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

2

PREPROCESSING OF HIGHLY MULTIPLEXED CELLULAR IMAGES

2.1. Normalization strategies for Imaging Mass Cytometry data

This chapter was adapted from:

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

2

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
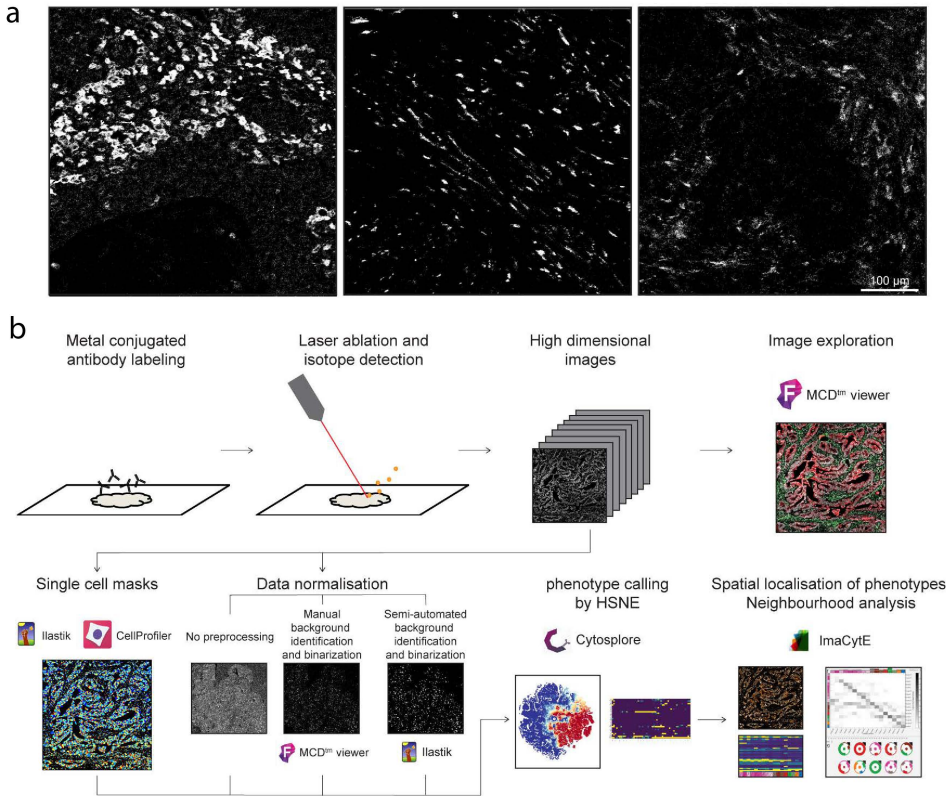
Figure 2.1: *(a) CD163 expression patterns in three samples. A signal range of 3-20 mean dual counts was set for all images in the MCD viewer, but differences in signal intensity and background are observed between images. (b) Workflow for IMC FFPE imaging and data analysis, including the three tested 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-
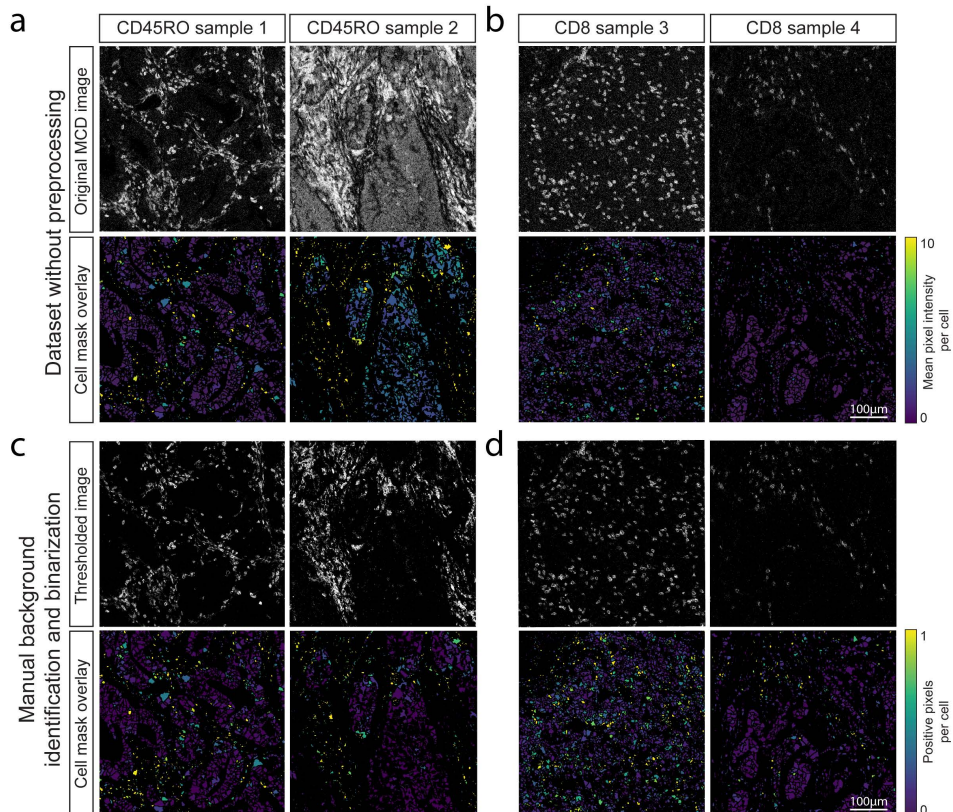
2

