2015.

[39] R. Mayeux, "Biomarkers: Potential uses and limitations," *NeuroRX*, vol. 1, no. 2, pp. 182–188, 2004.

[40] M. E. Ijsselsteijn, T. P. Brouwer, Z. Abdulrahman, E. Reidy, A. Ramalheiro, A. M. Heeren, A. Vahrmeijer, E. S. Jordanova, and N. F. de Miranda, "Cancer immunophenotyping by seven-colour multispectral imaging without tyramide signal amplification," *The Journal of Pathology: Clinical Research*, vol. 5, pp. 3–11, 2019.

[41] M. Brehmer and T. Munzner, "A multi-level typology of abstract visualization tasks," *IEEE Transactions on Visualization and Computer Graphics*, vol. 19, no. 12, pp. 2376–2385, 2013.

[42] V. van Unen, N. Li, I. Molendijk, M. Temurhan, T. Höllt, A. E. van der Meulen-de Jong, H. W. Verspaget, M. L. Mearin, C. J. Mulder, J. van Bergen, B. P. Lelieveldt, and F. Koning, "Mass cytometry of the human mucosal immune system identifies tissue- and disease-associated immune subsets," *Immunity*, vol. 44, no. 5, pp. 1227–1239, 2016.

[43] M. Allen, D. Poggiali, K. Whitaker, T. R. Marshall, and R. Kievit, "Raincloud plots: a multi-platform tool for robust data visualization," *Wellcome open research*, vol. 4, 2019.

[44] M. Blumenschein, L. J. Debbeler, N. C. Lages, B. Renner, D. A. Keim, and M. El-Assady, "v-plots: Designing hybrid charts for the comparative analysis of data distributions," *Computer Graphics Forum*, vol. 39, no. 3, pp. 565–577, 2020.

[45] M. Correll, M. Li, G. Kindlmann, and C. Scheidegger, "Looks good to me: Visualizations as sanity checks," *IEEE Transactions on Visualization and Computer Graphics*, vol. 25, no. 1, pp. 830–839, 2019.

[46] P. Kampstra, "Beanplot: A boxplot alternative for visual comparison of distributions," *Journal of Statistical Software*, vol. 28, no. 1, pp. 1–9, 2008.

[47] E. R. Tufte, *Envisioning information*. Graphics Press, 1990.

[48] P. J. Rousseeuw, "Silhouettes: A graphical aid to the interpretation and validation of cluster analysis," *Journal of Computational and Applied Mathematics*, vol. 20, no. C, pp. 53–65, 1987.

[49] J. C. Bezdek and N. R. Pal, "Cluster validation with generalized dunn's indices," in *Proceedings of Artificial Neural Networks and Expert Systems (ANNES)*, pp. 190–193, 1995.

[50] B. Shneiderman, "Dynamic queries for visual information seeking," *IEEE Software*, vol. 11, no. 6, pp. 70–77, 1994.

[51] C. Stolte and P. Hanrahan, "Polaris: A system for query, analysis and visualization of multi-dimensional relational databases," *IEEE Transactions on Visualization and Computer Graphics*, vol. 8, pp. 52–65, 2002.

[52] J. Heer and B. Shneiderman, "Interactive dynamics for visual analysis," *Communications of the ACM*, vol. 55, no. 4, p. 45–54, 2012.

[53] C. A. Brewer, G. W. Hatchard, and M. A. Harrower, "ColorBrewer in print: A catalog of color schemes for maps," *Cartography and Geographic Information Science*, vol. 30, no. 1, pp. 5–32, 2003.

[54] A. Somarakis and T. Höllt, "SpaCeCo open source software." https://www.doi.org/10.5281/zenodo.3885814, 2020. [Online; Accessed: 2021-07-14].

[55] GF, "Google forms." `https://www.google.com/forms/about/`. [Online; Accessed: 2020-04-20].

[56] J. Brooke, "SUS: a "quick and dirty" usability scale," in *Usability Evaluation in Industry* (P. W. Jordan, B. Thomas, B. A. Weerdmeester, and I. L. McClelland, eds.), pp. 189–194, Taylor and Francis, 1996.

[57] S. J. Luk, D. M. der Steen, R. S. Hagedoorn, E. S. Jordanova, M. W. Schilham, J. V. Bovée, A. H. Cleven, J. F. Falkenburg, K. Szuhai, and M. H. Heemskerk, "PRAME and HLA Class I expression patterns make synovial sarcoma a suitable target for PRAME specific t-cell receptor gene therapy," *OncoImmunology*, vol. 7, no. 12, p. e1507600, 2018.

[58] A. Bangor, P. Kortum, and J. Miller, "Determining what individual SUS scores mean: Adding an adjective rating scale," *Journal of Usability Studies*, vol. 3, pp. 114–1234, 1996.

5

Analysis of multiplexed cellular images in Alzheimer's study

This chapter was adapted from:

B. Kenkhuis*, **A. Somarakis***, L. de Haan, O. Dzyubachyk, M. E. IJsselsteijn, N. F. de Miranda, B. P. Lelieveldt, J. Dijkstra, W. M. van Roon-Mom, T. Höllt, and L. van der Weerd, "Iron loading is a prominent feature of activated microglia in Alzheimer's disease patients ," *Acta Neuropathologica Communications*, vol. 9, no. 1, pp. 1–15, 2021.

* Contributed equally

5.1. ABSTRACT

Brain iron accumulation has been found to accelerate disease progression in amyloid-β(Aβ) positive Alzheimer patients, though the mechanism is still unknown. Microglia have been identified as key players in the disease pathogenesis, and are highly reactive cells responding to aberrations such as increased iron levels. Therefore, using histological methods, multispectral immunofluorescence and an automated in-house developed microglia segmentation and analysis pipeline, we studied the occurrence of iron-accumulating microglia and the effect on its activation state in human Alzheimer brains. We identified a subset of microglia with increased expression of the iron storage protein ferritin light chain (FTL), together with increased Iba1 expression, decreased TMEM119 and P2RY12 expression. This activated microglia subset represented iron-accumulating microglia and appeared morphologically dystrophic. Multispectral immunofluorescence allowed for spatial analysis of FTL$^+$Iba1$^+$-microglia, which were found to be the predominant β-plaque infiltrating microglia. Finally, an increase of FTL$^+$Iba1$^+$-microglia was seen in patients with high β load and Tau load. These findings suggest iron to be taken up by microglia and to influence the functional phenotype of these cells, especially in conjunction with β.

5

5.2. INTRODUCTION

Alzheimer's disease is the most common cause of dementia, and is defined by the presence of amyloid-β (β) plaques and tau tangles. In addition, the brain's resident innate immune cells, microglia, have been found to be at the centre-stage of the disease, as most identified risk genes are predominantly or even exclusively expressed in microglia [1, 2].

Not only can microglia modulate Alzheimer's disease, but many transcriptomic studies showed microglia to undergo the most pronounced changes in response to pathology. In mice, a subset of responding microglia has been found to lose their homeostatic molecular signature and transition into a so-called 'disease-associated microglia' (DAM) state [3]. In humans, a comparable yet disparate state coined the human Alzheimer microglia (HAM) has been identified [4]. Upregulated genes in these subsets do not only indicate loss of homeostatic function and increased pro-inflammatory activation, but also dysregulated iron-metabolism, manifested via upregulation of the FTL-gene and downregulation of FTH1 and SLC2A11 [4, 5]. FTL encodes the ferritin light chain (FTL) protein, the component of the major iron-storage complex ferritin, responsible for the long term storage of iron. These transcriptomic findings coincide with previously observed ferritin$^+$ microglia in Alzheimer's disease [6, 7]. Though increased iron concentration likely plays a role, the exact link between the two has not yet been established.

Iron accumulation, irrespective of microglial activation, on the other hand, has been reported in disease-affected areas in Alzheimer's disease, using both in-vivo and post-mortem human MRI [8]. Several MRI and histology studies found high correlations between iron accumulation and cortical β and tau spreading [9–11]. Clinically, increased iron concentrations were shown to accelerate cognitive decline in β-positive Alzheimer patients, indicative of a disease-modifying role for iron accumulation [12, 13]. Again, how iron accelerates cognitive deterioration is poorly understood.

Therefore, in this study we aimed to research the possible link between iron accumulation and functionally activated microglia, and finally, its relation with β-plaques.

We performed a comprehensive investigation of iron-accumulating microglia, and first identified that the iron-storage protein FTL, specifically reflected increased iron accumulation in microglia. Secondly, by using multispectral immunofluorescence and an in-house automated cell-analysis pipeline, we found FTL^+ microglia to show significant activation, shown via both downregulation of homeostatic markers TMEM119 and P2RY12 and dystrophic morphology, and to predominantly infiltrate β-plaques. This provides evidence for iron dysregulation as a prominent feature of activated microglia in Alzheimer's disease in humans.

5.3. METHODS

5.3.1. TISSUE ACQUISITION

Brain autopsy tissue of the middle temporal gyrus (MTG) of 12 Alzheimer patients and 9 age-matched controls was collected at the Leiden University Medical Center (LUMC), Netherlands Brain Bank (NBB) and the Normal Aging Brain collection Amsterdam (NABCA). Patients were included based on clinical presentation and diagnosis was confirmed by a neuropathologist. The neuropathologists also evaluates Braak stage, based on Gallyas and Tau immunohistochemistry (IHC), and Thal phase based on Congo Red and Amyloid Beta IHC, in eighteen standard regions, according to the latest international diagnostic criteria [14–16]. Patient demographics are reported in Table 5.1. All material has been collected with written consent from the donors and the procedures have been approved by the Medical Ethical committee of the LUMC and the Amsterdam UMC.

5.3.2. HISTOLOGY AND IMMUNOHISTOCHEMISTRY

Formalin fixed paraffin embedded (FFPE) tissue was serially cut into ten 5-µm-thick and four 10-um-thick sections. Consecutive 10-µm-thick sections were used for histological detection of iron using an enhanced Perl's stain and IHC detection of Ferritin Light Chain (FTL). 5-µm-thick sections were used for staining of the microglia multispectral immunofluorescence (mic-mIF) panel (Table 5.2) to verify expression of FTL in microglia/macrophages (Iba1), look at the activation state of these cells (P2RY12/TMEM119) and study the interaction with Aβ-plaques. Finally, of three subjects, 20-µm-thick sections were obtained for 3D confocal imaging. Step-by-step histological and IHC optimization protocols, together with the imaging parameters, are reported in the Supplementary Methods. A step-by-step mIF protocol and further analysis of the described histological, IHC and mIF staining will be described in the following sections.

5.3.3. MICROGLIA MULTISPECTRAL IMMUNOFLUORESCENCE PANEL

One 5-µm-thick section of each subject was stained with the mic mIF panel with the following protocol, based on a previously described protocol by IJsselsteijn et al. [17]. Sections were deparaffinized with 3×5 min xylene, rinsed twice in 100 % alcohol and subsequently washed with 100% ethanol for 5 min. Endogenous peroxidases were blocked for 20 min in 0.3% H2O2/methanol, after which the slides were rinsed with 70% and 50% alcohol. Heat induced antigen-retrieval was performed by cooking the slides for 10 min in pre-heated citrate (10 mM, pH=6.0) buffer for 10 min. After cooking, excess buffer was removed and slides were cooled for 60 min. Non-specific antibody binding sites

were blocked with blocking buffer (0.1% BSA/PBS+0.05% Tween) for 30 min. Firstly, slides were incubated with anti-TMEM119 (1:250, Sigma Aldrich) diluted in blocking buffer overnight at RT. Slides were washed thrice with PBS and incubated with Poly-HRP secondary antibody for 30 min. Slides were washed again and incubated with the appropriate Opal tertiary antibody (1:100 in amplification diluent, Perkin Elmer) for 60 min, which causes permanent binding of the fluorophore to the antigen site. All subsequent steps are performed in the dark where possible. Finally, the slides are placed back in citrate buffer and cooked in the microwave for 15 min to wash the primary antibody off. The same steps are repeated for anti-P2RY12 (1:2500, Sigma Aldrich). After binding of the two antibodies amplified with Opal, slides are incubated with a primary antibody mix with anti-FTL (1:100, Abcam), anti-Aβ (17–24) (1:250, Biolegend) and anti-Iba1 (1:20, Millipore) antibodies, diluted in blocking buffer, overnight at room temperature. The next day, after three washes with PBS, slides are incubated with a secondary antibody mix of G-a-rIgG A594, G-a-mIgG2b A647 and G-a-mIgG1 CF680 (1:200, ThermoFisher), diluted in 0.1% BSA/BPS. Finally, the slides are washed and incubated with 0.1 µg/mL DAPI (Sigma Aldrich) for 5 min, after which they are mounted with 30 uL Prolong diamond (ThermoFisher).

5.3.4. POST-MORTEM MRI ACQUISITION AND ANALYSIS

MRI data and T2*-w severity scores were obtained from a previous study by Bulk et al. [10], on the same tissue-blocks. In this study, tissue blocks were put in proton-free fluid (Fomblin LC08, Solvay), and scanned at room temperature on a 7 T horizontal-bore Bruker MRI system equipped with a 23 mm receiver coil and Paravision 5.1 imaging software (Bruker Biospin, Ettlingen, Germany). A gradient echo scan was acquired with repetition time=75.0 ms, echo time=33.9 ms, flip angle=25° at 100 µm isotropic resolution with 20 signal averages. Subsequently, cortices were assessed for changes in MRI contrast following a pre-defined scoring system.

5.3.5. IRON-POSITIVE CELL IDENTIFICATION

Whole slide scans of the histochemical iron staining were exported from Philips Intellisite digital Pathology Solution platform (Philips, the Netherlands) and imported into ImageJ. RGB images were converted into 8-bit greyscale images. Subsequently, while blinded for diagnosis, for each subject an optimal threshold was set to include DAB-positive intracellular iron depositions, but exclude extracellular background signal. The cortex of the MTG was delineated and the number of positive cells was determined using the ImageJ particle analyser, with a size threshold of 4–100 pixels. Subject AD5 was excluded from this analysis, as iron-accumulating cells could not be distinguished due to high extracellular iron load.

