



HAL
open science

In tissue spatial single-cell metabolomics by coupling mass spectrometry imaging and immunofluorescences

Landry Blanc, Florent Grelard, Michael Tuck, Véronique Dartois, Antonio Peixoto,
Nicolas Desbenoit

► **To cite this version:**

Landry Blanc, Florent Grelard, Michael Tuck, Véronique Dartois, Antonio Peixoto, et al.. In tissue spatial single-cell metabolomics by coupling mass spectrometry imaging and immunofluorescences. 2024. ⟨hal-04730689⟩

HAL Id: hal-04730689

<https://hal.science/hal-04730689v1>

Preprint submitted on 10 Oct 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

In tissue spatial single-cell metabolomics by coupling mass spectrometry imaging and immunofluorescences

Acquisition and computational processing to correlate molecular and cellular data in complex microenvironments

Landry BLANC^{1,2*}, Florent GRELARD^{1*}, Michael TUCK¹, Véronique DARTOIS³, Antonio PEIXOTO^{2#}, Nicolas DESBENOIT^{1#}

(1) Univ. Bordeaux, CNRS, CBMN, UMR 5248, F-33600 Pessac, France.

(2) IPBS, UMR 5089 CNRS, Toulouse, France.

(3) Center for Discovery and Innovation, Hackensack Meridian Health (Nutley, New Jersey, USA).

* These authors contributed equally to this work

These authors contributed equally to this work

Corresponding authors: Landry.blanc@ipbs.fr and Nicolas.desbenoit@u-bordeaux.fr

Abstract

In this work, we introduce a multimodal imaging workflow that integrates Matrix-assisted Laser Desorption Ionization Mass Spectrometry Imaging (MALDI-MSI) combined with Immunofluorescence (IF) microscopy to enhance in tissue spatial single-cell metabolomics. The workflow allows to correlate cell populations with associated small molecule distributions by conducting on the same tissue section MSI before IF staining, addressing tissue integrity challenges and joint image analysis.

To process MSI data with IF guidance, we propose an original and advanced computational strategy utilizing Receiver Operating Characteristic (ROC) analysis, allowing to identify ions specific to targeted histological regions based on IF staining. Moreover, in a non-targeted strategy, we introduce a Spatial Coherence Measure (SCM) to distinguish genuine spatial patterns from noise within ion distributions, enhancing spatial metabolomics' robustness. Then spatial clustering techniques are employed to group ions sharing similar spatial distribution to reveal histological structures, providing complementary insights into metabolite distributions. We validated our workflow mouse spleen section as this organ presents a spatially complex but well-detailed microenvironment.

In conclusion, our multimodal and computational workflow opens new frontiers for diverse biomedical research applications by promoting precise spatial metabolomics in tissue sections.

Introduction

In the last 30 years, mass spectrometry imaging (MSI) has become one of the most attractive *ex vivo* techniques for the spatial characterization of molecules.¹ Many fields, such as pharmaceutical,² cancer,³ or neurodegenerative disease research,⁴ have profited from the ability of MSI to provide molecular maps of various molecules. The combination of MSI with other image techniques in multimodal imaging workflow has been a growing field of research for a decade,^{5,6} allowing to obtain wider biological and chemical spatial information from tissue samples. Multimodal imaging workflows combining on same tissue section MSI with microscopy techniques have proven valuable, in precisely linking molecular signatures to areas of tissues containing specific types of cells.⁷⁻¹² Recently, single-cell metabolomics has been achieved using MSI with in-vitro cultured cells.¹³ However to our knowledge, spatial single-cell metabolomic/lipidomic in a complex tissue microenvironment was still not reached. Such an approach, offering to fully understand the interplay between cellular identity and metabolite distribution within tissues, faces different challenges. MSI and microscopy acquisitions have to be done on the same tissue section at cellular resolution (below 10 μ m) to be able to link individual cells to their metabolite content. Metabolite distributions have to be highly preserved during data acquisition, imposing tissue processing techniques that prevent metabolite delocalization. On the other hand, MALDI-MSI scanning conditions at cellular resolution that require high laser power levels need to be optimized to preserve tissue integrity and allow further antibody labeling for identification of different cell populations. Moreover, MSI and Microscopy datasets have to be precisely aligned to match cellular and chemical data. Finally, advanced computational workflows have to be implemented to fully decipher the huge amount of correlative data.

Our approach presents a novel method that combines subcellular resolution MSI (5 μ m) compatible with immunofluorescent staining on the same tissue section. This allows for highly precise identification and analysis of individual cells within their tissue context. Additionally, we have developed a robust computational data processing pipeline to integrate these rich datasets, offering a holistic view of the tissue, encompassing both its chemical and histological makeup. To fully derive the benefits of our workflow as well as our original computational strategy, the mouse spleen was chosen as a biological model. This organ plays a role in blood filtration and the immune system. It presents a well-described complex microenvironment, with multiple layers of cells, including different cell types and subtypes.¹⁴ This histological heterogeneity constitutes an interesting environment to challenge and evaluate our multimodal workflow. Each spleen follicle (white pulp) is organized with a central vein surrounded by T lymphocytes, themselves surrounded

1 by B lymphocytes. Around these follicles, the red pulp contains red blood cells and macrophages, as well as migrating
2 lymphocytes.^{15–17} The junction of white pulp and red pulp constitutes the marginal zone, where various connective tissue
3 cells and immune cell subtypes expressing specific markers can be observed.¹⁵

4
5 To decipher computationally MSI/IF multimodal dataset, we used an open-source Python library, called Esmraldi,
6 dedicated to handling MSI data in multimodal workflow.¹⁸ We developed a new pipeline on Esmraldi focused on
7 deciphering histological information for multimodal datasets. We demonstrate first the potential of an automated
8 pipeline based on receiver operating characteristic (ROC) analysis to identify ions specific from targeted histological
9 regions,¹⁹ in opposition to more common univariate statistical tests such as T.test^{12,20} or Mann Whitney U test.²¹ This
10 approach relies on using IF as ground truth to generate binary masks of cell location and performs automatically ROC
11 analysis of all ion's distributions (called Volcano ROC). The ROC analysis determines the effectiveness of certain ions as
12 classifiers by assessing the overlap of intensity distributions between a specific tissue region and the surrounding tissue,
13 ultimately aiding in the identification of cell biomarkers. Secondly, we introduce an innovative clustering strategy as it
14 relies on ion spatial similarities in opposition to state-of-the-art approaches based on pixel spectral similarities.^{8,10,19,22} To
15 achieve this workflow, we first implement a new measure, the spatial coherence measure (SCM), to select *m/z* presenting
16 relevant spatial distribution. We then used hierarchical clustering to group these ions based on their spatial similarities
17 and revealed different molecular distributions highlighting as much histological area.

18
19 The multimodal imaging method presented in this article offers an exciting perspective by reaching *in situ* single-cell
20 spatial metabolomics. By coupling MSI with IF on the same tissue section and at high spatial resolution, this approach
21 allows for a paradigm shift in MSI, moving from pixel-centric thinking to a focus on cells, providing a much more precise
22 and detailed view of biological tissues with complex microenvironments. We detailed the meticulous optimization during
23 data acquisition to preserve small molecule distribution as well as the quality of cell immunostaining. For data processing,
24 we use histological information from microscopy and IF to guide MSI analysis and discover molecular signatures. Then,
25 we delve into the molecular distribution provided by MSI to identify spatial patterns that highlight relevant histological
26 areas. This approach has broad applicability across various biological studies. Enabling single-cell spatial omics within
27 tissues holds immense potential for uncovering novel biological insights and furthering our understanding of complex
28 cellular processes.

30 Results

31 1. Multimodal acquisition workflow and dataset alignment

32 To obtain a precise correlation between molecular profiles and cell types for *in situ* single-cell spatial metabolomics,
33 MALDI-MSI was performed at a subcellular resolution (down to 5 μm), and both modalities were applied sequentially on
34 a unique tissue section (**Figure 1a**). Consequently, the MSI acquisition was optimized to preserve the quality of the
35 subsequent IF acquisition. Several steps in the sample preparation of MSI can impair IF staining by damaging the IF
36 antibody targets, such as matrix deposition, matrix rinsing, and MSI scan conditions (laser parameters). For this
37 optimization, we have selected antibodies targeting different immune cells commonly found in the spleen: T
38 lymphocyte/CD4, B lymphocyte/B220, and Myeloid cells/CD11b. The results of these preliminary optimization processes
39 are described in the Method section and **Figures S1** and **S2**. Three of the four tested matrices were suitable for this
40 multimodal workflow: 1,5-Diaminonaphthalene (DAN), 2,5-Dihydroxybenzoic acid (DHB), and α -Cyano-4-
41 hydroxycinnamic acid (CHCA). 9-aminoacridine (9AA) is difficult to rinse from tissue and impaired IF acquisition with its
42 strong fluorescence at low wavelengths (**Figure S1 - 9AA panels**). The rinsing solution appears to be a key parameter, as
43 many tested solvents disturb antibody staining. Only acetone allows to keep the staining of all 3 tested antibodies. Finally,
44 laser parameters during MSI acquisition must be tuned to avoid tissue destruction. We observed that limiting its energy
45 to 7 nJ is the best compromise between MSI signal intensity and preservation of IF staining. Ultimately, DAN for negative
46 ionization mode and DHB for positive ionization mode offer the best annotation results based on Metaspace2020 queries.
47 Therefore, subsequent experiments focused on these two matrices (**Figure 1b** and **Figure S2**).