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
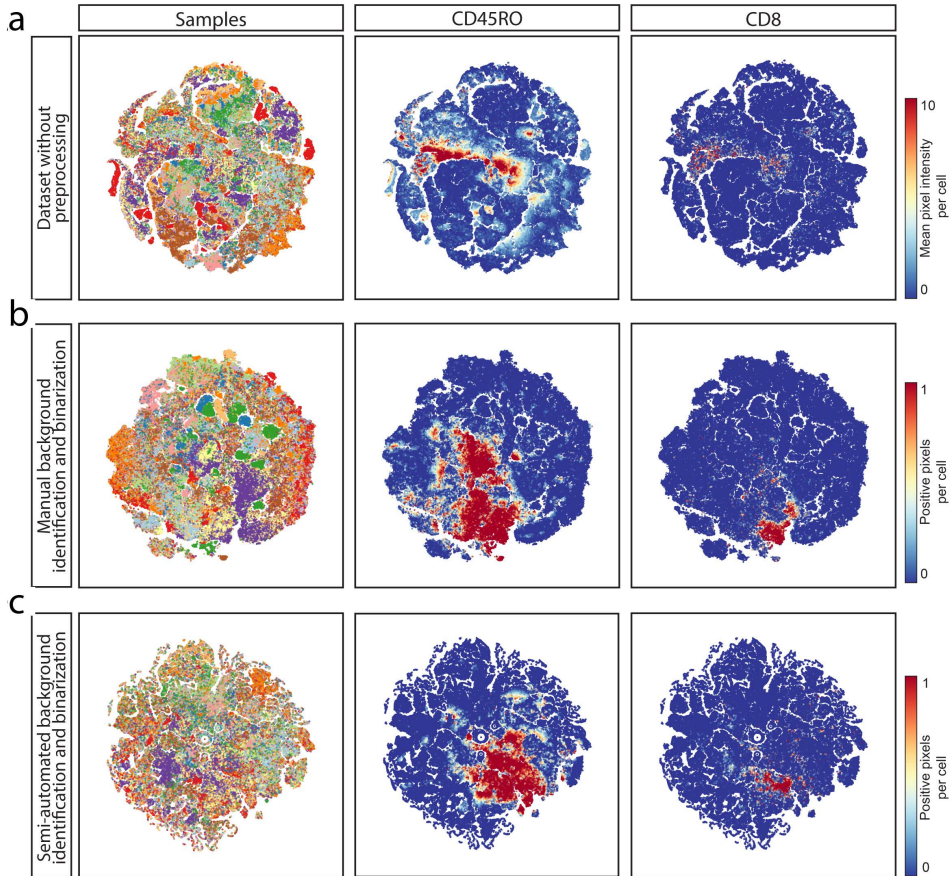
Figure 2.3: *(a) t-SNE analysis embedding of single cell data extracted from all 48 images without preprocessing. Each dot marks a cell and is coloured by sample of origin.To determine the effect of data normalisation, t-SNE analysis embeddings were generated for the same dataset after (b) manual background 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
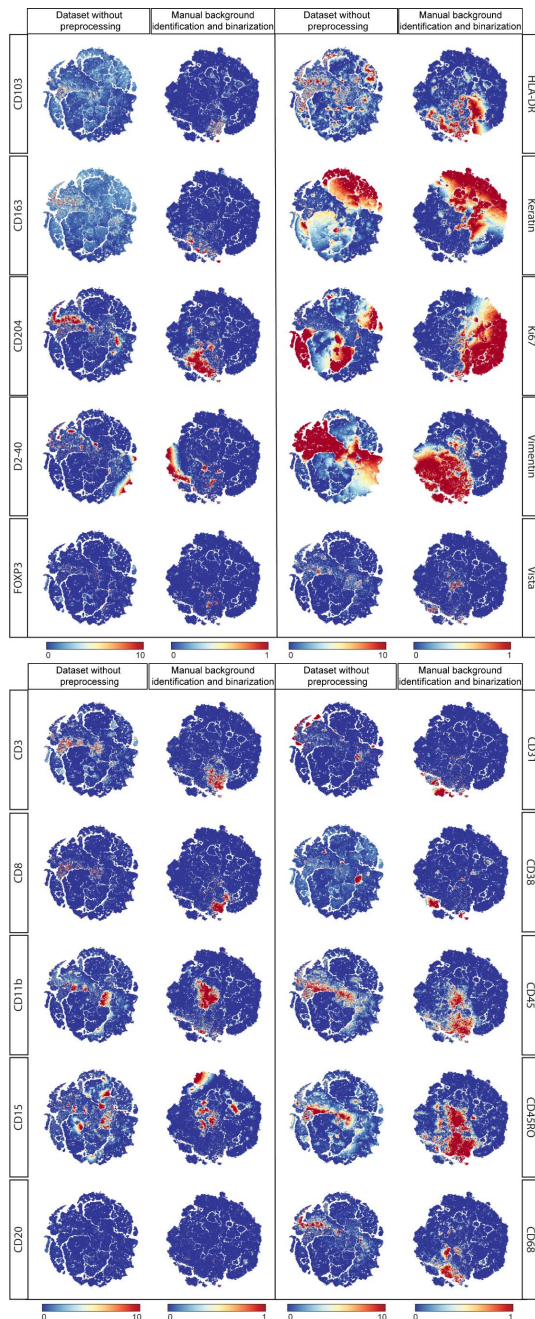