5.3.6. SINGLE CELL SEGMENTATION

For the presice identification of microglia's cell boundaries the algorithm, stated in Section 2.2, was utilized.

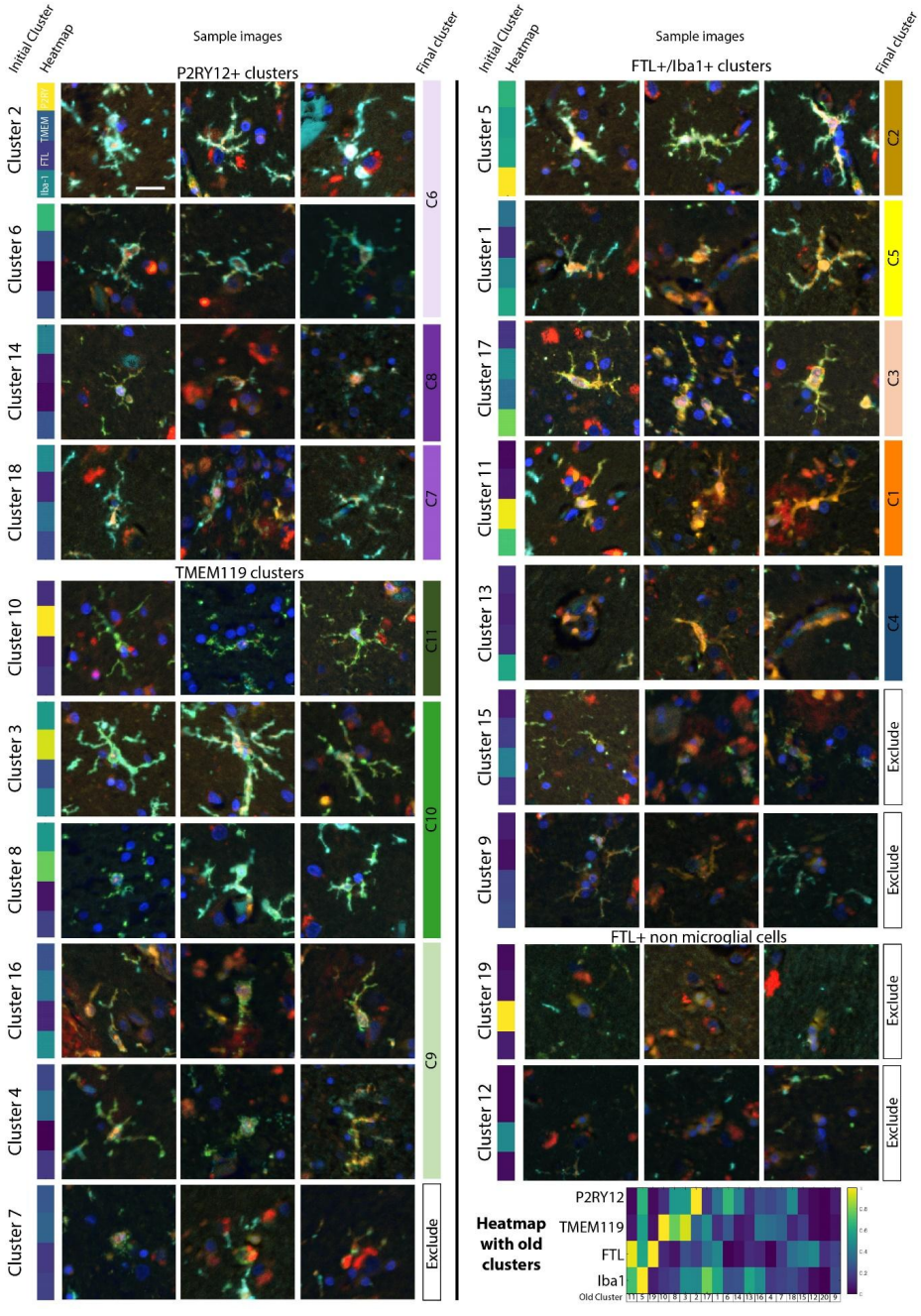

Figure 1: *Sample images of all original clusters, and how they were merged. Scale bar, 20* µm.

5.3.7. CELL PHENOTYPE IDENTIFICATION

Superimposing the segmentation masks onto the component image of all membrane markers, four mean intensity values were extracted for each cell. Afterwards, intensity values were normalized imposing Z-score transformation. For the definition of the different microglia cell types Phenograph [18], an unsupervised clustering method, was utilized. For Phenograph 100 nearest neighbours along with the default parameters were selected, in order to avoid overclustering due to the limited amount of markers. Subsequently, for each Phenograph identified cluster the variability of the single-cell marker expression values was examined (Figure 1) through a violin plot [19] indicating the variation in each cluster, in parallel with their expression patterns as illustrated in the composite images.

5.3.8. ANALYSIS OF CELLULAR PHENOTYPES

The median expression value of each marker for each phenotype was illustrated with a heatmap. The similarities among the identified phenotypes were observed from a t-SNE [20] embedding using the same input as in Phenograph and the default parameters. The t-SNE embedding was coloured according to the cluster of each cell, its cohort or its individual marker expression values [21].

To explore the differences between the Alzheimer patients and controls regarding their phenotypes and their spatial relationship with the Aβ-plaques, an interactive, data-driven pipeline described by [22] was utilized. First, using a version of raincloud plots [22] the phenotypes that exist predominantly in each cohort are identified and consequently, their relative position regarding the Aβ -plaques using a visual query system are explored. For the exploration of the variability in each subject and the validation of our findings, a customized version of the motif glyphs described in our previous work [23] was employed.

5.3.9. STATISTICAL ANALYSIS

Firstly, variables were inspected for being gaussian distributed. If normally-distributed, data plots represent the mean and the standard deviation. For not normally-distributed data, data plots show the median with the corresponding interquartile range. Comparison of two continuous variables was performed using a two-tailed unpaired Student's independent t-test (normally-distributed) or a Mann–Whitney U test (not normally-distributed). Paired normally distributed data were analysed using a two-tailed paired Student's t-test. Bonferroni post-hoc analysis was performed, and a significance level of $P<0.05$ was used. The linear correlation between identified number of cells and different pathological hallmarks was assessed using the Pearson correlation coefficient. All statistical tests were performed using GraphPad Prism (Version 8.00, La Jolla, San Diego, CA, USA).

5.4. RESULTS

5.4.1. FTL$^+$- MICROGLIA REFLECT IRON ACCUMULATING MICROGLIA IN ALZHEIMER'S DISEASE

An enhanced Perl's staining for iron revealed an abundance of iron-positive cells in the cortex of the MTG in Alzheimer's patients. On further inspection, iron-positive cells showed characteristic microglia morphology with a small soma and many thin processes (Figure 2a) and quantification indicated a significant increase of iron-positive cells in

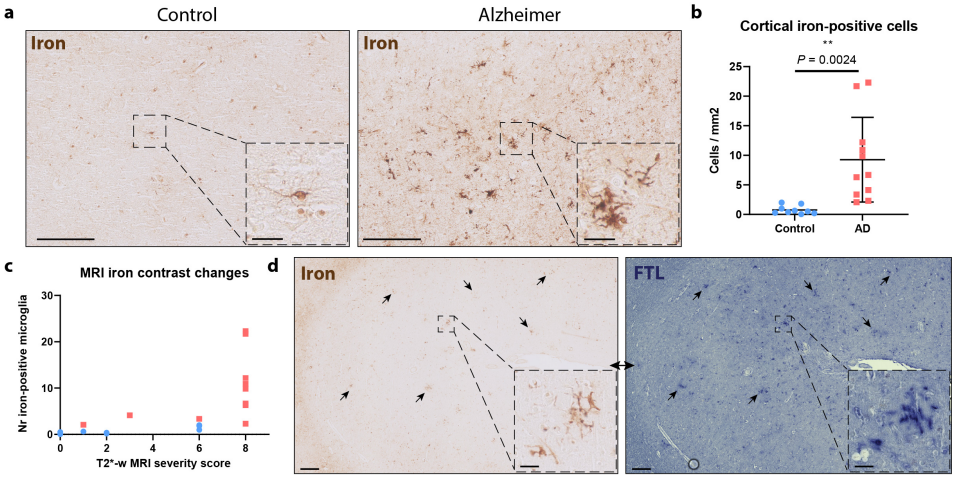

Figure 2: ***Increased iron-positive and corresponding FTL⁺-microglia in Alzheimer's disease.****(a) MTG cortex of Alzheimer patients shows increased positivity for iron inside cells with microglial morphology. (b) Significant increase of iron-positive cells in Alzheimer patients (n12) compared to controls (n9)(Mean, Student's t-test). (c) Iron-positive microglia number only increased in cases with severe signal alterations on iron-sensitive T2*-w MRI, reflected by MRI severity score. (d) FTL expression reflects intracellular iron accumulation. Scale overview images, 200 µm. Scale zooms, 30 µm.*

Alzheimer patients compared to controls (P=0.0024; Figure 2b). Additionally, iron-positive cells appeared to cluster in groups, something that was not observed in control patients (Figure 2a). All MTG tissue blocks have also previously been scanned using T2*-w MRI, sensitive for paramagnetic substances such as iron. MRI images were scored based on alterations in signal intensity reflecting overall parenchymal iron accumulation and focal iron depositions, and were published by Bulk et al. [10]. An increase of iron-positive microglia appeared to be only present in cases with the highest MRI severity score, indicating a significant increase of iron-positive microglia only to occur in subjects with a pronounced macroscopic iron-phenotype (Figure 2c). Subsequently we studied the correspondence of iron accumulation with altered expression of the main iron-storage protein ferritin light chain (FTL), as FTL is known to be expressed in microglia and oligodendrocytes, whereas heavy chain ferritin is primarily expressed by neurons in Alzheimer tissue [24]. The Perl's staining and the FTL staining showed a highly similar staining pattern, with focal clusters of cells representing microglia morphology (Figure 2d). Thus, increased expression of the main iron-storage protein FTL appears to reflect iron accumulation in microglial cells.

5.4.2. QUANTITATIVE ANALYSIS ENABLES MICROGLIA PHENOTYPING

To confirm the microglial origin of FTL$^+$ cells, study their activation state and potential interaction with Aβ, we designed the microglia multispectral immunofluorescence (mic-mIF) panel that can simultaneously detect 6 different markers (Table 5.2). The MTG of 12

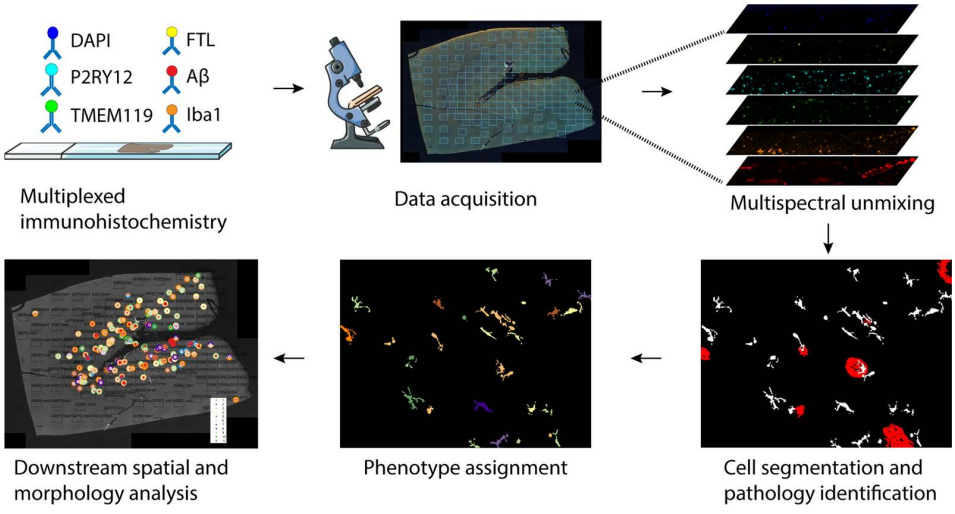

Figure 3: ***Schematic of mic-mIF acquisition and analysis pipeline.***

5

Alzheimer patients, both of early- and late onset, and 9 control subjects (Table 5.1) was stained and imaged. After image acquisition and multispectral unmixing of the data, images were exported for automated segmentation, phenotyping and spatial analysis (Figure 3). In total, 3149 images (110–236 per subject) were obtained. Multispectral unmixing allowed for simultaneous detection of FTL with the nuclear marker DAPI, TMEM119, P2RY12, Iba1 and Aβ at 0.5×0.5 µm resolution (Figure 4a). TMEM119 and P2RY12 are generally considered homeostatic microglia-specific markers, based on transcriptomic [3], in vitro [25–27] and post-mortem IHC studies [27–29], whose expression decreases when activated. Iba1, on the other hand, is a pan microglia/macrophage marker, which is upregulated upon activation. Finally Aβ stains the characteristic pathological Aβ-plaques that form in the parenchyma of Alzheimer patients. Images were segmented using a targeted in-house segmentation pipeline allowing segmentation of cells with processes (like microglia) in 2D images (Figure 4b; Section 2.2). After segmentation, unsupervised clustering using Phenograph assigned single segmented cells to 20 separate clusters. Following manual evaluation of the unsupervised clusters, 6 clusters were excluded based on non-microglial morphology and/or sub-threshold expression of all microglial markers (TMEM119/P2RY12/Iba1). In addition, three times two clusters were merged based on similarity in protein expression levels and their visual appearance (Figure 1). Exclusion of the non-microglial cells resulted in identification of 69,227 cells, with no significant differences in the number of microglia per mm2 between control and Alzheimer patients in either grey matter (GM) or white matter (WM) (Figure 4c). The remaining 11 clusters (C1–C11) were identified as major microglia phenotype clusters (Figure 4d). Though the 11 different phenotypes clustered on the t-SNE plot, the low degree of separation suggests a rather continuous spectrum of expression of the microglia markers (Figure 4e). The control and Alzheimer patients did cluster together, and the marker-based t-SNE plots already revealed more cells with high TMEM119 and P2RY12 expression in controls, but increased FTL expression in Alzheimer