48
49 After multimodal acquisition, both datasets are preprocessed. Contrasts of IF images are corrected, and peak selection
50 and alignment, as well as normalization to an internal standard, are done on MS images as described in the Method
51 section. At this stage, both datasets differ in terms of size and spatial orientation. Our optimized MSI acquisition method
52 preserves tissue integrity but does not generate visible ablation marks that could be used to align the two images. Thus,
53 fiducials are etched around the tissue just before MSI acquisition by using a MALDI source laser at full energy (**Figure**
54 **S3a**). The focalized laser allows the etching of fine marks (around 20 μm spots with the shape of a square or cross), which

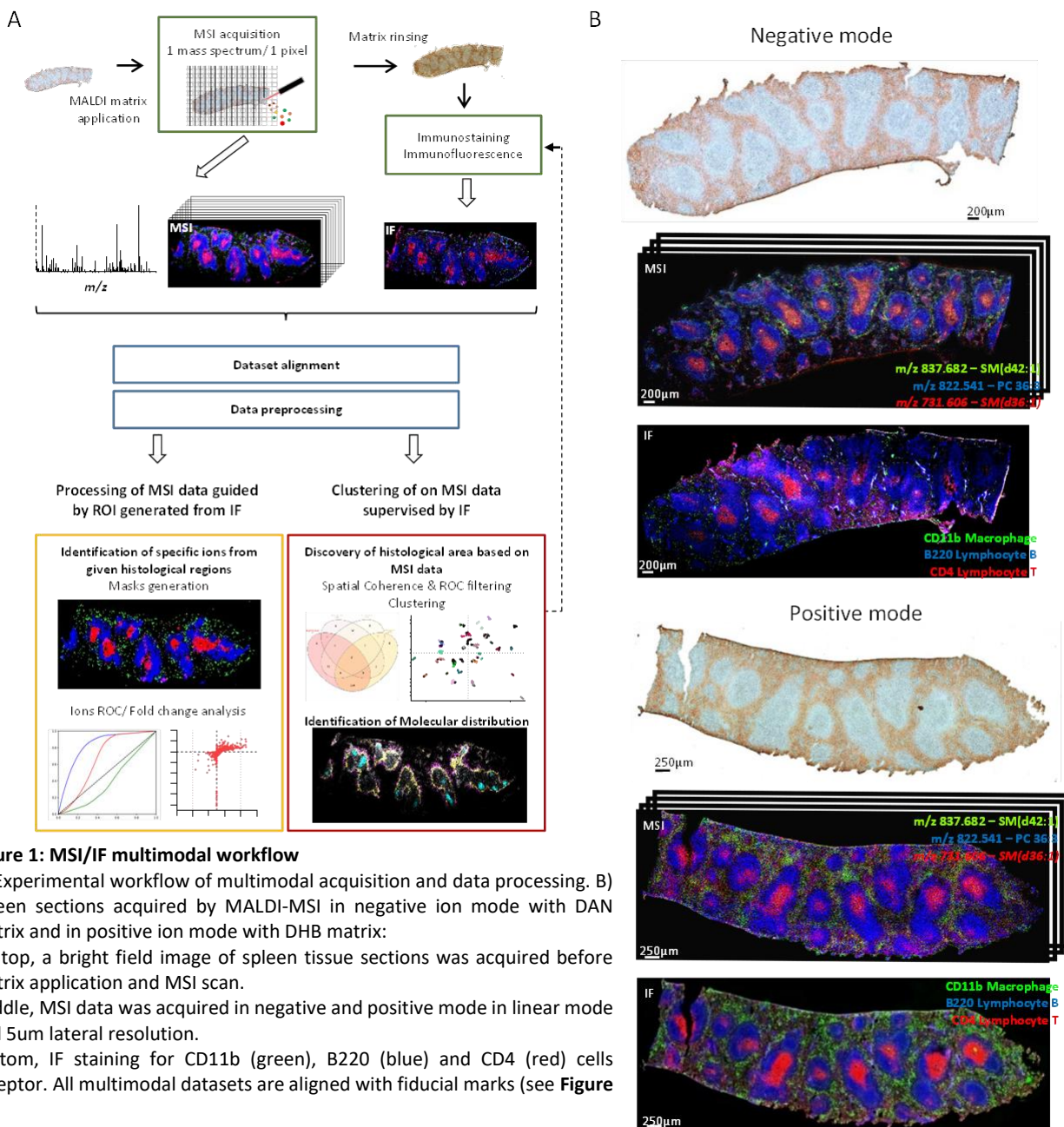


Figure 1: MSI/IF multimodal workflow

A) Experimental workflow of multimodal acquisition and data processing. B) Spleen sections acquired by MALDI-MSI in negative ion mode with DAN matrix and in positive ion mode with DHB matrix:

On top, a bright field image of spleen tissue sections was acquired before matrix application and MSI scan.

Middle, MSI data was acquired in negative and positive mode in linear mode and 5µm lateral resolution.

Bottom, IF staining for CD11b (green), B220 (blue) and CD4 (red) cells receptor. All multimodal datasets are aligned with fiducial marks (see **Figure S3**)

Ion annotations made with METASPACE 2020.

imzML file for this figure can be found here:

https://metaspace2020.eu/dataset/2021-11-09_15h54m35s (for negative mode dataset), https://metaspace2020.eu/dataset/2021-11-08_11h28m30s (for positive mode dataset)

1 can be used for precise alignment. These marks are visible in the brightfield image, as well as in transmission mode, which
 2 is inherently aligned with the IF image. During the etching process, the MALDI matrix and internal standard are fully
 3 ablated. Therefore, imprints of those marks are visible as no MSI signal can be observed from them. These marks are then
 4 used as fiducial markers to perform an affine registration (**Figure S3b**). The alignment was validated by measuring two
 5 metrics: precision and recall, which both evaluate the proportion of pixels in common between the two images. The
 6 resulting registration gave precision and recall values higher than 98%, this confirms the high quality of the registration
 7 and makes the subsequent analyses possible.

8

9 2. IF-based histology to guide processing of MSI dataset

10 To guide the MSI data processing, we selected regions of interest (ROIs) of different cell locations generated within the
 11 IF image (**Figure S4a**). First, average MS spectra from relevant areas were extracted (**Figure S4b**). A first insight into the
 12 MSI dataset is given by the comparison of average spectra from lymphocyte T in white pulp or in red pulp. Volcano plots
 13 were used to automatically find ions that are expressed differently across areas. These display the q-value of a t-test as a
 14 function of the fold change. (**Figure S4c**). The results identify metabolites with strong and significant fold change between

- the two cellular populations. Most of these metabolites are not specific for T cell populations but are significantly more abundant in either white or red pulp, as shown in **Figure S4d**.
-