Figure 2.4: *t-SNE analysis on single cells extracted from dataset without preprocessing and the dataset generated with manual background identification and binarization.*
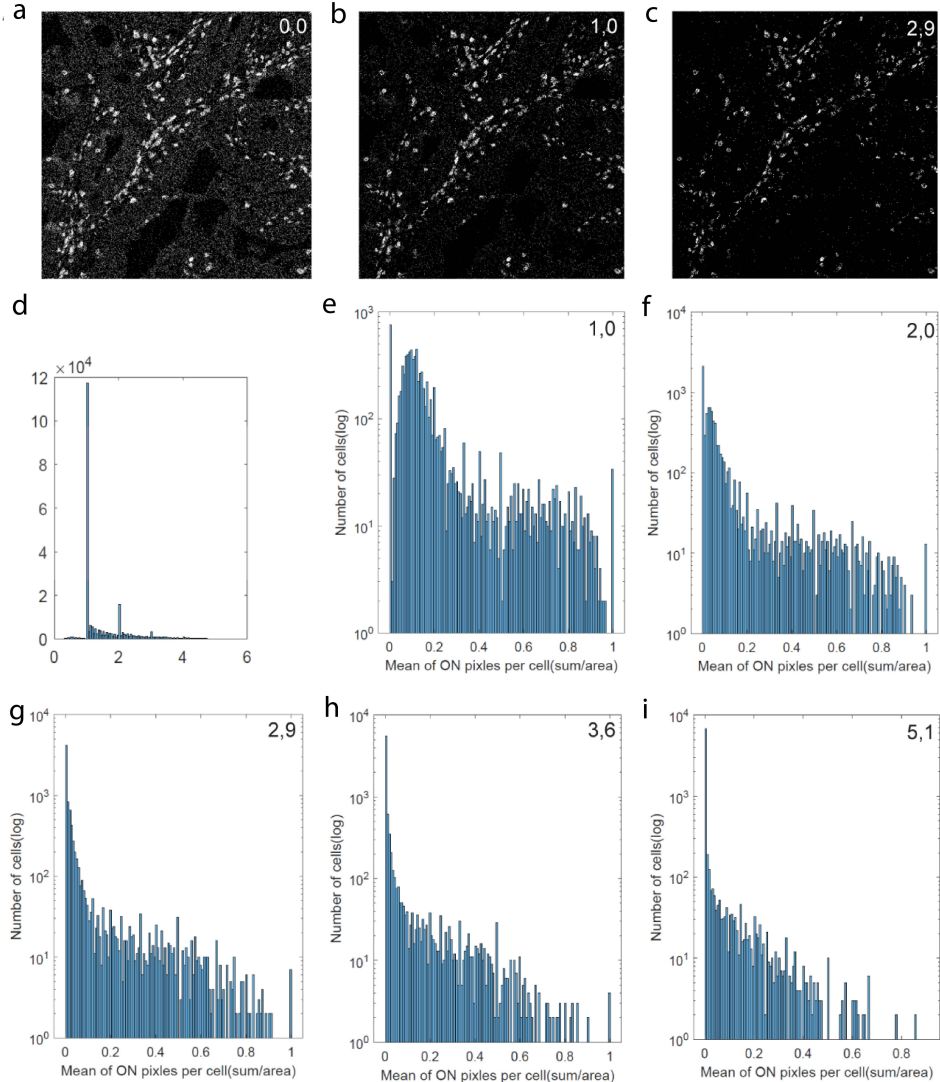
Figure 2.5: *(a-c) CD3 expression in a representative image with a pixel threshold of 0, 1 or 2,9. (d) pixel data plotted in a histogram displaying the relation between pixel intensity (x-axis) and frequency (y-axis). (e) Mean CD3 positive pixels per cell with a cut