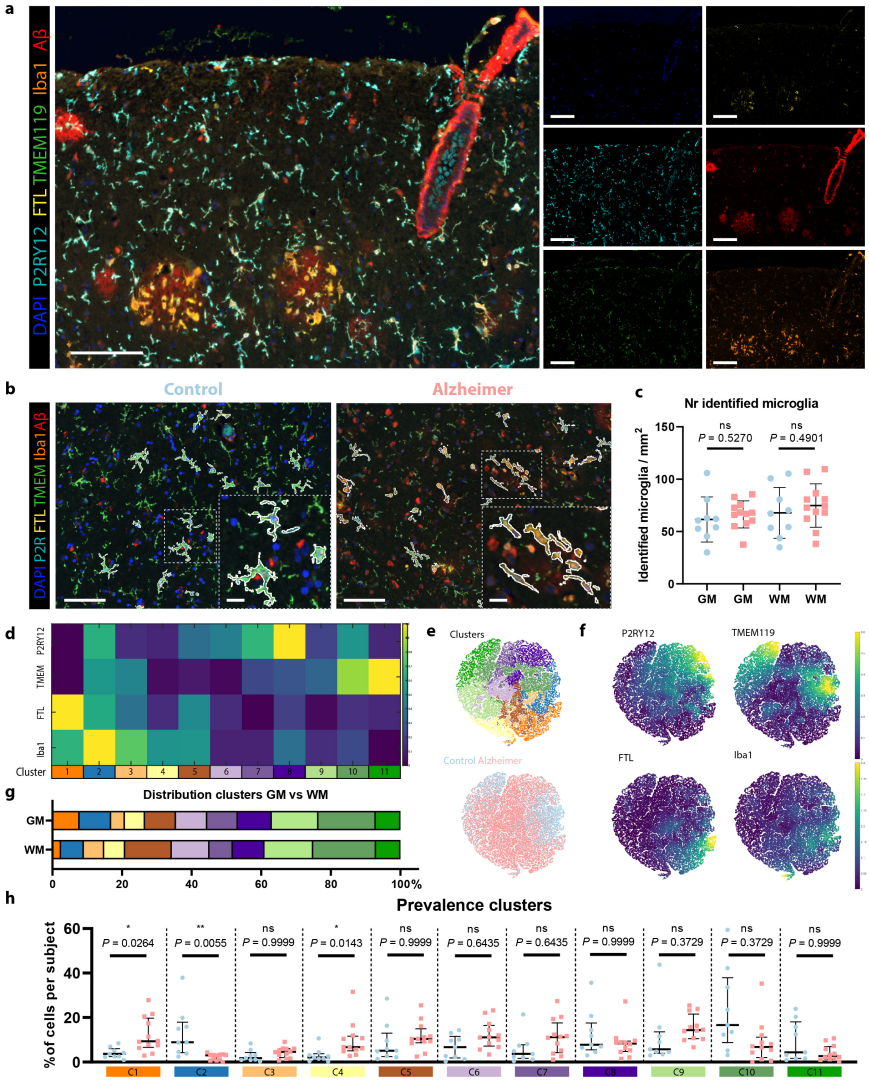

Figure 4: ***Identification of homeostatic and activated Alzheimer-associated microglia clusters.*** *(a) Example of mIF image of an Alzheimer patient. (b) Exemplary images of segmented microglia in a control and an Alzheimer patient. (c) Number of identified cells in the GM and WM of controls (blue; n=9) and Alzheimer patients (red; n=12)(Mean, Student's t-test). (d) Heatmap showing the expression of the four different markers (P2RY12, TMEM119, FTL and Iba1), in the 11 identified microglia clusters. (e) t-SNE plot of all individual cells showing the distinct colour-coded clusters and of control- vs. Alzheimer-patient-derived cells. (f) t-SNE plots colour-coded for intensity of the four individual markers. (g) Distribution of clusters in GM and WM. (h) Prevalence of identified clusters (C1–C11) in individual control (blue; n=9) and Alzheimer patients (red; n=12) (Median, Mann–Whitney U test). Scale bar, 100 µm. Scale bar zooms, 20 µm. GM Grey matter, WM White matter.*

patients (Figure 4f). With regard to anatomical region, only C1 and C2 appeared to be more present in the grey matter (GM), whereas C5 and C6 appeared to be proportionally more present in the white matter (WM) (Figure 4g). Four FTL^+ clusters (C1–C3, C5) were identified, with differing expression levels and co-expression levels of P2RY12, TMEM119 and Iba1 (Figure 4d). Cluster C1 (FTL^+Iba1^+) appeared significantly more present in Alzheimer patients (P=0.0264), while C2 ($P2RY12^+TMEM119^+FTL^+Iba1^+$) was more present in controls (P=0.0055; Figure 4h). FTL^+Iba1^+ clusters lacking either P2RY12 (C3) or TMEM119 (C5) did not differ significantly in prevalence between control and Alzheimer patients. Cluster C4 showed solely Iba1 expression, meaning that this cluster likely also consists of non-resident infiltrating macrophages. Additionally, three $P2RY12^+$ clusters (C6–C8) were identified, with the highest expressing cluster (C8) being more present in controls. The same applied for the $TMEM119^+$ clusters (C9–C11), with C10 and C11 having higher expression and being more present in control patients. These results indicate a small shift of homeostatic microglia positive for P2RY12 and TMEM119 in controls towards activated microglia, with downregulated expression of P2RY12 and TMEM119 in Alzheimer patients. In addition, a specific Alzheimer-associated cluster shows increased expression of a combination of FTL and Iba1.

5.4.3. SPATIAL ANALYSIS OF FTL^+- MICROGLIA CLUSTERS

After cell phenotype identification, all microglia were assessed for proximity to parenchymal Aβ-plaques. For visualization purposes, a second image was created, where infiltrated Aβ-plaques were plotted onto the original image as a 'glyph' (Figure 5a) [23], with the different colours corresponding to the respective cluster of the infiltrating microglia, to analyse which clusters predominantly infiltrated Aβ-plaques. Subsequently, all individual cells represented as cluster-colored dots or the cluster-colored glyphs were plotted back onto the original whole slide image (Figure 5a), to assess differences in cluster composition of microglial Aβ infiltration on a whole-section scale. As expected, quantification showed significantly more identified Aβ-plaques in Alzheimer patients, although some were found in controls as well (P=0.0002; Figure 5b). Furthermore, a higher percentage of the plaques showed microglia infiltration in Alzheimer patients (P=0.013; Figure 5c). Looking at the whole slide distribution, Aβ-plaques were found to be more present in the coronal sulcus rather than the gyrus. This also appeared to be associated with the regional microglia phenotype, as can be seen for the predominantly purple (C6–C8) microglia populating the Aβ-plaque deplete regions (Figure 5d). To quantify the influence of Aβ-plaques on microglia phenotype, we compared all phenotyped microglia (all-mic) with the subset of microglia infiltrating Aβ-plaques (Aβ-mic). Controls showed a slight percental increase of C1 and C5 in Aβ-mic compared to all-mic, and less Aβ-plaque infiltration of $TMEM119^+$ clusters C9–C11 (Figure 5e), though this was based on a limited total number of Aβ-plaques. Alzheimer patients on the other hand, showed a large percental increase of FTL^+ clusters C1 and C3 in the Aβ-mic population (Figure 5e), which was also statistically significant when looking at subject-specific proportional increases (C1: P<0.0001, C3: P=0.0004; Figure 5f, g). While C1 and C3-microglia together make up less than 20% of all-mic, they constitute almost 50% of the Aβ-mic population (Figure 5e). P2RY12+ clusters C6–C8, on the other hand, showed a small contribution to Aβ-mic compared to all-mic (Figure 5e). Finally, not only did C1 and C3 make up the majority of Aβ-mic, but

also when examining the proportions of these individual clusters that directly infiltrated Aβ-plaques, they showed much higher proportion of infiltration than all the other clusters (Figure 5h). A visual example of the C1 and C3-microglia infiltrating an Aβ-plaque on the original mic-mIF images can be found in Figure 5i. All in all, these results suggest Aβ-plaques to be predominantly infiltrated by a specific subset of microglia, characterized

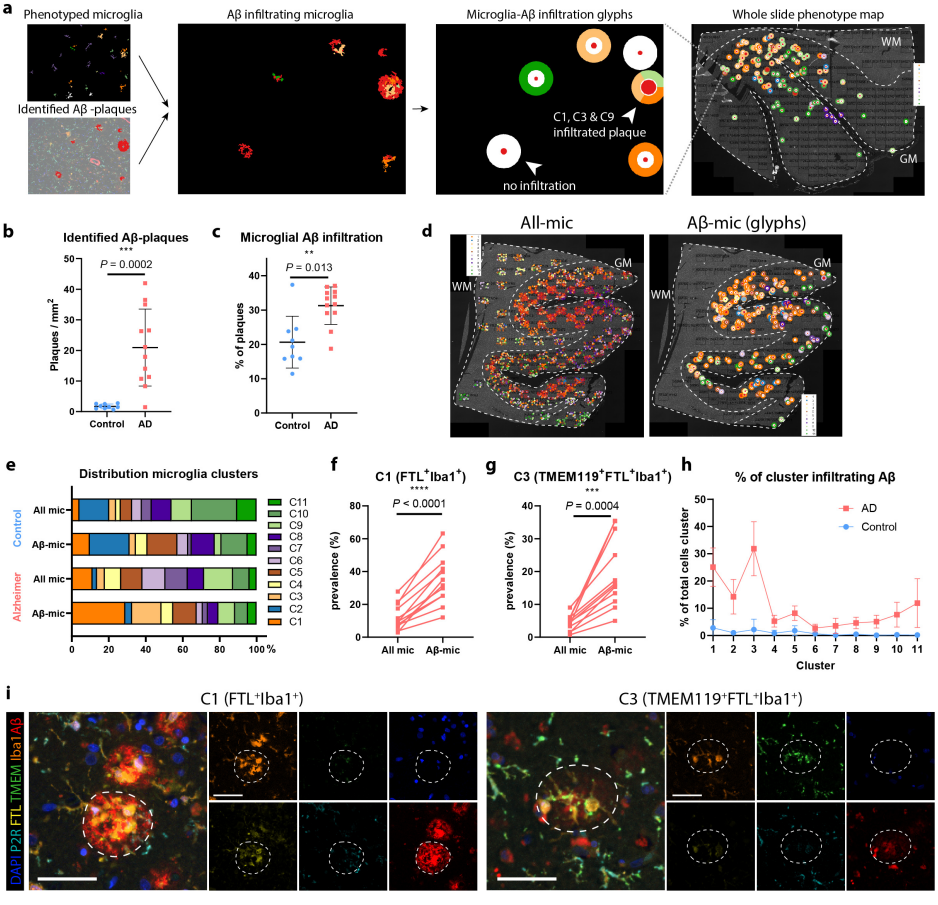

Figure 5: *FTL⁺-microglia show significant Aβ-plaque infiltration.* *(a) Schematic of how microglial Aβ-plaque infiltration is studied. Both cells and Aβ-plaques are identified and an interaction map showing 'glyphs' in the colour of the cluster of the infiltrating microglia is created. Subsequently glyphs are plotted back onto the whole slide image to also enable studying the spatial distribution pattern. Number of identified Aβ-plaques (Mean, Student's t-test) (b) and the percentage of microglia infiltrated Aβ-plaques (Mean, Student's t-test) (c) are increased in Alzheimer's disease (n=12) compared to controls (n=9). (d) Microglia clusters differ spatially, depending on the presence of Aβ-plaques in their proximity. (e) Distribution of all-mic clusters compared to Aβ-mic clusters of controls and Alzheimer patients. Comparison of prevalence of all-mic compared to Aβ-mic of C1- (f) and C3-microglia (g) of all individual Alzheimer patients (n=12) (paired Student's t-test). (h) Percentage of all identified clusters infiltrating Aβ-plaques. (i) Representative images of C1 and C3-microglia infiltrating an Aβ-plaque. Scale bar, 50 µm. All mic=all microglia, Aβ-mic=Aβ-plaque infiltrating microglia.*

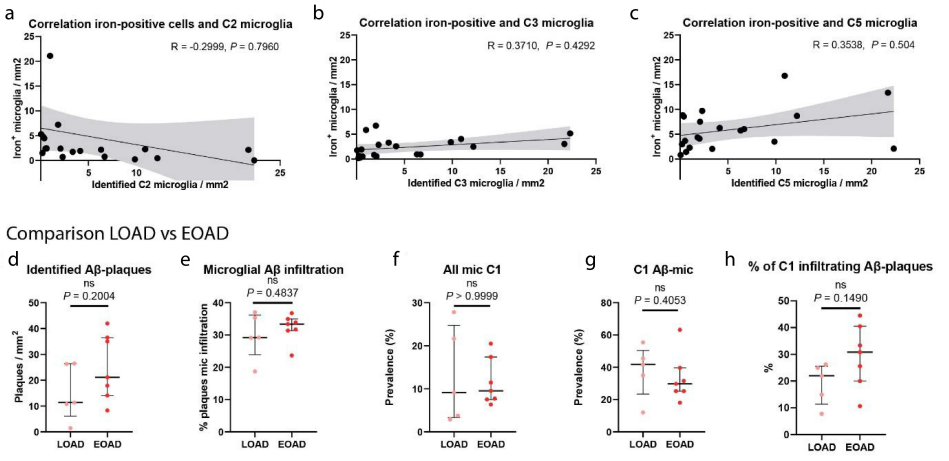

Figure 6: *Correlation of iron+-microglia with C2-microglia (a), C3-microglia (b) and C5-microglia (c) is non-significant (n=20, Pearson coefficient). Comparison between LOAD (n=5) and EOAD (n=7) patients shows no differences in number of identified Aβ-plaques (d), percentage of microglia infiltration (e), prevalence of C1-all-mic (f), prevalence of C1-Aβ-mic (g), and percentage of C1-microglia infiltrating Aβ-plaques (h) (Median, Mann-Whitney U test).*

5

by increased FTL and Iba1 expression and loss of expression of homeostatic markers P2RY12 or TMEM119 and P2RY12.