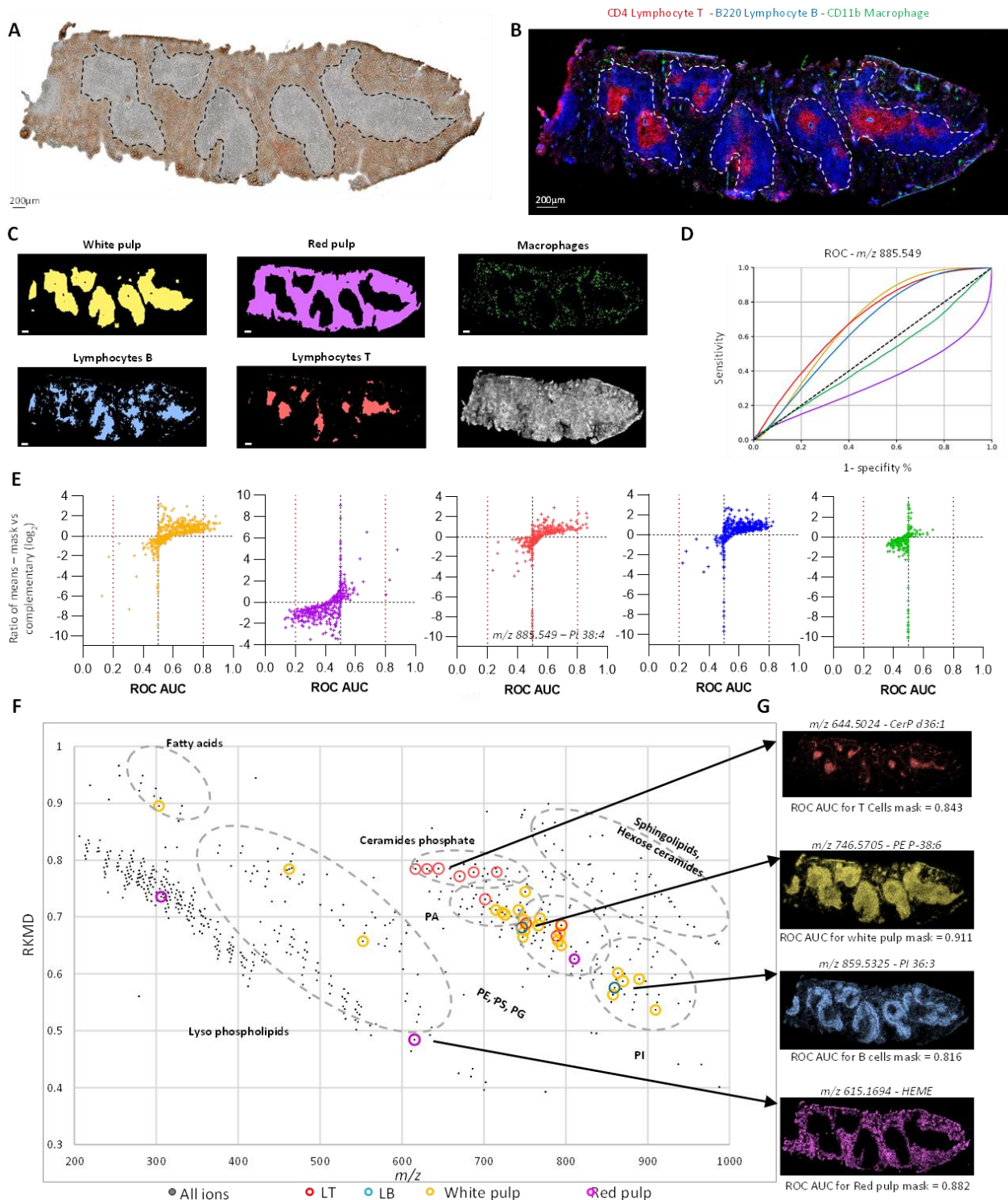


Figure 2: IF to guide MSI data processing: global strategy, multiple ROC analysis – Metabolic fingerprint

A) Bright-field image of spleen tissue section acquired before MSI analysis in negative mode B) IF staining for CD11b (green), B220 (blue) and CD4 (red) cell receptors made after MSI scan. C) IF and brightfield images have been used to generate masks of the location of the different cell populations and tissue structure: White pulp, Red Pulp, lymphocyte T, lymphocyte B and macrophage. MS image of ion m/z 885.549 annotated as PI 38:4 is shown as an example. The white scale bar corresponds to 200 μ m. D) ROC analysis has been used to evaluate the specificity and the sensitivity of ions density map for these masks, here ROC curve for m/z 885.549. E) For each ion and each mask, AUC below the ROC curve has been calculated, as well as the ratio of means ion abundance in a mask and its complementary. Similar to a volcano plot, the presented “volcano ROC” plot (ratio of means = $f[ROC\ AUC]$), allows a quick visual identification of ions that display large magnitude changes that are also statistically significant for the ROC analysis. F) Kendrick mass defect plot of the list of specific and sensitive ions (AUC of ROC curve >0.8) for each mask. This KMD plot (in base CH_2) is used organized ions by molecular families. G) MS image of ions with the best ROC AUC for each mask. Ion annotations made with METASPACE 2020. imzML file for this figure can be found here: https://metaspace2020.eu/dataset/2022-04-15_13h38m32s

1 To overcome the limitations in terms of specificity of t-test Volcano plots, we use receiver operating characteristic (ROC)
2 curve. In our context, ROC curves reflect the discriminative power of an ion with respect to a binary cell-type mask. Based
3 on microscopy (**Figures 2a** and **b**), we defined different ROIs representing tissue structures or IF-targeted cell populations
4 (binary masks of white pulp, red pulp, B220+ B cells, CD4+ T cells, and CD11b⁺ macrophages, **Figure 2c**). We reduced the
5 resolution of these binary masks (without interpolation) to match the resolution of the MSI dataset. We can perform for
6 each mask a ROC analysis of a given ion, for example, *m/z* 885.549 tentatively assigned as a phosphatidyl-inositol (PI)
7 (38:4) (**Figure 2d**). The area under the curve (ROC AUC) was used to evaluate sensitivity and specificity of an ion with
8 respect to a mask. We also calculated the ratio of mean ion abundance in a mask and its complementary (mask vs off
9 mask).

10 ROC analyses are automated for each mask and for each peak-picked ion. Then, similar to a volcano plot quickly visualizing
11 ions that display large magnitude changes, we generate a “Volcano ROC” plot: “ratio of means =f(ROC AUC)” (**Figure 2e**).
12 By selecting ions with high ROC AUC (>0.8),²³ we identify ions with specific and significative accumulation in considered
13 masks (see **Table S1**). Finally, we projected all these ions of interest in a Kendrick mass defect (KMD)²⁴ plot (**Figure 2f**) to
14 visualize the whole dataset organized by molecular families. For white pulp, multiple ions from different lipid classes were
15 seen, especially the PE group (**Figure 2g** and **Table S2**). In the red pulp, we highlight two metabolites, glutathione and
16 HEME-⁵⁶Fe. We did not find any ion with high specificity for macrophages, yet we found ions specific for B cells, tentatively
17 assigned phosphatidyl-ethanolamine (PE) P-38:6 (P for a plasmogen lipid side chain) and phosphatidyl-inositol (PI) 36:3.
18 In lymphocyte T, we observed a striking accumulation of ceramide phosphates and PE-ceramide (**Figure 2f** and **g**).
19 Performing a similar volcano ROC analysis in positive mode (**Figure S5**), we identified unique ions for the T cell area. Most
20 of these ions were annotated in Metaspaces as different ion adducts of sphingomyelin (SM) d34:1 and d36:1.
21

22
23 Here, we confirm how combining IF and MSI on the same section is a precise tool to spatially study molecular profiles of
24 cell populations. In addition, we show how automatic ROC and “volcano ROC” plots are a useful way to quickly identify
25 molecular biomarkers or to simply highlight metabolite accumulation in a targeted cell population.
26

27 3. Selection of promising *m/z* to highlight histological areas

28 MS ion images contain many spatial distributions that are not necessarily explained by regions in IF. For instance, from
29 METASPACE we noticed the distribution of the ion at *m/z* 837.549 strongly located at the border of red pulp and white
30 pulp (**Figure S6**). Such molecular distribution displays a striking organized spatial pattern. We hypothesize that such a
31 distribution highlighted other cell populations or cells with a specific metabolic state and constitute interesting targets to
32 decipher the tissue microenvironment in a molecular histology aspect. To sort the 584 peak-picked ions from the MSI
33 negative mode dataset, we implemented metrics to first automatically eliminate ions that are not specific to the sample
34 area and then to sort the remaining relevant ions based on their spatial coherence. These metrics are non-guided, in the
35 sense that microscopic or histologic masks are not required to sort ion distributions. On the opposite, we implemented
36 also guided measures using histological masks as references of tissue shape.
37

38 First, the **off-sample measure** (*OffM*) estimates how points are spread outside of a sample that is centered in the image
39 (**Figure S7a**). We start by estimating an off-sample image by using k-means with two classes and determine for each ion
40 if their signal is mostly off-sample or in-sample. In our metric, values range from 0 to 1, with values close to 0 indicating
41 on-sample ion images. Empirically, we select a threshold of 0.1, below which ion images are considered on-sample (left
42 panel of **Figure 3a**). For the guided approach (right panel of **Figure 3a**), we measure for each *m/z* the ratio of average
43 signal between inside the tissue mask (**Figure S6d**) or its complementary. Positive log₂(ratio) or close to 0 indicates a main
44 signal off-sample or nonspecific. A threshold of log₂(ratio) lower than -1 was defined to select *m/z* specific from the tissue.
45 We decided to challenge results from the guided and the non-guided in/off-sample (**Figure 3b** Venn diagram). Both
46 methods disagree with only 15 of 584 ions (around 2.6%), showing the efficacy in identifying tissue-specific and non-
47 specific ions. We excluded these 15 ions from further analysis as none appeared tissue-relevant and kept 224 ions for
48 further analysis (**Figure S8a**).
49

50 To detect spatially organized patterns in ion distributions, we propose a new measure, the SCM (**Figure S7b**). The **SCM**
51 selects spatially coherent images by estimating the amount of noise with respect to spatially coherent structures. This
52 measure ranges from 0 to 1, with high values indicating spatially coherent distributions and low values corresponding to
53 noisy ion images. Empirically, we set up a threshold of 0.65, above which ion images are considered spatially coherent.
54 In **Figure S8b**, the efficiency of SCM can be appreciated, with 117 high SCM ions showing sharper spatial structures
55 compared to 107 low SCM ions with noisier structures. In parallel, we implemented a guided strategy relying on IF masks
56 to sort ions based on their ROC AUC observed previously. As shown in **Figure 3c**, **most of the ions presented a strong**
57 **AUC in ROC analysis as well as a strong SCM score (33 vs. 3)**, confirming the potential of the SCM as a spatial measure.
58 Interestingly, on the 224 “in sample” ions, a large number (188 in total) do not present a high ROC AUC value whereas 84
59 of them present a coherent spatial distribution, potentially highlighting interesting histological structure.