off threshold at 1. All pixels below 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
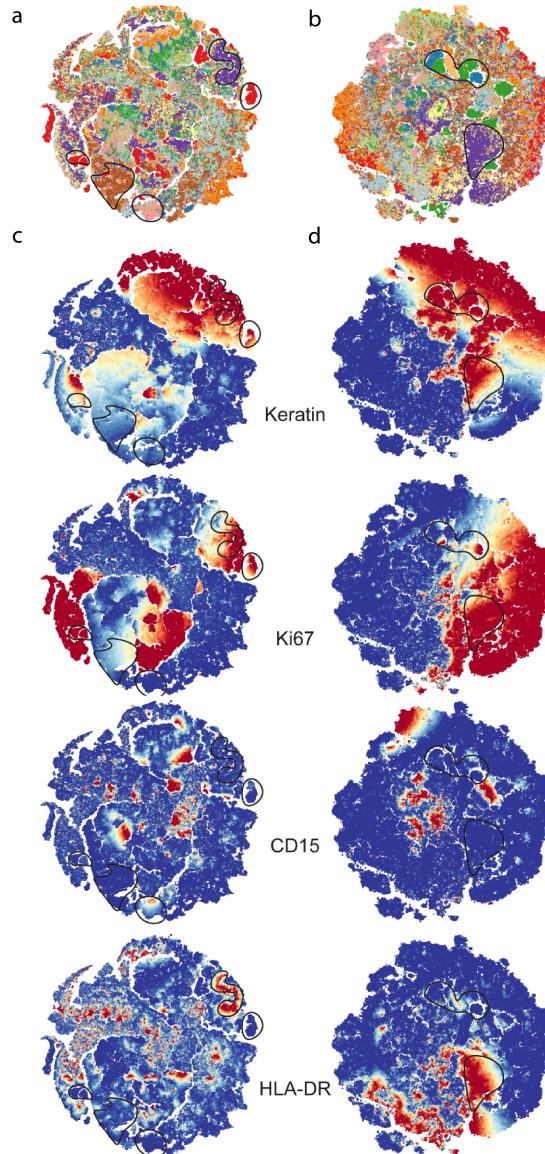
Figure 2.6: *(a) t-SNE embedding of single cell data extracted from dataset without pre-processing. Each colour represents a different sample and cells cluster together by phenotype. Sample-specific clustering is encircled. (b) t-SNE embedding of single cell data extracted from dataset generated with 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).*