5.4.4. CORRELATION OF FTL$^+$-MICROGLIA WITH PATHOLOGY

As already shown in Figure 2d, FTL staining closely followed the enhanced Perl's staining showing microglial iron loading. Therefore, we also checked the correlation of the number of iron-positive microglia with the number of identified microglia of different FTL$^+$ clusters. The number of identified C1 (FTL$^+$Iba1$^+$) microglia correlated well with number of iron–positive-cells (R=0.7601, p=0.0004; Figure 7a), while other FTL$^+$ clusters with lower expression (C2, C3) did not show correlation with number of iron-positive cells (Figure 6a, d). This suggests that it is especially the marked increase of FTL expression found in C1-microglia that reflects substantial iron loading, while moderate FTL expression is also found in non-iron accumulating cells in controls. Although we already found C1-microglia to significantly infiltrate Aβ-plaques, we also checked for its correlation with overall Aβ and Tau load, as assessed by a neuropathologist using Thal stage and Braak stage, respectively. A marked increase of the number of C1-microglia was solely found in high-pathology load subjects with Thal phase V, and Braak stage V/VI (Figure 7b, c), though not all high-pathology load subjects show increase of C1-microglia. C2-microglia were primarily found in controls with low Braak stage I/II and Thal I-II, whereas C3-microglia were present in both controls and Alzheimer patients with varying pathological burdens (Figure 6b, c, e, f). This is in line with the finding that iron-positive microglia were particularly present in Alzheimer patients with advanced iron loading. However, there is lack of Alzheimer patients with intermediate Thal- and Braak-scores, making it impossible to state that an increase of C1-microglia is exclusive to advanced stage disease,

and C3 represents an intermediate state between C2 in controls and C1 in advanced disease. Further investigation into the differences between early-onset Alzheimer's disease (EOAD, onset <65y) patients and late-onset Alzheimer's disease (LOAD, onset >65y) patients, showed no differences in Aβ load, microglia prevalence, or Aβ-infiltration of C1, C2 nor C3 (Figure 6g–q). In addition, we looked at differences between APOE3 and APOE4 carriers, as the latter have been found to have elevated ferritin levels in the CSF [30]. As expected, APOE4 carriers had more Aβ-plaques (Figure 7d), but did not show overall increased microglia infiltration (Figure 7e). Though sample sizes for both groups were small (n=4–6), a trend indicating higher prevalence of C1-microglia in the GM could be observed (P=0.0667; Figure 7f), which was not the case for C2 and C3-microglia (Figure 6r, s) However, no difference was observed when looking at the proportion of Aβ-plaques infiltrated by C1-microglia (Aβ-mic) (P=0.5096; Figure 7g). This suggests that even though a higher percentage of C1-microglia infiltrate Aβ-plaques (P=0.0381; Figure 7h), this is likely due to the increased number of Aβ-plaques present in the APOE4 carriers.

5.4.5. FTL$^+$IBA1$^+$-MICROGLIA HAVE A DYSTROPHIC MORPHOLOGICAL APPEARANCE

Finally, we visually evaluated the morphological appearance of all phenotyped microglia in the same dataset, as this provides additional information about the activation stage of

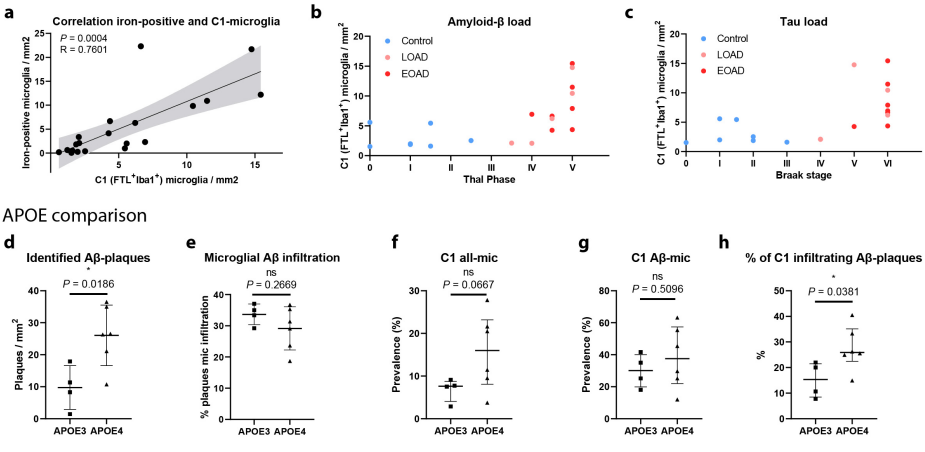

Figure 7: *C1-microglia (FTL$^+$Iba1$^+$) reflect iron-positive microglia.* *(a) Number of identified C1-microglia correlates well with number of identified iron-positive microglia (n=20, Pearson coefficient). Increased number of C1-microglia are associated with higher overall Aβ load (Thal) (b) and Tau Load (Braak) (c). Comparison between APOE4 (n=6) vs. APOE3 (n=4) carriers shows increased number of identified Aβ-plaques (d) and similar microglia infiltration (e). Increased prevalence of C1-all-mic (f), no increased proportion of Aβ-plaques infiltrated with C1-mic (Aβ-mic) (g), and significantly increased proportion of C1-microglia infiltrating Aβ-plaques (h). d and e Median, Mann–Whitney U test. Patients AD8 and AD12 were excluded from the APOE comparison analysis as they harbour a familial mutation in the APP and PSEN1 gene, respectively, which could be of more influence than the APOE-genotype.*

the microglia. Two authors (BK and LdH), evaluated the cells according to five distinctive morphological clusters: homeostatic, activated, dystrophic, phagocytic and perivascular macrophages (Figure 8a), based on previously described morphological phenotypes [31]. The parenchyma of controls was predominantly populated by C6–C11-microglia, which consistently expressed TMEM119 and/or P2RY12. These cells presented with homeostatic morphology, showing small circular or oval cell bodies, with thin highly ramified processes and extensive branches (Figure 8b). Morphological appearance therefore appeared to be in line with the homeostatic protein phenotype. Occasionally activated microglia were identified, which have larger cell bodies and noticeably fewer branches and ramifications

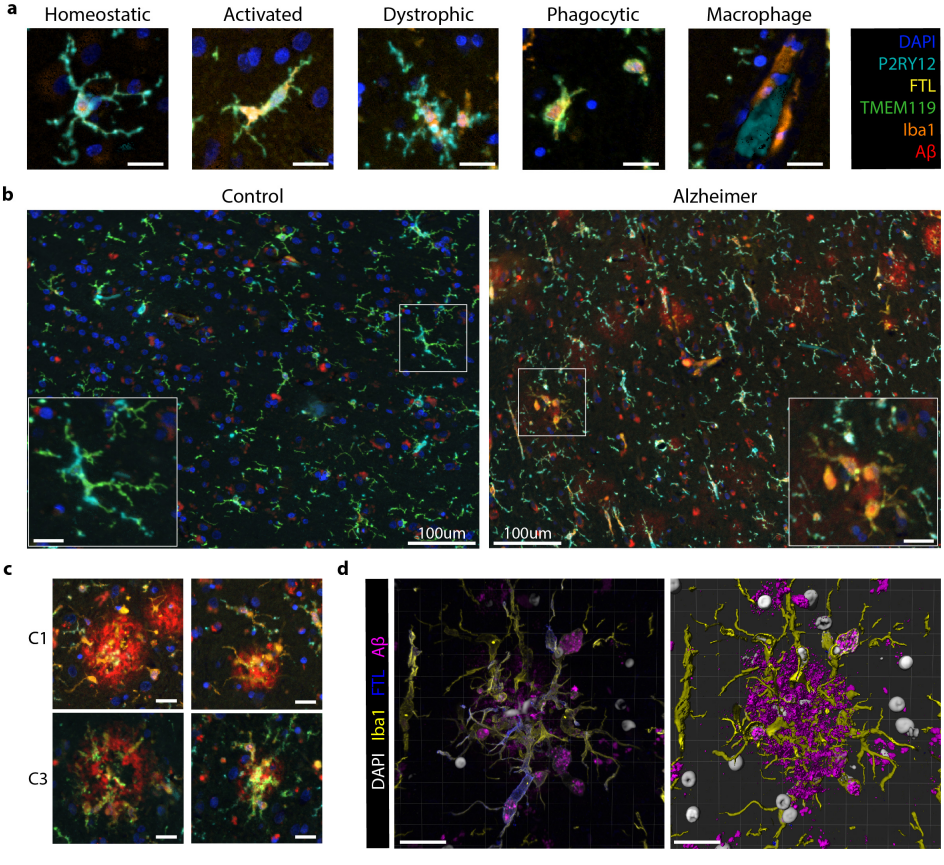

Figure 8: ***C1 and C3-microglia show distinct dystrophic morphology compared to homeostatic control microglia.*** *(a) Representative images of the five different morphological subtypes of microglia: homeostatic, activated, dystrophic, phagocytic and macrophage-like. (b) Controls show predominantly homeostatic and activated microglia, while Alzheimer patients show a variety of homeostatic, activated, phagocytic and dystrophic microglia. (c) Representative images of C1 and C3-microglia surrounding Aβ-plaques showing dystrophic morphology. (d) 3D confocal imaging confirms cytorrhexic appearance of FTL+Iba1+-microglia. Scale bar represents 20 µm unless otherwise stated. Colorcoding for IF-images in 6A-C are according to the box in the top right corner. Colorcoding of 3D confocal images are according to the legend adjacent to the images.*

(especially second degree) (Figure 8a). Activated cells generally showed higher Iba1 and FTL expression and were often phenotyped as C2-microglia (Figure 8a). Microglia in Alzheimer patients, on the other hand, had a much more heterogeneous appearance; homeostatic, activated, dystrophic and phagocytic microglia could all be observed within the coronal sulcus of a single patient (Figure 8b). Though almost all phenotype clusters and morphological clusters could be observed, we focussed on the C1-microglia, as they reflected iron-positive microglia. We found the striking majority of C1-microglia to have a dystrophic morphological appearance. The dystrophic cells show a very distinct phenotype, often with a cloudy or cytorrhexic (fragmentation of the cytoplasm) appearance which results in ill-defined processes (Figure 8a). There is often deramification and the remaining branches show spheroids and fragmentation. Especially microglia (both C1 and C3) infiltrating Aβ plaques showed highly dystrophic morphological characteristics, indicative of an advanced activated/neurodegenerative state (Figure 8c). The dystrophic morphology was also verified using 3D confocal microscopy, which also showed the same cytorrhexic appearance of microglia surrounding the Aβ-plaques (Figure 8d). All in all, the finding of a dystrophic phenotype in C1-microglia was in line with the increased Iba1 and decreased TMEM119 and P2RY12 expression, which accompanied the pronounced FTL expression. They also reflected the morphological appearance of the iron-positive microglia identified on the Perl's staining (Figure 2a).

5.5. DISCUSSION

In this manuscript, we confirmed that increased FTL expression reflects an increase in iron accumulation in microglia in the cortex of Alzheimer patients. Microglia with increased FTL expression also showed higher Iba1 expression, but loss of homeostatic markers TMEM119 and P2RY12, indicative of an activated phenotype. On further investigation this FTL^+Iba1^+ phenotype appeared to be increasingly present in Alzheimer patients and the predominant Aβ-plaque infiltrating microglia phenotype. Morphologically they appeared to be in a dystrophic activation stage.

Firstly, in this study we confirmed that previously identified iron-positive cells in Alzheimer patients [32, 33] are of microglial rather than astrocytic origin, and show high FTL expression. Subsequently, using multispectral fluorescence and unsupervised clustering, we identified several FTL^+ clusters, which were variably present in controls and Alzheimer disease stages. C2-microglia, which displayed positivity for all included microglia markers, were almost exclusively present in control patients. Conversely, C1-microglia (FTL^+Iba1^+) were significantly more present in AD patients, and C3-microglia ($TMEM119^+FTL^+Iba1^+$) were marginally present in either group. Interestingly, both C1 and C3-microglia showed a strong tendency to infiltrate Aβ-plaques. C1-microglia were almost exclusively present in advanced stage Alzheimer patients, whereas C2-microglia were primarily detected in controls (with low Thal/Braak stages), and C3-microglia were variably present across controls and Alzheimer patients of all stages. Regarding the temporal dynamics of these clusters, one could therefore hypothesize that in Alzheimer's disease microglia surround Aβ-plaques and lose P2RY12 expression, as has been observed previously by others (transition from C2 to C3) [28, 34]. As of yet we do not know what the relevance is of the preserved $TMEM119^-$expression. Over time, these microglia take up iron, causing a pronounced increase of FTL expression and loss of TMEM119. This

corresponds to the fact that only C1-microglia appeared to correlate with iron-accumulating microglia. However, our study population is not ideal to dissect the temporal dynamics of these clusters, since the majority of Alzheimer patients showed advanced disease (Braak V/VI) and only two patients showed mild to moderate (Braak III/IV). Future work studying these phenotypes in a larger cohort with a larger range of disease stages would be highly relevant to accurately determine at what stage of the disease C2-microglia prevalence decreases and C1 and C3-microglia prevalence increases.