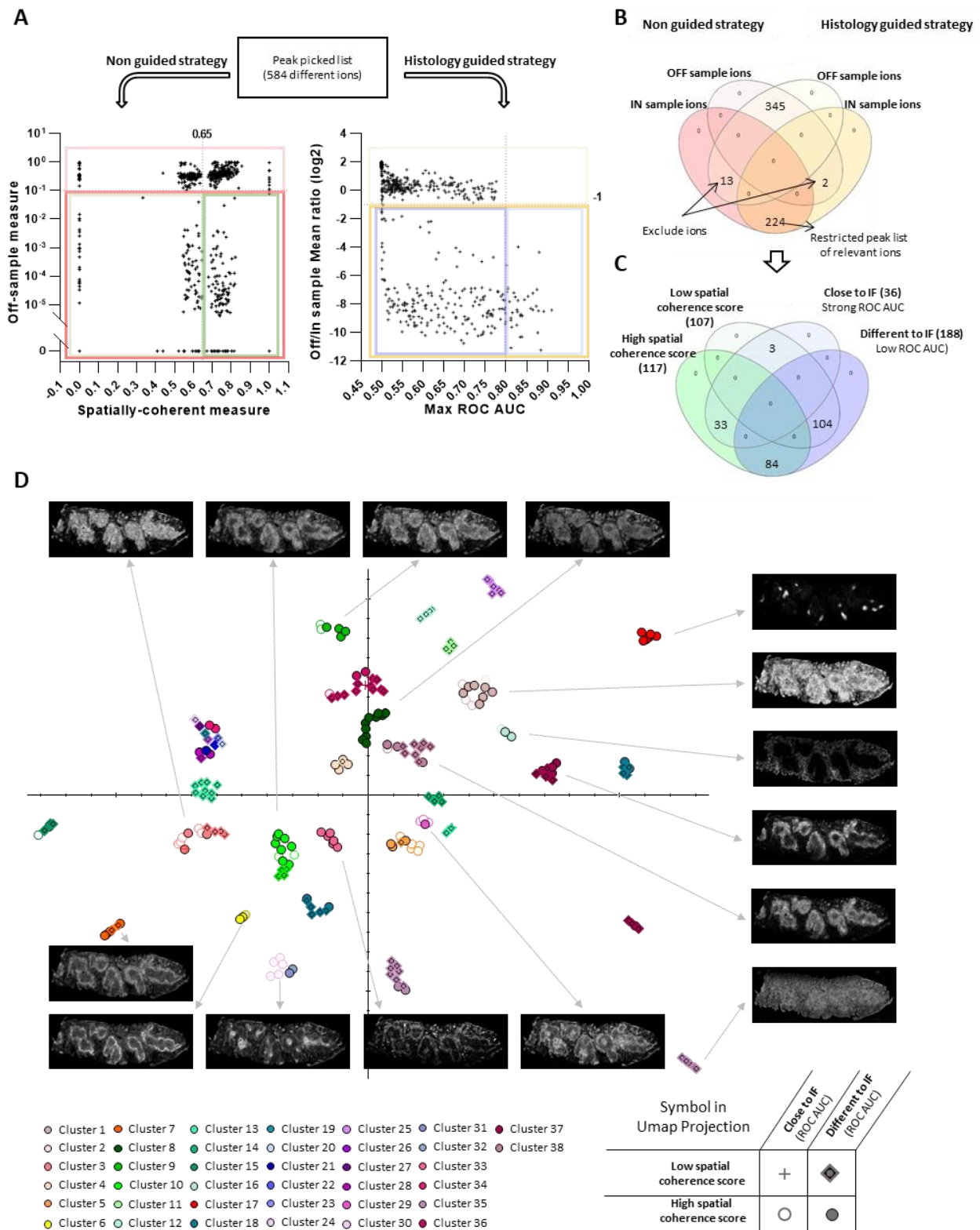


Figure 3: Peaks filtering based on guide and non-guided strategies followed by sorting and clustering based on spatial distribution

A) Partitioning plot obtained for the two strategies. On the left, is the evaluation of ion spatial distributions without any histological input from microscopy or IF. The Off/In measure define automatically where the tissue section is in the image and gives a score for each ion based on whether ion distribution is mostly inside or outside the tissue. The spatial coherence score distinguishes noisy spatial distributions compared to the highly spatially coherent. On the right side, is the evaluation with histologically guided strategy. For each m/z , the in/off sample mean ratio is based on the average ion signal inside or outside the whole sample mask (Figure S6d), and Max ROC AUC is the best AUC in the ROC curve among the different tested binary masks of the cell population (see figure 2c). B) Venn diagram of m/z for IN or OFF sample filtering considering the two strategies to select relevant ions. C) Venn diagram to sort relevant m/z based on their spatial coherence and similarities to IF masks. D) UMAP projection of hierarchical clustering of all relevant ions based on their spatial similarities + average images of some spatial clusters.

4. Spatial clustering and spectral clustering: complementary strategies to overview of histological structures highlight by molecular fingerprinting

To highlight histological areas, we implemented a spatial clustering strategy to group ions presenting similar spatial distributions. We used hierarchical clustering along with a projection of the clusters into a reduced space using uniform manifold approximation and projection (UMAP).²⁵ We used the cosine metric to build the distance matrix. The number of clusters was chosen with the inconsistency coefficient²⁶ of the hierarchical clustering tree (**Figure S7c**) set at a threshold of 1.7. This resulted in 38 clusters (listed in **Figure S8 d, e, and f**), which were further visualized by UMAP projections (**Figure 3d**) and modified to separate clusters in the reduced space using the previous distance matrix. The UMAP projection in **Figure 3d** estimates the similarities between ion distribution from each other and identify probable over-clustering (clusters #19 to #24, #26 to #27, and #34) or under-clustering (clusters #35 and #37). A detailed montage of each cluster's different ion distributions shows how efficient this clustering approach (**Figure S8f**) is to group based on spatial similarities. In addition, we sorted the 224 relevant ions into four distinct groups based on their ROC AUC and SCM scores to evaluate how they spread in the spatial clustering. It appears that ions presenting strong ROC AUC for the same mask can be split into clusters, like good biomarker ions of T.cells split between clusters # 29 and #30 or ions of white pulp between clusters #1, #3, and #5. More strikingly, clusters displaying the sharpest spatial structure, like #6, #17, #29, #30, or #33, exclusively contain ions with high SCM values, reinforcing the interest in this metric. Finally, we compared this spatial clustering strategy with the more common spectral clustering approach (**Figure S9**).²² For spatial clustering, we generated average images of each of these distribution groups to simplify the overview of clusters. For spectral clustering, we used the k-means algorithm²⁷ with the same number of clusters found by spatial clustering. Both clustering techniques were compared to identify the commonalities and specificities of each technique. On one hand, some tissue structures are identified by both methods (**Figure 4a**) like spatial cluster #30 and spectral cluster #37, highlighting the T.cells area. On the other hand, some are identified specifically by one or the other methods (**Figure 4b**), like some structures located at the border between red pulp and white pulp highlighted by spatial clusters #6, #20, and #33 or spectral clusters #6.

These results indicate that spatial and spectral clustering strategies are complementary approaches, and both should be applied in parallel to highlight histological structures efficiently.

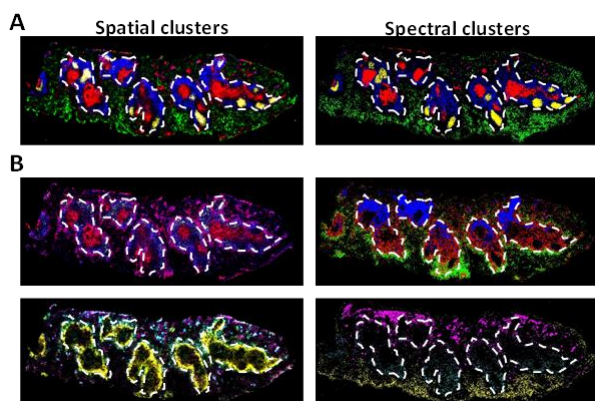


Figure 4: Spatial clustering vs spectral clustering: Complementary strategies to identify different relevant histological areas

A) Similar histological structures observed from both clustering strategies. On the left, average images of spatial clusters 30 (red), 16 (green), 37 (blue) and 17 (yellow). On the right, spectral clusters 37 (red), 12 (green), 24 (blue) and 36 (yellow).

B) Histological structures highlight only by one of the two the clustering strategies. On the left average images of spatial cluster 29 (red), 20 (green), 7 (blue), 33 (cyan), 34 (magenta) and 6 (yellow). On the right, spectral clusters 1 (red), 6 (green), 13 (blue), 14 (cyan), 34 (magenta), 30 (yellow).

White dash line highlights white pulp/red pulp border.

5. Identification of histological structure highlighted by MSI

As we identify various ions with consistent and interesting distributions that were previously unseen in IF, we wanted to annotate the associated cell populations, or in other words, select appropriate antibodies to show these cells in IF. We wished to identify cells associated with spatial clusters #6, #20, and #29, localized at the edge of the red pulp and the white pulp.

Candidates for cell populations can be selected by investigating the literature. We focused on the cell population observed at the edge of the white pulp, close to the marginal zone.¹⁵ We identified three cell populations that could match such distributions: marginal zone macrophages (MZM, positive to CD209),^{15,16} marginal metallophilic macrophages (MMM, positive to CD169)¹⁵⁻¹⁷ and marginal zone B cells (MZB, positive to CD1d).^{16,17} We selected antibodies targeting these cell populations, cut new spleen sections, and stained them directly without a prior MSI scan, to validate staining in comparison to literature (**Figure S10**). Staining for MZM (CD209) and MMM (CD169) appeared to be closer to the distribution of targeted clusters.