2

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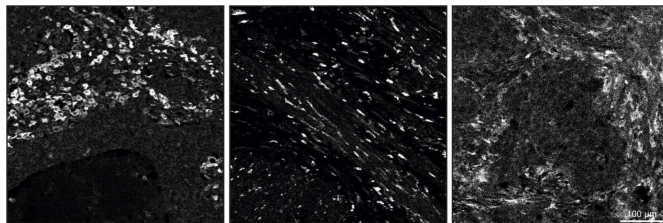
Figure 2.7: *CD163 expression pattern in the three samples after saturation of pixels below the 1st and above the 99tth percentile.*

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
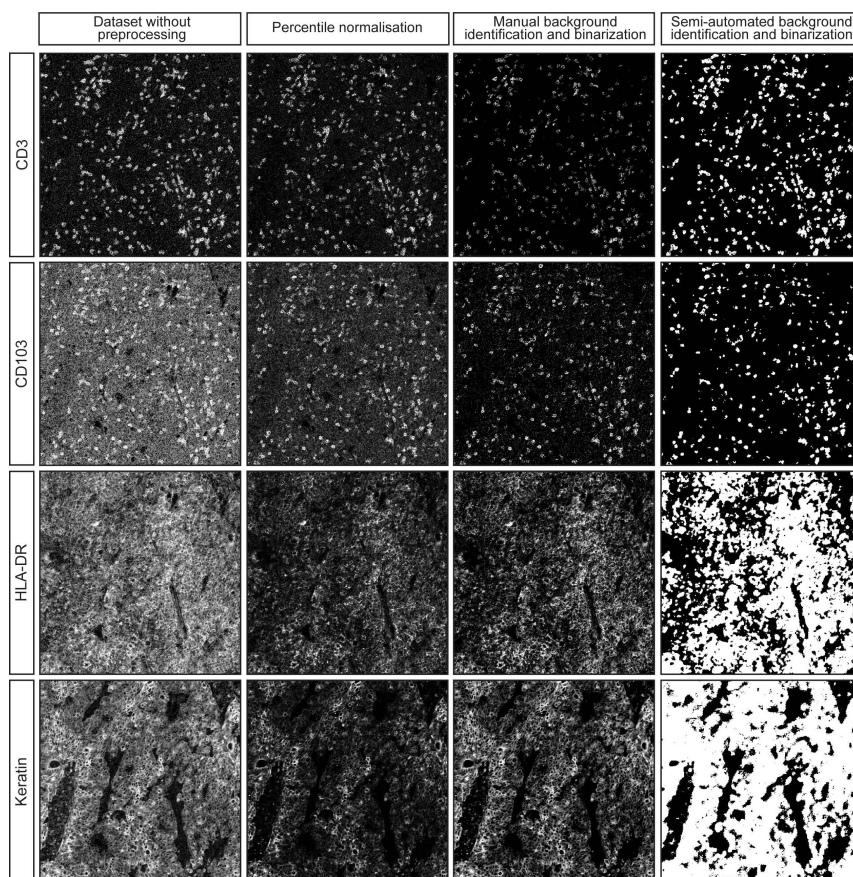

Figure 2.8: *Representative images of CD3, CD103, HLA-DR and Keratin signal without preprocessing, percentile normalisation, manual background removal or semi-automated background removal, both followed by pixel binarization*