Several qualitative studies had previously identified increased presence of dystrophic ferritin$^+$ microglia in brain tissue of Alzheimer patients [6, 32, 35, 36]. The dystrophic morphological appearance was also confirmed in this study, though the functional insights of these morphologically defined states remains debatable. Our spatial analysis revealed a strong tendency of FTL$^+$Iba1$^+$ to infiltrate Aβ-plaques; significantly more than can be expected based on prevalence of the cluster itself, and more than any other identified microglia cluster. Although some other studies had already looked into the association of dystrophic ferritin$^+$ microglia with Aβ-plaques [6, 7, 24, 31, 32], results were inconsistent, as none of these studies so far looked into the relative proportion of these microglia in the total population. The importance of this is also stressed in a recent study by Nguyen et al. [37], in which they found an amyloid-responsive microglia (ARM) subset, characterized by CD163, but did not pick up on the Aβ-plaque-infiltrating properties of their identified ferritin$^+$ microglia. Finally, we were able to further characterize iron-positive/FTL$^+$-microglia by analyzing co-expression of several other microglia markers on a single cell level. This revealed that C1-microglia, with the highest FTL protein expression and increased Iba1 expression, showed complete loss of expression of homeostatic markers TMEM119 and P2RY12. Although we acknowledge that our FTL$^+$Iba1$^+$(P2RY12$^-$TMEM119$^-$)-microglia were only characterized using four protein-markers, which is only a fraction compared to the total amount of genes used to define specific transcriptomic states such as the DAM/HAM-states, we do want to highlight the similarities. The DAM/HAM-subsets showed FTL among the highest upregulated genes, with coinciding downregulation of TMEM119 and P2RY12 [3, 4]. Additionally clustering around Aβ-plaques was also reported as a characteristic feature of DAM microglia [3], as is observed for the identified FTL$^+$Iba1$^+$-microglia.

To date, the reason for the observed increase of FTL-expression remains disputed. With FTL being the long-term storage component of ferritin, its expression is likely to be increased in response to increased intracellular labile iron concentrations. Yet, ferritin is also widely recognized as an acute phase reactant and it has also been suggested that microglia upregulate ferritin as a response to exhaustion, caused by the attempt to phagocytose aggregated Aβ [38]. However, our findings show that the identified FTL$^+$Iba1$^+$-microglia closely reflected microglia with high levels of the metal iron, and therefore suggest that the observed increased FTL-expression at least does not merely reflect inflammatory activation or exhaustion, but also increased iron levels. This is in line with a previous study, which found ferritin levels in the CSF to not be associated with an inflammatory response in Alzheimer patients and hypothesized ferritin levels to rather reflect changes in iron associated with tangle and plaque pathology [39].

Why iron increases with age and even more profoundly in neurodegenerative diseases is still largely unknown [8, 40]. It is hypothesized to be caused by several factors including

increased blood–brain barrier permeability and disorganization of the iron-dense myelin sheaths [10, 41, 42]. Alongside a general increase of iron in the parenchyma, iron was also shown to accumulate inside Aβ-plaques [10, 43]. Therefore, a possible hypothesis for why iron is sequestered in microglia surrounding Aβ-plaques, could be that the iron is taken up as byproduct while attempting to phagocytose the Aβ aggregates. Conversely, considering we only found approximately 25% of iron-accumulating C1-microglia to infiltrate Aβ-plaques, iron is more likely sequestered using either DMT1 or Transferrin-receptors and stored inside FTL, in an attempt to mitigate the potentially toxic effects of free iron, which in its free form is suggested to partake in Fenton's reaction to form hydroxyl radicals and cause toxic oxidate stress [42]. When iron is taken up by microglia, it first becomes part of the labile iron pool, where it can produce reactive oxygen species damaging the mitochondria and other cell organelles [44]. Studies performed using peripheral tissue cells showed the non-CNS equivalent of microglia, macrophages, to respond to intracellular iron accumulation by also activating the NLRP3 inflammasome [45]. Accordingly, in vitro and in vivo studies have shown that exposure to a combination of iron and Aβ induces the production of cytokine IL-1β and a switch to glycolytic metabolism in microglia, both of which can be interpreted as NLRP3-inflammasome activation [46, 47]. NLRP3-inflammasome activation in microglia was shown to be able to modify disease progression in two different Alzheimer mouse models [48, 49]. Our data support the in vitro and mouse model evidence that iron and Aβ can act together to accelerate disease progression via microglial inflammasome activation, by showing that in human brain tissue of Alzheimer patients, microglia are exposed to a combination iron and Aβ. Finally, these findings are also in line with recent clinical studies, in which iron was found to act as a potential disease modifier by accelerating deterioration in Alzheimer patients with high Aβ load [12, 13].

Thanks to the possibility to visualize up to six protein markers on the same section using mIF, we could better study the great heterogeneity in microglia phenotype and its spatial relationship with pathology. A limitation of mIF compared to other high-dimensional techniques such as single-cell or imaging mass cytometry is the limited number of markers available to characterize the complex microglial activation states. However, single-cell mass cytometry lacks the spatial component, which is essential when studying the relation with Aβ. Imaging mass cytometry, on the other hand, does capture the spatial distribution, however to date does not enable high-throughput analysis and offers limited resolution. Since microglia have very complicated and variable morphology, solely evaluating protein expression directly surrounding the nucleus is insufficient, and high-resolution images are required for proper segmentation and phenotyping. Secondly, as we are studying relatively rare activated microglia subtypes that will not be present in every ROI or even subject, we required high-throughput quantitative analysis methods. The mIF-mic panel, together with our optimized microglia segmentation pipeline for 2D-images, enabled accurate segmentation and analysis of >60,000 cells to carefully identify the FTL$^+$-microglia in an unbiased fashion.

In this study, we adopted an unsupervised learning approach to generate distinct clusters in our dataset, and avoid bias in the identification in clusters, as can be present in more classical IHC studies. However, as already indicated in the results section, even though distinct clusters were identified, the low degree of separation on the t-SNE mapping and similarity on the associated heatmap, suggest these clusters may be more of a continuum

rather that distinct subsets. This is in line with other transcriptomic and proteomic studies, in which they also showed the microglia clusters to be more of a continuum, even when studying substantially more genes or proteins [5, 50, 51]. However, employment of distinct clusters allows for studying the extreme ends of the continuum of the clusters to find meaningful changes in activation state. Finally, to verify that we were not looking at arbitrary differences in expression levels, we visually checked distinguishability of all independent clusters on the associated immunohistochemical images and merged clusters where this was not possible, as illustrated in Figure 1.

Future studies looking into the effect of iron and Aβ in humanized models such as iPSC-derived microglia would be extremely valuable to decipher the functional effect of this combination, and the influence of Alzheimer-associated genetic risk variants such as APOE. In addition, since microglia, as well as iron accumulation, are shown to be involved in many different neurodegenerative and neuro-immunological disease such as Parkinson's disease and multiple sclerosis, it would be worthwhile looking into this interaction as a common pathway in neurodegeneration. Like for Alzheimer disease, iron could interact with the accumulating protein of interest to affect microglia functioning and consequentially accelerate disease progression.

5.6. CONCLUSION

In summary, we showed that our multispectral immunofluorescence pipeline allowed for accurate identification of specific microglia clusters, and more importantly for the spatial analysis with respect to pathological hallmarks. In this specific study we identified dystrophic $FTL^{+}Iba1^{+}TMEM119^{-}P2RY12^{-}$-microglia to be significantly more present in Alzheimer's disease patient, and to be the predominant Aβ-plaque infiltrating microglia cluster. Finally, in correspondence with the increase of FTL-expression, $FTL^{+}Iba1^{+}$-microglia showed massive iron-loading.

The data that support the findings of this study are available from the corresponding author upon reasonable request.

5.7. ACKNOWLEDGEMENTS

We would like to thank all patients who donated their brain to the Leiden University Medical Center (LUMC), Netherlands Brain Bank (NBB) or the Normal Aging Brain collection Amsterdam (NABCA), and prof. A.J.M. Rozemuller for neuropathological evaluation of the brains. We would also like to thank I.M. Hegeman-Klein for technical assistance with histological and immunohistochemical techniques.

B.Kenkhuis is supported by an MD/PhD-grant from the Leiden University Medical Center. In addition, he has received funding from an early career fellowship from Alzheimer Nederland (WE.15-2018-13) and a Eurolife Scholarship for Early Career researcher. A.Somarakis has received funding through Leiden University Data Science Research Programme. L. van der Weerd received funding from The Netherlands Organization for Scientific Research (NWO) Innovational Research Incentives Scheme (VIDI 864.13.014).

5.8. SUPPLEMENTARY INFORMATION

Table 5.1: ***Patient demographics.*** *** Unknown pathology score. DO = Disease onset. PM = Post-mortem. LOAD = Late onset Alzheimer's disease. EOAD = Early onset Alzheimer's disease.*

	Diagnosis	Sex	DO	Age	Braak	Amyloid load	PM delay	APOE genotype	Familial genetic variant
C1	Control	F		91	2	2/3	03:47	3/3	
C2	Control	M		73	2	1	08:00	3/3	
C3	Control	M		82	**	**	05:30		
C4	Control	F		87	1	0	08:30		
C5	Control	F		72	0	0	07:15		
C6	Control	M		93	**	**	08:30		
C7	Control	M		82	1/2	1/2	07:30		
C8	Control	F		89	3	1/2	06:30	3/4	
C9	Control	F		72	1	1	06:50	3/3	
AD1	LOAD	F	80	88	6	5	04:40	3/4	
AD2	LOAD	M	69	73	5	5	04:45	4/4	
AD3	LOAD	F	>65	82	4	4	04:35	3/4	
AD4	EOAD	F	60	73	6	5	07:17	3/3	
AD5	EOAD	M	61	72	6	5	05:15	3/4	
AD6	EOAD	F	40	67	6	4	04:30	3/4	
AD7	EOAD	F	64	91	6	4/5	04:20	3/3	
AD8	EOAD	M	47	59	5	4/5	05:25	3/3	APP duplication
AD9	LOAD	F	85	90	6	4/5	03:55	2/3	
AD10	EOAD	F	51	70	6	5	04:20	4/4	
AD11	LOAD	F	87	89	4	3-4	04:30	3/3	
AD12	EOAD	F	34	43	6	5	04:15	3/3	PSEN1

Table 5.2: ***Details microglia multispectral immunofluorescence (mic-mIF) panel***

	Staining target	Antibody	Isotype	Antigen retrieval	dilution	Incubation time	Secondary Conjugate
1.	P2RY12	HPA014518, Sigma Aldrich	rIgG	10mM Citrate buffer (pH=6.0)	1:2500	2h	Poly HRP + Opal 520
2.	TMEM119	HPA051870, Sigma Aldrich	rIgG		1:250	ON	Poly HRP + Opal 570
3.	Light Chain Ferritin (FTL)	AB69090, Abcam	rIgG		1:100	ON	G--Rab Alexa 594
4.	A (17–24)	SIG-39220, Biolegend	mIgG2b		1:250	ON	Goat--mIgG2b Alexa 647
5.	Iba1	MABN92, Millipore	mIgG1		1:20	ON	G--mIgG1 CF680
6.	Nucleus, DAPI	D9542-1mg, Sigma Aldrich	n.a.		0.1 µg/ml	5 min	n.a.

Table 5.3: ***Key resources***

Product	Catalogue nr.	Supplier	Dilution
Primary antibodies			
Purified anti--Amyloid	SIG-39220	Bio legend	1:250
Anti-Iba1/AIF1	MABN92	Millipore	1:20
Anti-Ferritin Light Chain (rab)	ab69090	Abcam	1:100
Anti-Ferritin Light Chain (mIgG2a)	SC-74513	Santa Cruz	1:100
Anti-P2RY12	HPA014518	Sigma Aldrich	1:2500
Anti-TMEM119	HPA051870	Sigma Aldrich	1:250
Secondary antibodies			
Opal 520 Reagent Pack	FP1487001KT	Perkin Elmer	1:100
Opal 570 Reagent Pack	FP1488001KT	Perkin Elmer	1:100
Goat anti-Rabbit IgG (H+L), Alexa Fluor 594	A-11037	ThermoFisher	1:200
Goat anti-Mouse IgG2b, Alexa Fluor 647	A-21242	ThermoFisher	1:200
Goat anti-mouse IgG1, CF680	20253	Biotium	1:200
Goat anti-Mouse IgG (H+L), Alexa Fluor 546	A-11003	ThermoFisher	1:200
Goat anti-Rabbit IgG (H+L), Alexa Fluor 546	A-11010	ThermoFisher	1:200
Swine Anti-Rabbit Immunoglobulins/Biotin	E0353	DAKO	1:400
Rabbit anti-Mouse Immunoglobulins/Biotin	E0354	DAKO	1:200
VECTASTAIN Elite ABC-HRP Kit, Peroxidase	PK-6100	Vector Laboraties	n.a.
Extra resources			
BrightVision+ Poly- HRP-Anti Mouse/Rabbit IgG Biotin-free	VWRKDPVO110HRP	Immunologic	n.a.
1X Plus Amplificiation Diluent	FP1498A	Perkin Elmer	n.a.
ProLong Diamond Antifade mountant	P36961	ThermoFisher	n.a.
DAPI	D9542-1mg	Sigma Aldrich	0.1ug/mL
DAB sigma	D5637	Sigma Aldrich	0.5ug/mL

5.8.1. SUPPLEMENTARY METHODS

Histological staining protocol

10-µm-thick sections were used for histochemical iron detection using an enhanced Perl's reaction previously published by Van Duijn et al. [11]. Sections were incubated for 80 min in 1% potassium ferrocyanide, washed three times in 0.1 M phosphate buffer, followed by 100 min incubation in methanol with 0.01 M NaN3 and 0.3% H2O2. Subsequently, sections were washed again and incubated for 80 min in a solution containing 0.025% 3'3-diaminobenzidine-tetrahydrochloride (DAB, Sigma) and 0.005% H2O2 in 0.1 M phosphate buffer. The reaction was stopped by washing with tap water. A consecutive 10-µm-thick section was used for IHC detection of FTL. Sections were deparaffinized with xylene, washed with alcohol and blocked with 0.3% H2O2/methanol for 20 min. Subsequently, they were incubated with FTL (1:100, Santa Cruz) overnight at room temperature and after washing with PBS incubated with anti-rabbit alkaline phosphatase (1:50, Vector) for 1 h. Sections were washed with PBS and incubated with Vector Blue for 20 min in the dark. Finally, they were rinsed and covered with aqua mount.