Therefore, we performed a multimodal acquisition of a new spleen section, still with MSI scan first followed by IF staining with new antibodies targeting MMM and MZM (**Figure 5a**). We generated the binary image of the IF signal of CD169 and CD209 for the mask of the MMM (**Figure 5b**) and MZM (**Figure S11a and b**) respectively. We performed a Volcano ROC analysis by comparing binary masks of CD169+ cells (MMM, **Figure 5b**) and CD209+ cells (MZM, **Figure S11b**) with each peak-picked ion (**Figure 5c and Figure S11d**). For the MZM mask, none of the peak-picked ions presents relevant ROC

1 AUC (**Figure S11d, e and f**). For MMM, however, 3 ions of
2 the spatial cluster#6; m/z 837.549 / 863.569 / 889.581
3 present ROC AUC >0.85. Ion m/z 837.549 (**Figure 5d**)
4 presents an interesting overlay of MMM in the spleen as
5 shown in **Figure 5e**. These 3 ions have a similar hot spot
6 distribution (**Figure 5d, e, and f**) and interestingly they are
7 annotated in METASPACE as part of the same family:
8 respectively PI 34:0 – PI 36:1 – PI 38:2. Therefore, we
9 validated the annotation of these ions by MSMS imaging
10 (**Figure S12**). We confirmed that they are respectively PI
11 18:0/16:0, PI 18:0/18:1, and PI 18:0/20:2, hinting at a family
12 of molecules, that share similar features in a subclass of
13 macrophage (MMM). We demonstrate that our spatial
14 clustering strategies of molecular distribution can point at
15 specific cell sub-populations, like spatial cluster #6
16 highlighting the MMM area.

18 Discussion

19 Spatial metabolomics is a cutting-edge analytical technique
20 that focuses on studying the small molecules within
21 microenvironments, up to individual cells.²⁸ Improvements
22 in instrument sensitivity and spatial resolution^{29,30} alongside
23 the development of algorithms and software capable of
24 effectively managing and analyzing the substantial volume
25 of generated data³¹ have streamlined the realization of
26 single-cell spatial metabolomics in cultured cells.¹³ The
27 multimodal imaging method presented in this article
28 presents a thrilling perspective on complex
29 microenvironment spatial metabolomics. By seamlessly
30 integrating MSI with IF on the same tissue section, and
31 achieving a high level of spatial resolution, this approach
32 ushers in a paradigm shift in MSI thinking. It moves away
33 from pixel-centric considerations and, instead, centers on
34 individual cells, providing an extraordinarily precise and
35 detailed depiction of biological tissues. This combination
36 delivers a comprehensive understanding of biological
37 samples, allowing molecular features to be intricately
38 mapped onto specific cellular structures. Such
39 advancements are pivotal for spatial metabolomics, as they
40 empower researchers to correlate metabolites and
41 bioactive molecules with specific cells and tissues, unlocking
42 new dimensions in the study of cellular dynamics.

44 Regarding MSI data analysis guided by IF, our work demonstrates the potential of the ROC approach, which allows for
45 the identification of ions specific to targeted histological regions. This approach addresses the limitations of traditional
46 univariate analysis by focusing on the discriminative ability of ions rather than average differences in ion values between
47 groups.¹⁹ The "volcano ROC" plot presented in this work offers a useful way to identify quickly molecular biomarkers of
48 IF-stained targeted cell population, like an accumulation of ceramide phosphate (**Table S2**) and SM (**figure S5**) in T cell
49 area, a lipid family know to be very important in T cell phenotype.^{32,33} We think that the open-source Python algorithm
50 we propose in this work might be useful for the community to achieve this goal.

52 In addition, another contribution of our study lies in the introduction of a non-targeted strategy with SCM metric and its
53 combination with spatial clustering, allowing a quick overview of the different histological distributions displayed by
54 metabolites detected by MSI. In opposition to pixel spectral clustering,²² this nontargeted approach offers to distinguish
55 between ion distributions that are mere noise and those that represent genuine spatial patterns within the tissue. By
56 quantifying the degree of spatial coherence in ion distributions, we can identify, from MSI data, regions where specific
57 metabolites accumulate (**Figure S8d and f**). This measure helps investigate new spatial structures and ensures that we
58 focus on biologically relevant ion distributions that correspond to distinct histological structures or cell populations. This
59 computational innovation opens exciting opportunities for the discovery of previously unrecognized biomarkers and the

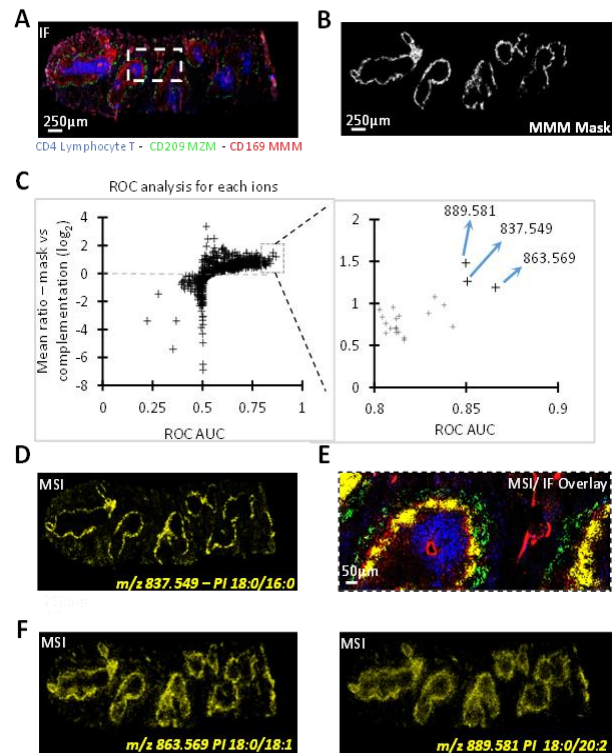


Figure 5: Identification of cell population highlighted by MSI

A) IF image of the same spleen tissue section stained for CD4, CD209, and CD169 cell receptors.

B) Display of the mask of MMM cells used for ROC analysis.

C) Multiplexed ROC analysis of all detected ions for MMM cells mask. The right panel is a zoom on the best MMM biomarkers, highlighting the 3 best ions (m/z 837.549mm, m/z 863.569, and m/z 889.581, all belonging to spatial cluster#6). D) MSI image of m/z 837.549 (PI 18:0/16:0) on a new spleen section. E) Overlay of MSI and IF image. Distribution of m/z 837.549 matches with CD169⁺ MMM cells.

F) MSI distribution of ions in the spleen m/z 863.569 and m/z 889.581. Ion annotation made with METASPACE 2020 and confirmation by MSMS imaging (**Figure S12**).

imzML file for this figure can be found here: https://metaspace2020.eu/dataset/2022-04-17_19h49m12s.

1 elucidation of intricate spatial relationships in complex biological systems. Finally, the use of spatial and spectral
2 clustering techniques provides a powerful way to highlight histological structures emerging from MSI data. These
3 complementary approaches uncover regions of interest and complex spatial patterns, thus revealing crucial insights into
4 the distribution of metabolites and their correlation with cellular biology.

5
6 The impact of this method on spatial metabolomics is significant because it enhances our understanding of metabolite
7 distribution in complex and heterogeneous tissues, which is essential for studying metabolic responses in diverse
8 biological contexts. This multimodal imaging approach and its associated computational processing have demonstrated
9 its potential to highlight specific biological features within the spleen. For instance, it enables the identification of
10 metabolites associated with distinct cell populations, such as T cells and MMM. ROC analysis in conjunction with IF masks
11 revealed ions with significant and distinctive accumulation patterns within T cell-rich regions. Notably, ceramide
12 phosphates PE-ceramides and SM accumulated prominently in T cell areas. This result sheds light on the importance of
13 this lipid class in T cells and their potential roles in immune responses.^{34,35} Similarly, SCM and spatial clustering highlighted
14 the accumulation of metabolites in specific but untargeted histological areas (**figure S6c**). Thanks to new IF staining and
15 ROC analysis, we identify that some clustered ions exhibit strong associations with MMM, notably ions from spatial
16 cluster #6, *m/z* 837.549, 863.569, and 889.581. Subsequent MS/MS imaging confirmed these ions as members of the
17 phosphatidylinositol (PI) lipid family, respectively PI 18:0/16:0, PI 18:0/18:1, and PI 18:0/20:2 (**figure 5 and S12**).
18 Interestingly, we observed overexpression of a lipid-containing eicosadienoic acid (EDA) in MMM. EDA is a rare fatty acid
19 observed in animal tissues and involved in pro-inflammatory modulator production in murine macrophages.³⁶ These
20 findings offer valuable insights into the metabolic characteristics of MMM and their functional roles within the spleen.³⁷
21 In summary, this multimodal imaging approach not only provides a comprehensive view of the spleen's histological
22 structures but also allows for the identification of metabolites specific to key cell populations.

23
24 In conclusion, the multimodal imaging method presented in this article opens new avenues for spatial metabolomics in
25 MSI. It pushes the boundaries of MSI data analysis by combining chemical and histological information, using innovative
26 algorithms, and enabling detailed cellular analysis. This advancement is critical for better comprehending the distribution
27 of metabolites in biological tissues and its impact on cellular biology, paving the way for new discoveries in various fields
28 of biomedical research. In addition, we provide an open-source Python library to ensure the reproducibility of our
29 workflow. We believe that these tools will be highly valuable to the spatial omics community, particularly with the
30 emergence of multiplexed imaging strategies such as spatial transcriptomics, cyclic immunofluorescence, and imaging
31 mass cytometry.^{38,39} Furthermore, the algorithms we propose are not limited to MSI data (.imzML) but can also handle
32 images in .tif format containing a large number of channels, making these tools more versatile for such multiplexed
33 imaging applications.