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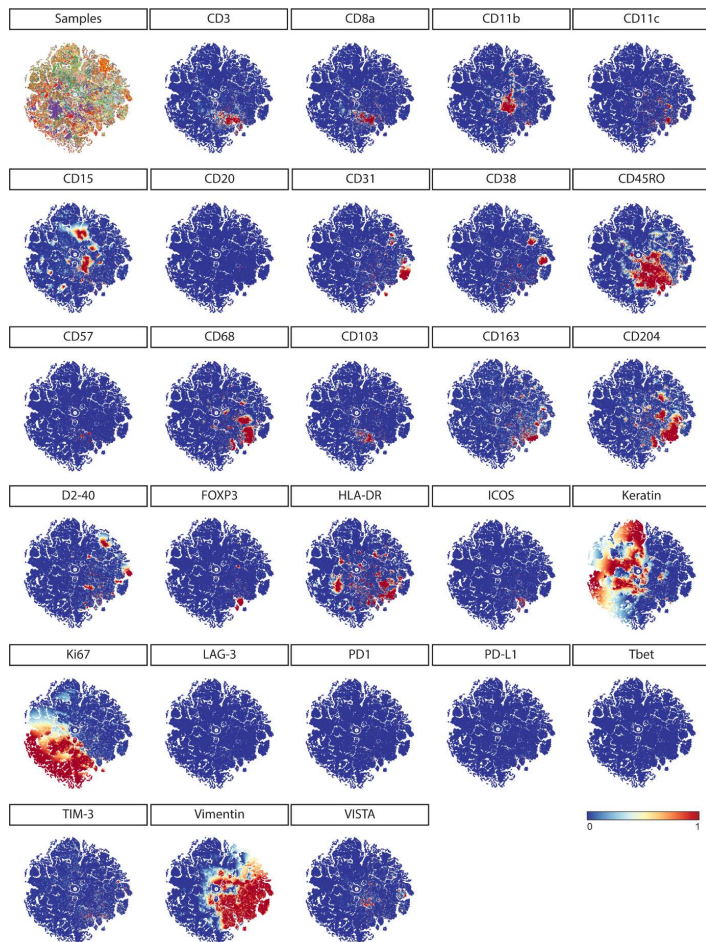
Figure 2.9: *t-SNE analysis on single cells extracted from the dataset generated through semi-automated background identification and binarization. Data is shown in a range of 0 (blue) to 1 (red).*

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
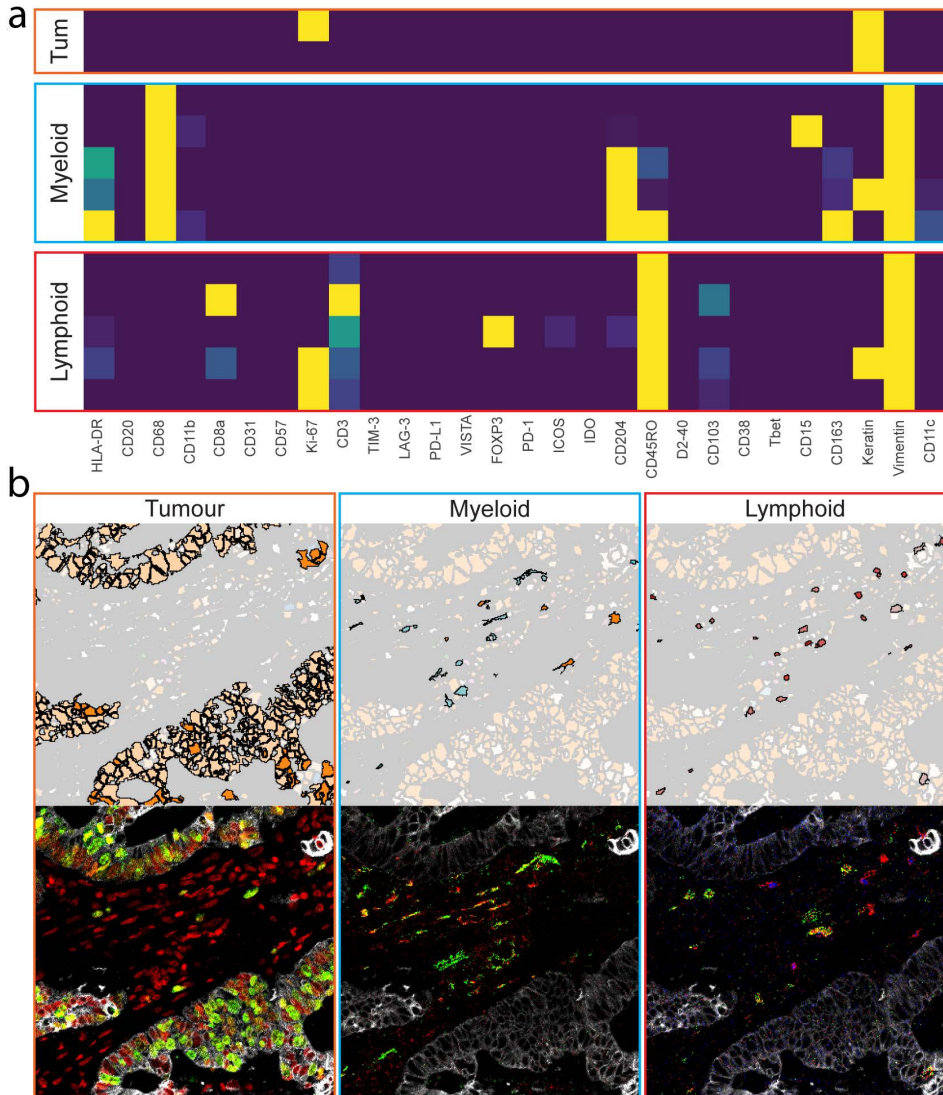
Figure 2.10: *(a) Lymphoid, myeloid and tumour (tum) clusters identified in the t-SNE analysis from Figure 2.3c . (b) the clusters were mapped back onto a representative image and compared to their corresponding cell types in the MCD files. The images contain the following markers: for tumour keratin (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