Brightfield microscopy

Chromogenically stained slides were imaged using a Philips IntelliSite Ultra Fast Scanner (Philips, the Netherlands). Whole slide images could be viewed in 2×–40× magnification using the Philips Intellisite digital Pathology Solution system. Whole slide images were exported at 4× magnification for quantification of number of iron-positive cells. Snapshots of higher magnification were taken directly in the image-viewer.

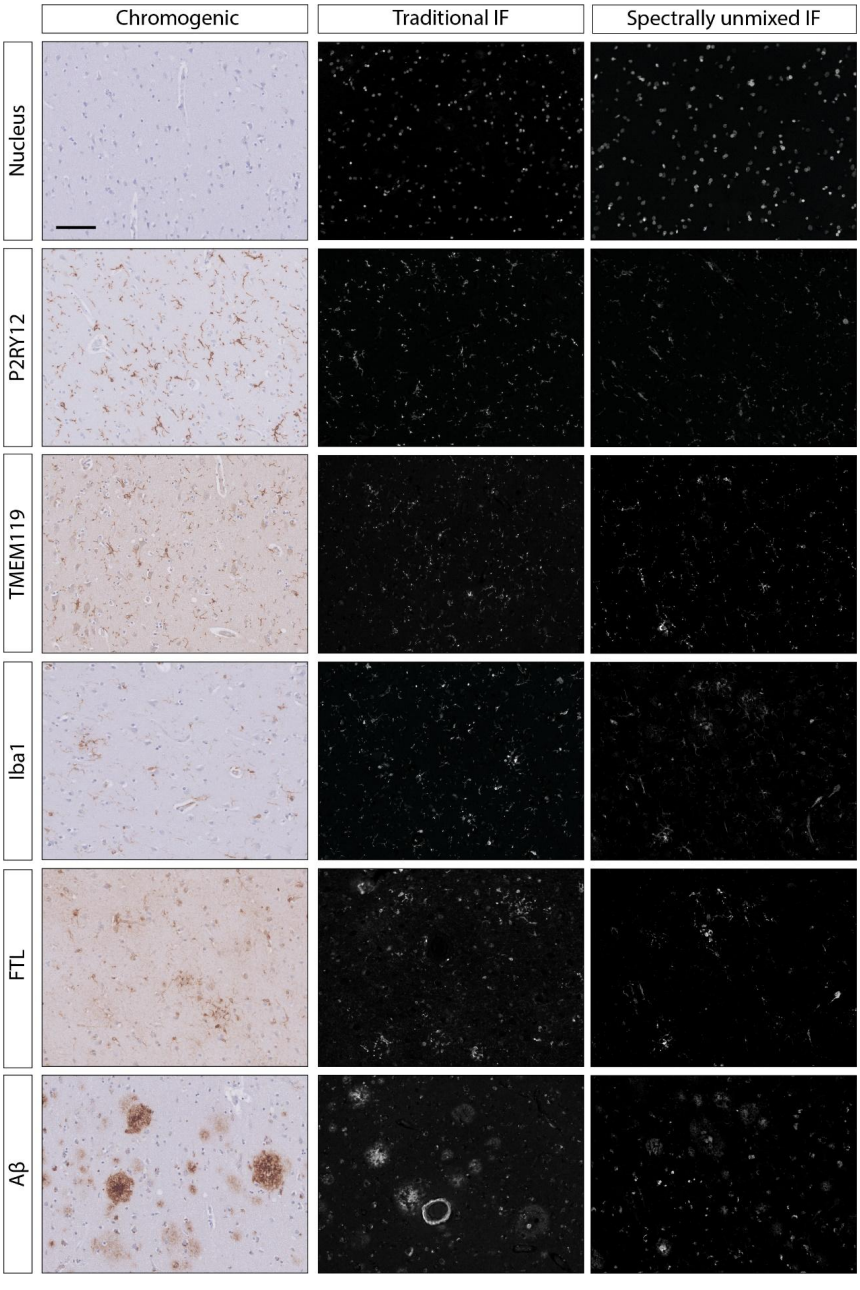

Figure 9: *Sample images of DAB-enhanced, single immunofluorescence and multiplexed immunofluorescence stainings of all antibodies used in our multispectral immunofluorescence panel. Scale bar, 100 µm.*

Antibody validation for multiplexed IHC

The six-colour mic mIF panel (Table 5.2) was created and optimized following a previously described protocol ([17]). Individual antibody conditions were optimized using single IHC and IF. Firstly, individual antibodies were tested with chromogenic and fluorescent detection for optimal antibody concentration and antigen retrieval method with the following protocol. 5-µm-thick FFPE tissue sections were deparaffinized with xylene, washed with alcohol and endogenous peroxidase was blocked by incubating in 0.3% H202 in methanol for 20 min. Antigen retrieval was performed with either citrate buffer (10 mM, pH = 6.0) or EDTA buffer (10 mM, pH 9.0). After cooling, the sections were blocked with 0.1% BSA/PBS with 0.05% Tween for 30 min and incubated with primary antibody diluted in blocking buffer in a range of concentrations overnight at room temperature. For chromogenic detection, the slides were washed with PBS and incubated with Sw-a-Rb/biotin (1:400, Dako) or Sw-a-Mouse/biotin (1:200, DAKO) for 60 min, followed by 30 min incubation with VECTAstain elite ABC. Chromogenic substrate was developed with 0.05% DAB (Sigma) with 0.005% H2O2 for 10 min. The reaction was stopped with tap water, after which the sections were counterstained with haematoxylin for 5 min and mounted with micromount. For fluorescent detection, slides were incubated with the appropriate Alexa 546 fluorophore (1:200, ThermoFisher) for 1 hour. Subsequently, the slides were incubated with 0.1 µg/mL DAPI (Sigma) for 5 min and mounted with prolong diamond (ThermoFisher). After individual antibody testing, antibodies were combined to test the viability for multiplexed immunofluorescence. The full protocol described in Section 5.3.9 was performed for 7 slides of one control and one AD subject. One slide of each was incubated with all primary antibodies. For each of the other slides the exact same protocol was performed, but only 1 out of the 6 primary antibodies was added in the primary incubation step. These slides were used to analyse the fluorescent spectrum of the individual fluorophores on our slides, which are subsequently used to un-mix the 6 different spectra. To verify specificity, the extracted signal was compared with single immunofluorescence (Figure 9).

Multispectral microscopy and image-acquisition

Mic-mIF-stained tissue slides were scanned at 4× magnification using the Vectra 3.0 Automated Quantitative Pathology Imaging system (PerkinElmer). Following whole slide scanning, a 50% and 25% ROI grid was placed on the cortex and white matter, respectively. For each ROI, high-resolution 20× magnification images are obtained of all subjects. Spectral separation of the 6 individual dyes was performed automatically using InForm Cell Analysis software (PerkinElmer), using spectral libraries obtained with single-marker IF detection of the different fluorophores. In total, six different raw component images of the extracted spectra were exported from Inform for further analysis.

Confocal microscopy

20 µm sections were imaged using an Andor Dragonfly 200 spinning disk confocal system (Andor, Oxford Instruments). Sections were stained with DAPI, Alexa 488, Alexa 546 and Alexa 647, which were imaged with a 405 nm, 488 nm, 561 nm and 637 nm laser respectively. Subsequently images were exported and processed using Imaris (Bitplane, Oxford instruments). First, a Gaussian filter was applied to all channels, after which snapshots of the reconstructed 3D projection were taken.

REFERENCES

[1] I. E. Jansen, J. E. Savage, K. Watanabe, J. Bryois, D. M. Williams, and S. Steinberg, "Genome-wide meta-analysis identifies new loci and functional pathways influencing alzheimer's disease risk," *Nat Genet*, vol. 51, 2019.

[2] J. C. Lambert, C. A. Ibrahim-Verbaas, D. Harold, A. C. Naj, R. Sims, and C. Bellenguez, "Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for alzheimer's disease," *Nat Genet*, vol. 45, 2013.

[3] H. Keren-Shaul, A. Spinrad, A. Weiner, O. Matcovitch-Natan, R. Dvir-Szternfeld, and T. K. Ulland, "A unique microglia type associated with restricting development of alzheimer's disease," *Cell*, vol. 169, 2017.

[4] H. Mathys, J. Davila-Velderrain, Z. Peng, F. Gao, S. Mohammadi, and J. Z. Young, "Single-cell transcriptomic analysis of alzheimer's disease," *Nature*, vol. 570, 2019.

[5] Y. Zhou, W. M. Song, P. S. Andhey, A. Swain, T. Levy, and K. R. Miller, "Human and mouse single-nucleus transcriptomics reveal trem2-dependent and trem2-independent cellular responses in alzheimer's disease," *Nat Med*, vol. 26, 2020.

[6] K. O. Lopes, D. L. Sparks, and W. J. Streit, "Microglial dystrophy in the aged and alzheimer's disease brain is associated with ferritin immunoreactivity," *Glia*, vol. 56, 2008.

[7] I. Grundke-Iqbal, J. Fleming, Y. C. Tung, H. Lassmann, K. Iqbal, and J. G. Joshi, "Ferritin is a component of the neuritic (senile) plaque in alzheimer dementia," *Acta Neuropathologica*, vol. 81, 1990.

[8] A. Damulina, L. Pirpamer, M. Soellradl, M. Sackl, C. Tinauer, and E. Hofer, "Cross-sectional and longitudinal assessment of brain iron level in alzheimer disease using 3-t mri," *Radiology*, vol. 296, 2020.

[9] N. Spotorno, J. Acosta-Cabronero, E. Stomrud, B. Lampinen, O. T. Strandberg, and D. Westen, "Relationship between cortical iron and tau aggregation in alzheimer's disease," *Brain*, vol. 143, 2020.

[10] M. Bulk, B. Kenkhuis, L. M. Graaf, J. J. Goeman, R. Natté, and L. Weerd, "Post-mortem t2*- weighted mri imaging of cortical iron reflects severity of alzheimer's disease," *J Alzheimer's Dis*, vol. 65, 2018.

[11] S. Duijn, M. Bulk, S. G. Duinen, R. J. A. Nabuurs, M. A. Buchem, and L. Weerd, "Cortical iron reflects severity of alzheimer's disease," *J Alzheimer's Dis*, vol. 60, 2017.

[12] S. Ayton, A. Fazlollahi, P. Bourgeat, P. Raniga, A. Ng, and Y. Y. Lim, "Cerebral quantitative susceptibility mapping predicts amyloid--related cognitive decline," *Brain*, vol. 140, 2017.

[13] S. Ayton, Y. Wang, I. Diouf, J. A. Schneider, J. Brockman, and M. C. Morris, "Brain iron is associated with accelerated cognitive decline in people with alzheimer pathology," *Mol Psychiatry*, vol. 66, 2019.

[14] H. Braak and E. Braak, "Neuropathological stageing of alzheimer-related changes," *Acta Neuropathologica*, vol. 66, 1991.

[15] H. Braak and E. Braak, "Staging of alzheimer's disease-related neurofibrillary changes," *Neurobiol Aging Neurobiol Aging*, vol. 16, 1995.

[16] B. T. Hyman, C. H. Phelps, T. G. Beach, E. H. Bigio, N. J. Cairns, and M. C. Carrillo, "National institute on aging-alzheimer's association guidelines for the neuropathologic assessment of alzheimer's disease," *Alzheimer's Dement NIH Public Access*, vol. 8, 2012.

[17] M. E. Ijsselsteijn, T. P. Brouwer, Z. Abdulrahman, E. Reidy, A. Ramalheiro, A. M. Heeren, A. Vahrmeijer, E. S. Jordanova, and N. F. de Miranda, "Cancer immunophenotyping by seven-colour multispectral imaging without tyramide signal amplification," *The Journal of Pathology: Clinical Research*, vol. 5, pp. 3–11, 2019.

[18] J. Levine, E. Simonds, S. Bendall, K. Davis, E. ad D. Amir, M. Tadmor, O. Litvin, H. Fienberg, A. Jager, E. Zunder, R. Finck, A. Gedman, I. Radtke, J. Downing, D. Pe'er, and G. Nolan, "Data-driven phenotypic dissection of aml reveals progenitor-like cells that correlate with prognosis," *Cell*, vol. 162, no. 1, pp. 184–197, 2015.

[19] J. L. Hintze and R. D. Nelson, "Violin plots: A box plot-density trace synergism," *The American Statistician*, vol. 52, no. 2, pp. 181–184, 1998.

[20] L. v. d. Maaten and G. Hinton, "Visualizing data using t-sne," *Journal of machine learning research*, vol. 9, no. Nov, pp. 2579–2605, 2008.

[21] T. Höllt, N. Pezzotti, V. van Unen, F. Koning, E. Eisemann, B. P. F. Lelieveldt, and A. Vilanova, "Cytosplore: Interactive immune cell phenotyping for large single-cell datasets," *Computer Graphics Forum (Proceedings of EuroVis)*, vol. 35, no. 3, pp. 171–180, 2016.