34 35 Method

36 1. Tissue sectioning

37 Mouse spleen biopsies were purchased from Charles Rivers (Collected from 5 mice male or female). Samples were stored
38 at -80°C until sectioning. Sections (12 μm -thick) were cut from the spleen using an NX70 Star cryostat (Thermo Fisher
39 Scientific) and thaw-mounted onto standard glass microscope slides.

40 41 2. Matrix and internal standards application

42 For the MALDI-MSI, a homogenous matrix layer was deposited using a home-built pneumatic sprayer⁴⁰. Briefly, 150 μL of
43 a matrix solution was applied to the tissues under the following optimized conditions: a 0.05 mL/min matrix solution flow
44 rate and 0.9 bar of nitrogen gas flow rate. Matrix solutions are detailed in the table below:

Matrix	Concentration (mg/ml)	Solvent composition	Solvent Ratio	Quantity of matrix on slide ($\mu\text{g}/\text{mm}^2$)
DAN	5	Acetone/water	7:3	1.5
CHCA	5	Acetonitrile/acetone/water	4:3:3	1.5
DHB	20	Propanol/Toluene	7:3	4.8
9-AA	5	Methanol/water	5:5	1.5

46
47
48 For MS signal normalization, deuterated internal standards have been used. Matrix solutions were supplemented with
49 0.2 mg/ml of 1-palmitoyl-d31-2-oleoyl-sn-glycero-3-phosphate (16:0-d31-18:1 PA, Avanti Polar Lipid, [M - H]⁻, *m/z*
50 703.6697 \pm 0.003) or 1-palmitoyl-d31-2-oleoyl-sn-glycero-3-phosphoethanolamine (16:0-d31-18:1 PE, Avanti Polar Lipid,
51 [M - H]⁻, *m/z* 746.710 \pm 0.003) for negative mode MS. A 1-palmitoyl-d31-2-oleoyl-sn-glycero-3-phosphocholine (16:0-d31-

1 18:1 PC, Avanti Polar Lipid, m/z 790.7726-[M+H]⁺, m/z 812.754-[M+Na]⁺ and m/z 827.7397-[M+K]⁺) was used for positive
2 mode MS.
3

4 3. MSI instrument and acquisition

5 Acquisitions were performed using an atmospheric pressure imaging ion source named AP-SMALDI 5 AF (TransMIT
6 GmbH) connected to an orbital trapping mass spectrometer (QExactive Orbitrap, Thermo Fisher Scientific).⁴¹ This latter
7 was operated in negative or positive ion mode at a mass resolution of 70 000 @ m/z 400 over a mass range of m/z 200-
8 2000. The ion source is equipped with a diode laser (Flare NX343, λ = 343nm), operating at a repetition rate of 2 kHz.
9 Imaging data were acquired in continuous mode with a pixel size of 5 μ m, and a speed rate of 3.7 pixels/s. Laser frequency
10 and shots intensity were measured previously with a pyro-electrical receiver (μ -joule meter, Lasertechnik Berlin),
11 mounted directly on the source stage. For our acquisition parameters, source drivers adjust laser frequency to 95Hz.
12 Concerning laser energy, 20% filter was ON and the attenuator angle was set a 39°, giving ultimately 7nJ per shot. Matrix
13 peaks have been used as lock mass during acquisition to compensate for mass shifting.

14 For species MSMS identification, we define small areas in the spleen section containing splenic follicles and perform
15 MSMS imaging by selecting parent ions and generating images of ion fragments (**Figure S12a**). Fragmentation is
16 performed with HCD at normalized collision energy set at 28% (performed on Q-Exactive). To have a precise image of
17 parent ions distribution, a consecutive MS1 image is performed on the same tissue section with the acquisition pattern
18 flipped by 90° (**Figure S12b**).
19

20 4. Matrix rinsing, Immunofluorescence staining and acquisition

21 After MSI acquisition, tissue sections on glass slides are recovered from the MSI instrument and rinsed with two baths of
22 acetone, 2 min each. Acetone rinsing was chosen over various solvent conditions according to our optimization presented
23 in the Results section and in **Figure S2**.

24 Tissue sections were washed three times with Cell Stain Buffer (CSB) (Biolegend Netherlands, #420201) and incubated
25 for 30 minutes at room temperature with Fc Block (Biolegend Netherlands, # 101320). After 3 washes with CSB buffer,
26 tissue sections were incubated with either CD45R/B220 Alexa 488 (clone RA3-6B2, #103225), CD4 Alexa 647 (clone GK1.5
27 # 100424) and CD11b Brilliant Violet 421 (clone M17/0, # 101251) or CD169 BV421 (clone 3D6.112, #142421) and CD209b
28 (clone 22D1, #53-2093-82) antibodies and CD11b Brilliant Violet 421 (clone M17/0, #101251). All antibodies were
29 purchased from BioLegend Netherlands, except CD209b purchased from eBioscience. Fluorescently labeled antibodies
30 were incubated overnight at 4° in CSB buffer. After three washes with CSB buffer, tissue sections were mounted with a
31 coverslip (Zeiss #1.5H) in Prolong Glass Gold mounting media (ThermoFisher, #P36980). After immunostaining, tissue
32 sections images were acquired in tile mode using a Zeiss LSM710 NLO or LSM900 confocal microscope equipped with a
33 20x Zeiss NA 0.8 objective, at a zoom of 0.6x, an image size of 512x512 pixels and with a pinhole partially opened (108.7
34 or 107.4 μ m).
35

36 5. Optimization of MSI workflow to turn it compatible with IF staining

37 First, we assessed the impact of matrix deposition and its rinsing on the IF staining. To do so, we decided to spray on
38 tissue the four main matrices used by the MSI community on tissue section, 1,5-Diaminonaphthalene (DAN) and 9-
39 aminoacridine (9AA) for negative mode or 2,5-Dihydroxybenzoic acid (DHB) and α -Cyano-4-hydroxycinnamic acid (CHCA)
40 for positive mode. We then rinsed matrices after a few hours a room temperature prior to performing IF staining with
41 antibody targeting immune cell receptors (B220/Alexa488, CD4/Alexa 647, CD11b/brilliant violet 421, **Figure S1**). For this
42 first evaluation of matrix impact, no MSI scan was performed. Interestingly, in our negative control, without any wash
43 prior to IF staining, only CHCA did not impair the staining of all the chosen antibodies. In contrast, we see that the 3 other
44 matrices induced strong autofluorescence in wavelength close to 400nm. We do not observe major interference for two
45 other antibodies coupled with Alexa probes 488 and 647nm. The matrix 9AA is more difficult to remove as revealed by
46 the strong autofluorescence in the low wavelength observed all over the slide.

47 We then tried to rinse MALDI matrices with different solvents (Acetone, ethanol, methanol, isopropanol, and toluene).
48 Only acetone presented suitable results, preventing damage to the epitope of the targeted cell receptor and efficiently
49 removing 3 matrices (9AA still present some non-specific/autofluorescent signal). All the other solvents do not remove
50 matrices efficiently or impair the IF staining specificity. For example, the ethanol wash of DAN matrix seems acceptable
51 for B220 and CD4 staining specific but impaired the CD11b staining.
52

53 We then optimized the MSI scan condition to limit tissue damage induced by the laser. We decided to pursue this work
54 with the 3 matrices for which acetone wash works well: DAN, DHB, and CHCA. We scan delimited areas on spleen tissue
55 (previously sprayed with matrices) in continuous mode at different lateral resolutions (5 and 25 μ m) and different laser
56 intensities (**Figure S2**). The source driver adjusts automatically laser frequency depending on pixel rate (here 3.7pixels/s)
57 and lateral resolution, for 5 and 25 μ m frequency is set at 95Hz and 430Hz, respectively. Once the different areas are
58 scanned, the tissue is rinsed with acetone, and IF staining is performed. The laser does not touch the tissue between the
59 different scanned areas, allowing us to properly evaluate its impact on IF staining with the different MSI scan conditions.

1 We observed that with the highest energy, the laser damaged the tissue and IF staining is impaired on the laser path
2 (**Figure S2**; area #4, #5, #12, #13, and #14). The damage of the laser is even more striking at our best lateral resolution
3 (5 μ m, corresponding to the size of the laser spot) where most of the IF signal is lost (**Figure S2**; area #7, #8, #17 and strong
4 attenuation in area #16). On the other hand, we demonstrate that at lower MSI laser energy, IF signal can be preserved
5 while getting and satisfying MSI signal (**Figure S2**; area #1, #2, #6, #9, #10, #12, #15, and #18). Interestingly CHCA matrix
6 seems to provide the best protection from the laser, as demonstrated in areas #23 and #26 (**Figure S2**) where even at the
7 maximum tested laser energy the IF signal is barely affected, however, the MSI signal is low. By the same time to properly
8 evaluate the best compromise between IF staining and quality in MSI signal, we load the different MSI datasets obtained
9 from the different areas in METASPACE2020 website to perform automated annotation. Interestingly, for all the matrices
10 and lateral resolution conditions, we found that lower energy allows to achieve more species annotation (Table in **Figure**
11 **S2**). At high laser energy, the MS signal is more intense but such as the noise, impairing proper annotation.
12 We observe our best annotation results with DAN in negative mode and DHB in positive mode at lower energy and for
13 both at 25 μ m and 5 μ m lateral resolution. To conclude this first step of MSI optimization to be included in our multimodal
14 workflow, we chose DAN and DHB matrices sprayed on tissue and rinsed with acetone baths. The MSI scan can be
15 performed at any pixel size but should not exceed 7nJ.