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

* 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 (ν=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).
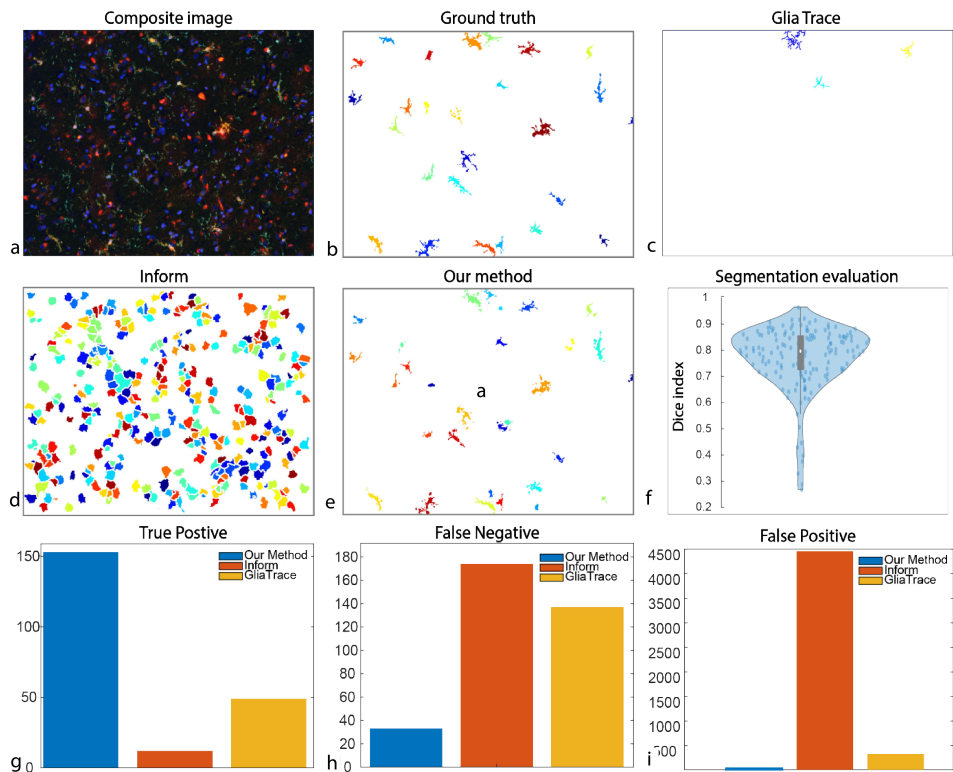

Figure 2.11: *(a) Example of multispectral immunofluorescence data extracted from the grey matter of an AD-patient. The segmentation mask derived: manually by the specialist (b), using the GliaTrace toolbox (c), automatically from the "Inform" image acquisition software (d) and from our proposed segmentation 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).*

2

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.

2