[22] A. Somarakis, M. E. Ijsselsteijn, S. J. Luk, B. Kenkhuis, N. F. C. C. de Miranda, B. P. F. Lelieveldt, and T. Höllt, "Visual cohort comparison for spatial single-cell omics-data," *IEEE Transactions on Visualization and Computer Graphics*, vol. 27, no. 2, pp. 733–743, 2021.

[23] A. Somarakis, V. van Unen, F. Koning, B. P. Lelieveldt, and T. Höllt, "ImaCytE: visual exploration of cellular microenvironments for imaging mass cytometry data," *IEEE Transactions on Visualization and Computer Graphics*, vol. 27, no. 1, pp. 98–110, 2021.

[24] M. Meadowcroft, J. Connor, and Q. Yang, "Cortical iron regulation and inflammatory response in alzheimer's disease and appswe/ps1e9 mice: a histological perspective," *Frontiers in Neuroscience*, vol. 9, p. 255.

[25] P. Banerjee, E. Paza, E. M. Perkins, O. G. James, B. Kenkhuis, A. F. Lloyd, K. Burr, D. Story, D. Yusuf, X. He, R. Backofen, O. Dando, S. Chandran, and J. Priller, "Generation of pure monocultures of human microglia-like cells from induced pluripotent stem cells," *Stem Cell Research*, vol. 49, p. 102046, 2020.

[26] T. A. van Wageningen, E. Vlaar, G. Kooij, C. A. M. Jongenelen, J. J. G. Geurts, and A. M. van Dam, "Regulation of microglial TMEM119 and P2RY12 immunoreactivity in multiple sclerosis white and grey matter lesions is dependent on their inflammatory environment," *Acta Neuropathologica Communications*, vol. 7, no. 1, p. 206, 2019.

[27] M. L. Bennett, F. C. Bennett, S. A. Liddelow, B. Ajami, J. L. Zamanian, N. B. Fernhoff, S. B. Mulinyawe, C. J. Bohlen, A. Adil, A. Tucker, I. L. Weissman, E. F. Chang, G. Li, G. A. Grant, M. G. Hayden Gephart, and B. A. Barres, "New tools for studying microglia in the mouse and human cns," *Proceedings of the National Academy of Sciences*, vol. 113, no. 12, pp. E1738–E1746, 2016.

[28] A. Mildner, H. Huang, J. Radke, W. Stenzel, and J. Priller, "P2y 12 receptor is expressed on human microglia under physiological conditions throughout development and is sensitive to neuroinflammatory diseases," *Glia*, vol. 65, pp. 375–387.

[29] T. Zrzavy, S. Hametner, I. Wimmer, O. Butovsky, H. L. Weiner, and H. Lassmann, "Loss of 'homeostatic' microglia and patterns of their activation in active multiple sclerosis," *Brain*, vol. 140, no. 7, pp. 1900–1913, 2017.

[30] S. Ayton, N. G. Faux, A. I. Bush, and A. D. N. Initiative, "Ferritin levels in the cerebrospinal fluid predict alzheimer's disease outcomes and are regulated by apoe," *Nature Communications*, vol. 6, no. 1, p. 6760, 2015.

[31] W. J. Streit, Q. S. Xue, J. Tischer, and I. Bechmann, "Microglial pathology," *Acta Neuropathologica Communications*, vol. 2, p. 142, 2014.

[32] J. R. Connor, S. L. Menzies, S. M. St. Martin, and E. J. Mufson, "A histochemical study of iron, transferrin, and ferritin in alzheimer's diseased brains," *Journal of Neuroscience Research*, vol. 31, no. 1, pp. 75–83, 1992.

[33] M. M. Zeineh, Y. Chen, H. H. Kitzler, R. Hammond, H. Vogel, and B. K. Rutt, "Activated iron-containing microglia in the human hippocampus identified by magnetic resonance imaging in alzheimer disease," *Neurobiol Aging*, vol. 36, pp. 2483–2500, 2015.

[34] D. G. Walker, T. M. Tang, A. Mendsaikhan, I. Tooyama, G. E. Serrano, L. I. Sue, T. G. Beach, and L.-F. Lue, "Patterns of Expression of Purinergic Receptor P2RY12, a Putative Marker for Non-Activated Microglia, in Aged and Alzheimer's Disease Brains," vol. 21, no. 2, 2020.

[35] Y. Kaneko, T. Kitamoto, J. Tateishi, and K. Yamaguchi, "Ferritin immunohistochemistry as a marker for microglia," *Acta Neuropathologica*, vol. 79, no. 2, pp. 129–136, 1989.

[36] K. Jellinger, W. Paulus, I. Grundke-Iqbal, P. Riederer, and M. B. H. Youdim, "Brain iron and ferritin in parkinson's and alzheimer's diseases," *Journal of Neural Transmission - Parkinson's Disease and Dementia Section*, vol. 2, no. 4, pp. 327–340, 1990.

[37] A. T. Nguyen, K. Wang, G. Hu, X. Wang, Z. Miao, J. A. Azevedo, E. Suh, V. M. Van Deerlin, D. Choi, K. Roeder, M. Li, and E. B. Lee, "Apoe and trem2 regulate amyloid-responsive microglia in alzheimer's disease," *Acta Neuropathologica*, vol. 140, no. 4, pp. 477–493, 2020.

[38] W. J. Streit, H. Braak, K. Del Tredici, J. Leyh, J. Lier, H. Khoshbouei, C. Eisenlöffel, W. Müller, and I. Bechmann, "Microglial activation occurs late during preclinical alzheimer's disease," *Glia*, vol. 66, no. 12, pp. 2550–2562, 2018.

[39] S. Ayton, S. Janelidze, B. Roberts, S. Palmqvist, P. Kalinowski, I. Diouf, A. A. Belaidi, E. Stomrud, A. I. Bush, and O. Hansson, "Acute phase markers in csf reveal inflammatory changes in alzheimer's disease that intersect with pathology, apoe 4, sex and age," *Progress in Neurobiology*, vol. 198, p. 101904, 2021.

[40] J. Acosta-Cabronero, M. J. Betts, A. Cardenas-Blanco, S. Yang, and P. J. Nestor, "In vivo mri mapping of brain iron deposition across the adult lifespan," *Journal of Neuroscience*, vol. 36, no. 2, pp. 364–374, 2016.

[41] A. J. Farrall and J. M. Wardlaw, "Blood–brain barrier: ageing and microvascular disease—systematic review and meta-analysis," *Neurobiol Aging*, vol. 25, pp. 337–352, 2009.

[42] R. J. Ward, F. A. Zucca, J. H. Duyn, R. R. Crichton, and L. Zecca, "The role of iron in brain ageing and neurodegenerative disorders," *The Lancet Neurology*, vol. 13, no. 10, pp. 1045–1060, 2014.

[43] R. J. Nabuurs, I. Hegeman, R. Natté, S. G. van Duinen, M. A. van Buchem, L. van der Weerd, and A. G. Webb, "High-field mri of single histological slices using an inductively coupled, self-resonant microcoil: application to ex vivo samples of patients with alzheimer's disease," *NMR in Biomedicine*, vol. 24, no. 4, pp. 351–357, 2011.

[44] S. J. Dixon and B. R. Stockwell, "The role of iron and reactive oxygen species in cell death," *Nature Chemical Biology*, vol. 10, no. 1, pp. 9–17, 2014.

[45] K. Nakamura, T. Kawakami, N. Yamamoto, M. Tomizawa, T. Fujiwara, T. Ishii, H. Harigae, and K. Ogasawara, "Activation of the nlrp3 inflammasome by cellular labile iron," *Experimental Hematology*, vol. 44, no. 2, pp. 116–124, 2016.

[46] A. McIntosh, V. Mela, C. Harty, A. M. Minogue, D. A. Costello, C. Kerskens, and M. A. Lynch, "Iron accumulation in microglia triggers a cascade of events that leads to altered metabolism and compromised function in app/ps1 mice," *Brain Pathology*, vol. 29, no. 5, pp. 606–621, 2019.

[47] I. C. Nnah, C.-H. Lee, and M. Wessling-Resnick, "Iron potentiates microglial interleukin-1β secretion induced by amyloid-β," *Journal of neurochemistry*, vol. 154, no. 2, pp. 177–189, 2020.

[48] C. Ising, C. Venegas, S. Zhang, H. Scheiblich, S. V. Schmidt, A. Vieira-Saecker, S. Schwartz, S. Albasset, R. M. McManus, D. Tejera, A. Griep, F. Santarelli, F. Brosseron, S. Opitz, J. Stunden, M. Merten, R. Kayed, D. T. Golenbock, D. Blum, E. Latz, L. Buée, and M. T. Heneka, "Nlrp3 inflammasome activation drives tau pathology," *Nature*, vol. 575, no. 7784, pp. 669–673, 2019.

[49] M. T. Heneka, M. P. Kummer, A. Stutz, A. Delekate, S. Schwartz, A. Vieira-Saecker, A. Griep, D. Axt, A. Remus, T.-C. Tzeng, E. Gelpi, A. Halle, M. Korte, E. Latz, and D. T. Golenbock, "Nlrp3 is activated in alzheimer's disease and contributes to pathology in app/ps1 mice," *Nature*, vol. 493, no. 7434, pp. 674–678, 2013.

[50] R. Sankowski, C. Böttcher, T. Masuda, L. Geirsdottir, E. Sindram, T. Seredenina, A. Muhs, C. Scheiwe, M. J. Shah, D. H. Heiland, *et al.*, "Mapping microglia states in the human brain through the integration of high-dimensional techniques," *Nature neuroscience*, vol. 22, no. 12, pp. 2098–2110, 2019.

[51] C. Böttcher, S. Schlickeiser, M. A. M. Sneeboer, D. Kunkel, A. Knop, and E. Paza, "Human microglia regional heterogeneity and phenotypes determined by multiplexed single-cell mass cytometry," *Nat Neurosci*, vol. 22, 2019.

6

Discussion

6.1. DISCUSSION

In each chapter of this thesis, we discussed in detail possible limitations and future directions specifically for each part of our end-to-end pipeline. Here, we will place this discussion in a wider context, and reflect on the overall strengths and weaknesses of our work to fully understand the possibilities and limitations and propose directions for future research.

As we have defined in Chapter 1, the cell is the basic unit of life. Therefore, the whole pipeline is based on the exploration of cellular patterns and characteristics. Furthermore, the high-resolution of the raw, highly multiplexed images offers information about the extracellular matrix and the distribution of molecular characteristics inside the cell. All this information is omitted in favor of the cellular analysis, introducing a limitation of our work. A "cell-free" analysis approach could replace the error-prone cell segmentation step and add information to the analysis pipeline. On the other hand, as the amount of high-dimensional cellular data extracted from an image is two to three orders of magnitude smaller than its contained high-dimensional pixel values, it would cause scalability issues. Fortunately, current developments [1] have enabled the interactive exploration of millions of data points enabling the possibility of a "pixel-based" analysis, circumventing the need for cell segmentation [2]. An adaptation of our proposed Visual Analytics methods for the inclusion of "pixel-based" analysis in the pipeline is an interesting direction for further development.

A factor that limits the broad applicability of our method is related to the novelty and consequently shortage of (annotated) data. The well-defined nature of the preprocessing problems make them ideal candidates for the efficient machine learning algorithms, given there are plenty of annotated data. With the current boom in the use of spatial-omics modalities [3] a huge volume of annotated data, ideal for the efficient development of fully-supervised solutions, is expected. Such solutions could increase the generality and efficiency of the presented preprocessing algorithms.

For the detailed in-depth exploration of the highly multiplexed cellular images, the cellular microenvironment characteristics are crucial. The exploration of the cellular microenvironment, though, requires first its definition. Before starting exploring the cellular microenvironment characteristics, in both Visual Analytics tools; ImaCytE and SpaCeCo, the microenvironment is defined. A cell's microenvironment is composed of cells that are below a specific distance. This way each cell of a cell's microenvironment, regardless of its distance, contributes the same in its microenvironment. However, the biological significance of a cell to its neighbor varies regarding their distance. Therefore, developing a weighted, or a multi-level approach of the algorithms proposed in this thesis may be more suitable for the exploration of cellular microenvironments.

Our visual analytics tools enable experts to generate hypotheses for a tissue's functionality. Mere hypotheses are not enough for the experts draw safe conclusions. Therefore, integrating hypotheses testing methods into the pipeline would facilitate the interpretation of our observation and relate them to specific clinical information.

The introduction of novel visual designs for the exploration of cellular characteristics, such as the glyphs for the motifs representation in ImaCytE and the visual query system in SpaCeCo, is challenging. Their perception and usefulness follows a steep learning

curve and is highly correlated with the complexity of the analyzed data. Hence, it is important for the designer to provide to the users continuous support and explanation of the design concepts and be willing to "transfer" the design concepts according to the needs of the experts. A characteristic example is the utilization of the glyphs designed for the representation of the motifs for the representation of Abeta plaques microenvironment on the tissue image.

The above example highlights a second challenge that we should deal with during visual design process, the discrimination of visualization design and visual data analysis. A difficult task to be accomplished in the context of a PhD project, where close and continuous collaboration with clinical researchers is part of the daily routine. However, during the phase of visual design it is important not to slip under the pressure of experts for the analysis of specific datasets and abstract as much as possible the design solutions. To deal with this challenge and meet both the imminent and more generic needs of our collaborators, we developed two pure Visual Analytics projects (ImaCytE and SpaCeCo), where the exploration of spatial omics data can take place and contributed equally with our clinical collaborators to two other projects related with the analysis of specific data(sets).