17 6. Dataset alignment and validation

18 Registration was performed manually using handle transform guided by fiducial marks choosing internal standards m/z
19 for the MS image (**figure S3b**). Then, the registration was assessed using the precision (p), recall (r) and F-score = $\frac{2 \times p \times r}{p+r}$
20 on binarized images, using an average image for the MSI dataset (**Figure S3c**).

22 7. MS data preprocessing: ion annotation, peak picking, internal standard normalization

23 Ion annotation was performed using METASPACE web interface⁴² based on peak exact mass and isotope profiles. MALDI
24 images were generated using MSiReader software⁴³ or Esmraldi,¹⁸ an in-house software; accessible at
25 <https://github.com/fgrelard/Esmraldi>.

26 Peak picking was performed on the mean spectra using the prominence of local maxima. The prominence is the height
27 of a peak relative to the lowest contour line separating it from a higher peak within a window. This ensures peaks from
28 signal irregularities are discarded. Here, the window size w depends on the chemical tolerance (tol , in ppm). Let x be the
29 current m/z ratio, and $\Delta(f(x)) = f(x+1) - f(x)$ be the finite difference operator, the window size w is given as: $w(x)$
30 = $\frac{tol \times x}{\min(\Delta(x))}$. Next, we apply spectral alignment to map each pixel's spectra onto the detected peaks. All m/z ratios that fall
31 within the chemical tolerance window of a detected peak are averaged and mapped to the m/z of that peak. The spectra
32 were deisotoped. For each peak, isotopes were found by searching for peaks with a 1 Da +/- 14 ppm adduct within a 8
33 Da window. The first peak was kept while the others were removed.

34
35 Signal normalization is performed by dividing the signal of a given ion by the signal of the deuterated internal standards.
36 The following calculation involving MSI data used normalized data.
37 The resulting data structure is dense, as opposed to sparse prior to preprocessing. Thus, the images were exported in Tiff
38 format to facilitate visualization in traditional imaging software, such as ImageJ.

41 8. IF preprocessing / Cellular region extraction

42 Red pulp and white pulp region are obtained by thresholding the blue channel bright field (**Figure S5a**) image in ImageJ,
43 blue channel offering the better contrast for blood signature. Cellular regions are obtained by thresholding each channel
44 of the IF image (**Figure S5b**) with ImageJ. The red channel corresponds to lymphocyte T, the blue channel to lymphocyte
45 B, and the green channel to myeloid cells.

47 9. Univariate analysis

48 For the Volcano plots, to compare two regions of interest, pixels in those areas are considered as individual data points.
49 For each ion, we determine the average signal and standard deviation of the signal as well as the number of pixels in
50 selected regions. We used GraphPad Prism software to perform a multiple t.test with FDR threshold (q value) set at 1%,
51 with Benjamini, Krieger, Yekutieli method and assuming no consistent SD. We then calculate $\log_2(\text{ratio})$ between the two
52 regions to generate a volcano plot (**figure S4d**).

53 For Receiver-Operating Characteristic (ROC) analysis⁴⁴ of a selected MSI dataset, we considered pixels as individual data
54 points. We generate an ROC curve for a given ion given delimited histological regions. The ROC curve is generated as
55 follows: for each possible intensity level in the ion image, the ion image is thresholded and the specificity (Sp) and
56 sensitivity (Se) are calculated with respect to the mask of the region. The ROC curves plot Se against $1-Sp$ (**Figure 2d**) from
57 which we can determine the area under the curve (ROC AUC) which measures the specificity of an ion with regard to a

1 region. Then, we used a routine to automate this process for all ions in all regions. ROC AUC and the ratio of ion mean
2 intensity in the mask and its complementary are calculated, and then we plot them to obtain the “volcano-like” plot
3 (Figure 2f). We consider an ion as a good biomarker of a region when its ROC AUC is above an empirical threshold of
4 0.8.²³

6 10. Spatially coherent distribution filtering algorithm

7 At this stage, the MS image still encloses interesting spatial distributions that do not correlate with IF-targeted cells. We
8 propose a new method that automatically identifies these ions. For all metrics, we compute binarized versions of the
9 image over various thresholds. Indeed, the spatial distributions enclosed by each ion are apparent at different intensity
10 levels. Intensity distributions across ions have a different number of modes, so automatic thresholding methods such as
11 Otsu’s are not suited.⁴⁵ In the following, we choose the set of percentiles $P = [30, 50, 60, 70, 80, 90, 95]$. In the following,
12 the ion images were normalized between 0 and 255.

13 We define new distance-based measures. First, the **off-sample measure** ($OffM$) estimates how points are spread
14 outside of a sample that is centered in the image (Figure S6a). We start by estimating an off-sample image. For each ion
15 image, the minimum distance between the center of the image and the points in the binarized image is computed. There
16 are two distinct groups of distances with low values, corresponding to on-sample ions, and high values, corresponding to
17 off-sample ions. We separate them using K-means with two classes. Then, a binary object B is built adding every point
18 in the class of ions images with high off-sample values. Finally, the off-sample measure is defined as the average of the
19 intensity of an ion image I in the binary off-sample object B , i.e., $OffM(I) = \frac{1}{|B|} \sum_{p \in B} I(p)$. Values range from 0 to 1,
20 with values close to 0 indicating on-sample ion images. Empirically, we select a threshold of 0.1, below which ion images
21 are considered on-sample.

22
23 Next, our **spatially-coherent measure** (SCM) estimates the amount of noise with respect to spatially coherent structures
24 (Figure S7b). It builds on distance transformations (DT) of the ions. The DT is a mapping of points in a binary object to
25 their distance to the complement of the object. Let O be an object, \bar{O} be its complement, the Euclidean DT maps the
26 following value to each point $p \in O$: $DT(p) = \min_{q \in \bar{O}} \|p - q\|_2$. The spatially-coherent measure of an ion image estimates
27 the prominence of noise, i.e., small structures with low DT values, with respect to larger structures. This is done by
28 comparing the average of the DT values in the objects with the maximum of those values. It is defined as the minimum
29 of $M(O, t) = 1 - \frac{\frac{1}{|O|} \sum_{p \in O} DT(p) - t}{\max(DT(p) - t)}$ over the thresholds $t \in P$ (set of percentile). Then, we range from 0 to 1 SCM
30 values, with high values indicating spatially coherent distributions and low values corresponding to noisy ion images. We
31 set a threshold of 0.65, above which ion images are considered to be spatially coherent.

32 For subsequent analysis, we used thresholded ion images. For each ion, we used as threshold t the one allowing to
33 minimized $M(O, t)$.

36 11. Clustering and UMAP projection

37 We wanted to visualize the distributions of the resulting ions and the impact of each criterion using two clustering
38 techniques: spatial and spectral. Spatial clustering consists of grouping ion images, and spectral clustering consists of
39 grouping pixels. Both techniques are complementary since less frequent spatial distributions might be better highlighted
40 by spatial clustering, while small structures sharing similar spectra might be better evidenced by spectral clustering.

41 For spatial clustering, we used hierarchical clustering along with a projection of the clusters into a reduced space using
42 UMAP.²⁵ We used the cosine metric to build the distance matrix. We chose the number of clusters by using the
43 inconsistency coefficient²⁶ of the hierarchical clustering tree. We decide to set a threshold of 1.7 in inconsistency leading
44 to 38 clusters. Clusters were further visualized by UMAP projection, using the previous distance matrix, modified to
45 separate clusters in the reduced space. For each row corresponding to an ion m_i with a cluster number k_i , the cell values
46 associated to a cluster number $k_j = k_i$ were not modified. The other cell values ($k_j \neq k_i$) were multiplied by a factor of
47 10.

48
49 For spectral clustering, we used the K-means algorithm²⁷ with the same number of clusters found by spatial clustering.
50 Both clustering techniques were compared to identify the commonalities and specificities of each technique. More
51 specifically, we estimate whether a spatial structure was found by both methods or was specific to either method. We
52 used our distance similarity measure between the average images from the spatial clustering and the cluster images from
53 the spectral clustering, with a threshold of 5.

54

12. Code availability

The algorithm codebase is accessible on GitHub at <https://github.com/fgrelard/esmraldi>. The repository contains documentation as well as examples of how to use the methods presented in this article. The methods can be found in the following files:

- Pre-processing (peak picking, spectra alignment): Esmraldi GUI, in `gui/Esmraldi.py`
- ROC: `examples/roc.py`
- Offsample measure and SCM: `examples/find_structured_images.py`; and the definition of the measures in `esmraldi/segmentation.py`, using the `find_similar_images_distance_map` function.
- Spatial clustering with UMAP: `examples/hierarchical_clustering.py`

Acknowledgments

This work was funded by a grant from the “Fondation pour la Recherche Médicale” (FRM, grant #ARF201809007123) to support L.B.'s post-doctoral fellowship. Part of this work was supported by grant #INV-004704 from the Bill and Melinda Gates Foundation to V.D. Part of this work was financially supported by the Agence National de la Recherche (France, Grants ANR-19-CE29-0010601 “MultiRaMaS”), the CNRS (International Emerging Actions (IEA) “MultiRaMaS”, and Groupement de Recherche “GdRMSI” GDR2125) to N.D. Part of this work was supported by Fondation “Toulouse Cancer Santé” (TCS, grant #2023CS130).

Bibliography

1. Porta Siegel *et al.* MALDI Mass Spectrometry Imaging. (Royal Society of Chemistry, 2021). doi:10.1039/9781839165191.
2. Blanc *et al.* Impact of immunopathology on the antituberculous activity of pyrazinamide. *Journal of Experimental Medicine* **215**, 1975–1986 (2018).
3. Ferraro *et al.* Fatty acid synthesis is required for breast cancer brain metastasis. *Nat Cancer* **2**, 414–428 (2021).
4. Kaya *et al.* Novel Trimodal MALDI Imaging Mass Spectrometry (IMS3) at 10 μm Reveals Spatial Lipid and Peptide Correlates Implicated in A β Plaque Pathology in Alzheimer’s Disease. *ACS Chem Neurosci* **8**, 2778–2790 (2017).
5. Tuck *et al.* Multimodal Imaging based on Vibrational Spectroscopies and Mass Spectrometry Imaging applied to Biological Tissue : A Multiscale and Multi-omics Review . Pre-Analytical Factors. *Anal Chem* **93**, 445–477 (2021).
6. Porta Siegel *et al.* Mass Spectrometry Imaging and Integration with Other Imaging Modalities for Greater Molecular Understanding of Biological Tissues. *Mol Imaging Biol* **20**, 888–901 (2018).
7. Brieu *et al.* Context-based interpolation of coarse deep learning prediction maps for the segmentation of fine structures in immunofluorescence images. in *Medical Imaging 2018: Digital Pathology* (eds. Gurcan *et al.*) 24 (SPIE, 2018). doi:10.1117/12.2292794.
8. Carlred *et al.* Probing amyloid- β pathology in transgenic Alzheimer’s disease (tgArcSwe) mice using MALDI imaging mass spectrometry. *J Neurochem* **138**, 469–478 (2016).
9. Kaya *et al.* Histology-Compatible MALDI Mass Spectrometry Based Imaging of Neuronal Lipids for Subsequent Immunofluorescent Staining. *Anal Chem* **89**, 4685–4694 (2017).
10. Kaya *et al.* Delineating Amyloid Plaque Associated Neuronal Sphingolipids in Transgenic Alzheimer’s Disease Mice (tgArcSwe) Using MALDI Imaging Mass Spectrometry. *ACS Chem Neurosci* **8**, 347–355 (2017).
11. Dufresne *et al.* Multimodal detection of GM2 and GM3 lipid species in the brain of mucopolysaccharidosis type II mouse by serial imaging mass spectrometry and immunohistochemistry. *Anal Bioanal Chem* **409**, 1425–1433 (2017).
12. Prade *et al.* De novo discovery of metabolic heterogeneity with immunophenotype-guided imaging mass spectrometry. *Mol Metab* **36**, 100953 (2020).
13. Rappez *et al.* SpaceM reveals metabolic states of single cells. *Nat Methods* **18**, 799–805 (2021).
14. Mebius *et al.* Structure and function of the spleen. *Nat Rev Immunol* **5**, 606–616 (2005).
15. Bellomo *et al.* Macrophage-fibroblast circuits in the spleen. *Immunol Rev* **302**, 104–125 (2021).
16. Matsuda *et al.* The immunosenescence-related gene Zizimin2 is associated with early bone marrow B cell development and marginal zone B cell formation. *Immunity and Ageing* **12**, 1–12 (2015).
17. Recher *et al.* B cell-intrinsic deficiency of the Wiskott-Aldrich syndrome protein (WASp) causes severe abnormalities of the peripheral B-cell compartment in mice. *Blood* **119**, 2819–2828 (2012).
18. Grélard *et al.* Esmraldi: efficient methods for the fusion of mass spectrometry and magnetic resonance images. *BMC Bioinformatics* **22**, 56 (2021).
19. Jones *et al.* Imaging mass spectrometry statistical analysis. *J Proteomics* **75**, 4962–4989 (2012).
20. Yan *et al.* Cell-Type-Specific Metabolic Profiling Achieved by Combining Desorption Electrospray Ionization Mass Spectrometry Imaging and Immunofluorescence Staining. *Anal Chem* **92**, 13281–13289 (2020).
21. Rujchanarong *et al.* Defining the Tumor Microenvironment by Integration of Immunohistochemistry and Extracellular Matrix Targeted Imaging Mass Spectrometry. *Cancers (Basel)* **13**, 4419 (2021).
22. Hu *et al.* Emerging Computational Methods in Mass Spectrometry Imaging. *Advanced Science* **9**, 2203339 (2022).

- 1 23. Mandrekar. Receiver Operating Characteristic Curve in Diagnostic Test Assessment. *Journal of Thoracic Oncology*
2 **5**, 1315–1316 (2010).
- 3 24. Kune *et al.* Rapid Visualization of Chemically Related Compounds Using Kendrick Mass Defect As a Filter in Mass
4 Spectrometry Imaging. *Anal Chem* **91**, 13112–13118 (2019).
- 5 25. McInnes *et al.* UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. (2018).
- 6 26. Jain *et al.* *Algorithms for Clustering Data*. (Prentice-Hall, Inc., 1988).
- 7 27. Lloyd. Least squares quantization in PCM. *IEEE Trans Inf Theory* **28**, 129–137 (1982).
- 8 28. Eisenstein. Seven technologies to watch in 2023. *Nature* **613**, 794–797 (2023).
- 9 29. Soltwisch *et al.* Mass spectrometry imaging with laser-induced postionization. *Science (1979)* **348**, 211–215
10 (2015).
- 11 30. Heiles *et al.* Atmospheric-Pressure MALDI Mass Spectrometry Imaging at 213 nm Laser Wavelength. *J Am Soc*
12 *Mass Spectrom* **31**, 326–335 (2020).
- 13 31. Alexandrov. Spatial Metabolomics and Imaging Mass Spectrometry in the Age of Artificial Intelligence. *Annu Rev*
14 *Biomed Data Sci* **3**, 61–87 (2020).
- 15 32. Beyersdorf *et al.* Sphingomyelin breakdown in T cells: role in activation, effector functions and immunoregulation.
16 *Biol Chem* **396**, 749–758 (2015).
- 17 33. Gassert *et al.* Induction of Membrane Ceramides: A Novel Strategy to Interfere with T Lymphocyte Cytoskeletal
18 Reorganisation in Viral Immunosuppression. *PLoS Pathog* **5**, e1000623 (2009).
- 19 34. Peest *et al.* S1P-lyase independent clearance of extracellular sphingosine 1-phosphate after dephosphorylation
20 and cellular uptake. *J Cell Biochem* **104**, 756–72 (2008).
- 21 35. Alyamani *et al.* Alkaline sphingomyelinase (NPP7) impacts the homeostasis of intestinal T lymphocyte
22 populations. *Front Immunol* **13**, 1050625 (2022).
- 23 36. Huang *et al.* Eicosadienoic acid differentially modulates production of pro-inflammatory modulators in murine
24 macrophages. *Mol Cell Biochem* **358**, 85–94 (2011).
- 25 37. Liu *et al.* Functions of CD169 positive macrophages in human diseases (Review). *Biomed Rep* **14**, 26 (2020).
- 26 38. de Souza *et al.* Multiplex protein imaging in tumour biology. *Nat Rev Cancer* **24**, 171–191 (2024).
- 27 39. Einhaus *et al.* High-multiplex tissue imaging in routine pathology—are we there yet? *Virchows Archiv* **482**, 801–
28 812 (2023).
- 29 40. Bouschen *et al.* Matrix vapor deposition/recrystallization and dedicated spray preparation for high-resolution
30 scanning microprobe matrix-assisted laser desorption/ionization imaging mass spectrometry (SMALDI-MS) of
31 tissue and single cells. *Rapid Communications in Mass Spectrometry* **24**, 355–364 (2010).
- 32 41. Kompauer *et al.* Atmospheric pressure MALDI mass spectrometry imaging of tissues and cells at 1.4- μ m lateral
33 resolution. *Nat Methods* **14**, 90–96 (2016).
- 34 42. Palmer *et al.* FDR-controlled metabolite annotation for high-resolution imaging mass spectrometry. *Nat Methods*
35 **14**, 57–60 (2016).
- 36 43. Bokhart *et al.* MSiReader v1.0: Evolving Open-Source Mass Spectrometry Imaging Software for Targeted and
37 Untargeted Analyses. *J Am Soc Mass Spectrom* **29**, 8–16 (2018).
- 38 44. Hoo *et al.* What is an ROC curve? *Emergency Medicine Journal* **34**, 357–359 (2017).
- 39 45. Otsu. A Threshold Selection Method from Gray-Level Histograms. *IEEE Trans Syst Man Cybern* **9**, 62–66 (1979).
- 40
- 41