In general, multiplexed cellular images have not been utilized to their full potential. Until now, experts analyze such data solely to identify the contained cell types or to observe in a qualitative manner the proximity of two cellular types. A detailed quantitative analysis of cellular microenvironments for the extraction of important information about tissue functionality is still missing, but it is highly anticipated [2, 4] and the development of high quality open-source software applications is necessary for developments in this direction. To that end, the daily usage of the open-source tools developed for our holistic pipeline by our collaborators in LUMC and other research labs [5] indicate that our work has been established as an important tool for the generation of significant biological hypotheses from the analysis of multiplexed cellular images. Moreover, the awarding of the third Dirk Bartz Prize for Visual Computing in Medicine 2021 and the honourable mention for the Karl-Heinz Höhne Award for Medical Visualization 2021 indicate the recognition of our solutions for the Medical Visualization community as well.

REFERENCES

[1] V. Unen, T. Höllt, N. Pezzotti, N. Li, M. J. Reinders, E. Eisemann, F. Koning, A. Vilanova, and B. P. Lelieveldt, "Visual analysis of mass cytometry data by hierarchical stochastic neighbour embedding reveals rare cell types," *Nature communications*, vol. 8, no. 1, p. 1740, 2017.

[2] N. L. de Vries, A. Mahfouz, F. Koning, and N. F. C. C. de Miranda, "Unraveling the complexity of the cancer microenvironment with multidimensional genomic and cytometric technologies," *Frontiers in Oncology*, vol. 10, p. 1254, 2020.

[3] V. Marx, "Method of the year: spatially resolved transcriptomics," *Nature Methods*, vol. 18, no. 1, pp. 9–14, 2021.

[4] N. Guo, V. van Unen, M. E. Ijsselsteijn, L. F. Ouboter, A. E. van der Meulen, S. M. Chuva de Sousa Lopes, N. F. C. C. de Miranda, F. Koning, and N. Li, "A 34-marker panel for imaging mass cytometric analysis of human snap-frozen tissue," *Frontiers in Immunology*, vol. 11, p. 1466, 2020.

[5] J. O. Garcia, "Data analysis at Newcastle Unicersity flow cytometry core facility." `https://www.ncl.ac.uk/fccf/research-development/data-analysis/`, 2021. [Online; Accessed: 2021-02-05].

Nederlandse Samenvatting

Een dieper inzicht in de pathologie van een organisme is belangrijk voor de ontwikkeling van behandelingen. In de loop van eeuwen van systematisch onderzoek hebben klinische onderzoekers aangetoond dat hoe meer informatie zij verwerven over de celeigenschappen en hun organisatie in het weefsel, hoe beter zij de functionaliteit van een organisme en de ziekteprogressie kunnen begrijpen. De laatste jaren heeft de komst van hoge-resolutie beeldvormingstechnieken onderzoekers voorzien van nieuwe informatie over individuele cellen, die onderzoekers in staat stellen om de cellen nauwkeurig te karakteriseren en te onderzoeken hoe ze zijn verdeeld in het weefsel. Echter, de extractie van bruikbare biologische inzichten uit de analyse van zulke nieuwe en complexe data, waar experts de intrinsieke karakteristieken van de data niet kennen noch de patronen die ze willen identificeren, vereist een exploratieve data-analyse aanpak. Het doel van dit proefschrift is dan ook de ontwikkeling van een end-to-end pijplijn voor de analyse van deze hoogdimensionale cellulaire beelden; van de voorbewerking van de ruwe data tot de exploratie van cellulaire patronen en hun associatie met klinische karakteristieken. In Hoofdstuk 1 motiveren we het belang en de noodzaak van onze studie, presenteren we de belangrijkste kenmerken van de data, bespreken we de voorbewerkingsalgoritmen en introduceren we de visual analytics benadering voor de exploratieve analyse van complexe data. Vervolgens beschrijven we een nieuwe methodologie voor de voorbewerking van deze hoogdimensionalecellulaire beelden (Hoofdstuk 2). In Hoofdstuk 2.1 beschrijven we de semi-supervised workflow die we hebben ontwikkeld voor de normalisatie van Imaging Mass Cytometry data tussen verschillende monsters en dimensies. Daarnaast laten we zien hoe de binaire classificatie van een beeldpixel in voorgrond en achtergrond niet-biologische variatie tussen weefselmonsters kan elimineren en de traceerbaarheid van cellen met laag-geexpresseerde eiwitniveaus kan verbeteren. In hoofdstuk 2.2 presenteren we een segmentatie-algoritme, gericht op de precieze identificatie van microglia, zoals deze is vastgelegd met de Vectra multiplexed immunofluorescentie modaliteit. In het bijzonder wordt een combinatie van traditionele segmentatie-algoritmen (Level-Set, Watershed) en morfologische operaties gebruikt om een geautomatiseerd algoritme te creëren voor de precieze identificatie van microglia's cellulaire grenzen. In de Hoofdstukken 3 en 4 stellen we een Visual Analytics aanpak voor voor de analyse van deze zeer complexe ruimtelijke cellulaire gegevens, die twee data-gedreven interactieve tools omvat, ImaCytE en SpaCeCo. ImaCytE (Hoofdstuk 3) stelt de experts in staat om meerdere weefselmonsters te onderzoeken en nieuwe celtypen te identificeren, om op meerdere niveaus de ruimtelijke patronen die zij vormen te onderzoeken en om specifieke celtypen te stratificeren op basis van hun micro-omgeving kenmerken. SpaCeCo (Hoofdstuk 4) is een aanvulling op ImaCytE en stelt deskundigen in staat om te bepalen welke cellulaire kenmerken twee cohorten monsters met verschillende klinische kenmerken van elkaar onderscheiden. De workflow voor de vergelijking is verdeeld in twee stappen. De eerste stap omvat de vergelijking van cohorten op basis van de abundantie van de verschillende celtypes en de tweede stap op basis van de ruimtelijke pa-

tronen die de verschillende celtypes vormen. De vergelijking vindt plaats op twee niveaus, het cohort- en het monsterniveau, zodat de deskundige in staat is eventuele uitschieters in elk cohort te identificeren. Zowel in ImaCytE als in SpaCeCo kan de gebruiker elke bevinding op het weefsel lokaliseren om ze te verifieren en in haar biologische context te plaatsen. In Hoofdstuk 5 laten we zien hoe delen van onze pijplijn zijn gebruikt voor een studie naar de ziekte van Alzheimer. Het segmentatie algoritme voor de efficiënte identificatie van microglia, de motief glyphs, voor het eerst gepresenteerd in ImaCytE, voor de exploratie van cellulaire patronen in het weefsel en SpaCeCo voor de vergelijking van Alzheimer patiënten met gezonde individuen. Tot slot hebben we Visual Analytics oplossingen ontwikkeld om de experts in staat te stellen diepgaande ruimtelijke cellulaire gegevens te verkennen en hypotheses te genereren voor de oorsprong van de functionaliteit van het weefsel in een zieke of een gezonde toestand. Naast de kerntaken op het gebied van exploratie, stellen we de experts in staat om hun gegevens te pre-processen, waardoor een end-to-end pijplijn ontstaat voor de analyse van multiplexed weefselbeelden op cellulaire resolutie. Bij elkaar beschouwd vormt het werk in dit proefschrift de basis voor een Visual Analytics-benadering voor de analyse van hoogdimensionale beeldvorming om cellen en weefsels in groot detail te kunnen bestuderen.

List of Publications

Journal Publications

A. Somarakis, M. E. Ijsselsteijn, B. Kenkhuis, V. v. Unen, S. J. Luk, F. Koning, L. v. d. Weerd, N. F. C. C. d. Miranda, B. P. F. Lelieveldt, and T. Höllt, "Visual Analysis of Tissue Images at Cellular Level," in *EuroVis 2021 - Dirk Bartz Prize*, The Eurographics Association, 2021. doi: 10.2312/evm.20211074

M. E. Ijsselsteijn*, **A. Somarakis***, B. Lelieveldt, T. Höllt[+], and N. F. de Miranda[+], "Semi-automated background removal limits data loss and normalizes imaging mass cytometry data," *Cytometry Part A*, vol. 99, no. 12, pp. 1187–1197, 2021. doi: 10.1002/cyto.a.24480

B. Kenkhuis*, **A. Somarakis***, L. de Haan, O. Dzyubachyk, M. E. IJsselsteijn, N. F. de Miranda, B. P. Lelieveldt, J. Dijkstra, W. M. van Roon-Mom, T. Höllt, and L. van der Weerd, "Iron loading is a prominent feature of activated microglia in alzheimer's disease patients," *Acta neuropathologica communications*, vol. 9, no. 1, pp. 1–15, 2021. doi: 10.1186/s40478-021-01126-5

A. Somarakis, M. E. Ijsselsteijn, S. J. Luk, B. Kenkhuis, N. F. C. C. de Miranda, B. P. F. Lelieveldt, and T. Höllt, "Visual cohort comparison for spatial single-cell omics-data," *IEEE Transactions on Visualization and Computer Graphics*, vol. 27, no. 2, pp. 733–743, 2021. doi: 10.1109/tvcg.2020.3030336

A. Somarakis, V. V. Unen, F. Koning, B. P. Lelieveldt, and T. Höllt, "ImaCytE: visual exploration of cellular microenvironments for imaging mass cytometry data," *IEEE Transactions on Visualization and Computer Graphics*, vol. 27, no. 1, pp. 98–110, 2021. doi: 10.1109/tvcg.2019.2931299

Z. Abdulrahman, S. Santegoets, G. Sturm, P. Charoentong, M. Ijsselsteijn, **A. Somarakis**, T. Höllt, F. Finotello, Z. Trajanoski, S. v. Egmond, D. Mustafa, M. Welters, N. d. Miranda, and S. v. d. Burg, "35 Chemokine-driven spatial organization of immune cell microaggregates marks oropharyngeal squamous cell carcinomas containing tumor-specific T cells," *Journal for ImmunoTherapy of Cancer*, vol. 9, no. Suppl 2, pp. A41–A41, 2021. doi: 10.1136/jitc-2021-SITC2021.035

B. Kenkhuis, **A. Somarakis**, L. R. Kleindouwel, W. M. van Roon-Mom, T. Höllt, and L. van der Weerd, "Co-expression patterns of microglia markers Iba1, TMEM119 and P2RY12 in Alzheimer's disease," *bioRxiv*, 2021. doi: 10.1101/2021.05.31.446375

---

*,[+] Contributed equally

D. Lähnemann, J. Köster, E. Szczurek, D. J. McCarthy, S. C. Hicks, M. D. Robinson, C. A. Vallejos, K. R. Campbell, N. Beerenwinkel, A. Mahfouz, L. Pinello, P. Skums, A. Stamatakis, C. S. O. Attolini, S. Aparicio, J. Baaijens, M. Balvert, B. de Barbanson, A. Cappuccio, G. Corleone, B. E. Dutilh, M. Florescu, V. Guryev, R. Holmer, K. Jahn, T. J. Lobo, E. M. Keizer, I. Khatri, S. M. Kielbasa, J. O. Korbel, A. M. Kozlov, T. H. Kuo, B. P. Lelieveldt, I. I. Mandoiu, J. C. Marioni, T. Marschall, F. Mölder, A. Niknejad, L. Raczkowski, M. Reinders, J. de Ridder, A. E. Saliba, **A. Somarakis**, O. Stegle, F. J. Theis, H. Yang, A. Zelikovsky, A. C. McHardy, B. J. Raphael, S. P. Shah, and A. Schönhuth, "Eleven grand challenges in single-cell data science," *Genome Biology*, vol. 21, 2020. doi: 10.1186/s13059-020-1926-6

N. Li, V. van Unen, N. Guo, T. Abdelaal, **A. Somarakis**, J. Eggermont, A. Mahfouz, S. M. C. de Sousa Lopes, B. P. Lelieveldt, and F. Koning, "Early-life compartmentalization of immune cells in human fetal tissues revealed by high-dimensional mass cytometry," *Frontiers in Immunology*, vol. 10, 2019. doi: 10.3389/fimmu.2019.01932

Awards

Third Prize for the Dirk Bartz Award for Visual Computing in Medicine 2021

Honorable mention for the Karl Heinz Höhne Award for Medical Visualization 2021

Curriculum Vitæ

Antonios Somarakis was born in Heraklion, Crete at the 11th of November in 1992. After graduating from the 2nd General High School of Heraklion, he studied Electrical and Computer Engineering in the National Technical University of Athens. In April 2016, he obtained his 5-year Diploma with a major in Signal Processing and minor in Biomedical Engineering. He conducted his Diploma Thesis on the compression of ECG signal in collaboration with Erasmus University Medical Center.

After his introduction to the Dutch Academia, he started his PhD project in February 2017 on the visual analytics for spatial omics data under the supervision of Boudewijn Lelievledt and Thomas Höllt in Leiden University Medical Center (LUMC). During the PhD project, he focused, in close collaboration with clinical researchers from LUMC, on problems related to the analysis of highly multiplexed cellular images.

In parallel with his PhD work, he remained socially active being, among others, member of the Leo Leiden PhD association and TU Delft Ambassadors.

Acknowledgements

First of all, I would like to express my gratitude towards Boudewijn Lelieveldt and Thomas Höllt. Dear Boudewijn, thank you for believing in me and providing me with the opportunity to start this PhD project. Your research enthusiasm and support all these years was the ideal context for conducting high-end research. Dear Thomas, you are for me a paradigm of scientific excellence and integrity. Your scientific perfectionism shaped, to a great extent, my research profile, thank you. Also, I am more than grateful having Fons Verbeek as my copromotor.

I would like to thank all the members of LKEB family for our daily associations, especially my roommates in our "Crete"-office- island and my collaborators in LUMC, Marieke, Boyd, Sietse, Ziena, Vincent, Juliette, Li Na, Nannan, Noel, Frits, Louise. Special thanks to my students, Jan and Ismigul. They taught more than I did.

Moreover, being part of the DSO community was a blessing. The scientific plurality of projects broadened my perception on the usefulness of Data Science for the promotion of scientific research.

Finally, I would like to thank those that mattered most during this journey. The friends that I met in the Netherlands, those that continue to honour me with their friendship even they were far away and of course, the cornerstone of all, my family.