TU Delft

Delft University of Technology

msiFlow: automated workflows for reproducible and scalable multimodal mass spectrometry imaging and microscopy data analysis

Spangenberg, Philippa; Bessler, Sebastian; Thiebes, Stephanie; Migas, Lukasz G.; Kasarla, Siva Swapna; Kleesiek, Jens; Van de Plas, Raf; Shevchuk, Olga; Engel, Daniel R.; More Authors

DOI
10.1038/s41467-024-55306-7

Publication date
2025

Document Version
Final published version

Published in
Nature Communications

Citation (APA)
Spangenberg, P., Bessler, S., Thiebes, S., Migas, L. G., Kasarla, S. S., Kleesiek, J., Van de Plas, R., Shevchuk, O., Engel, D. R., & More Authors (2025). msiFlow: automated workflows for reproducible and scalable multimodal mass spectrometry imaging and microscopy data analysis. *Nature Communications*, *16*(1), Article 1065. https://doi.org/10.1038/s41467-024-55306-7

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.
For technical reasons the number of authors shown on this cover page is limited to a maximum of 10.

nature communications

Article

https://doi.org/10.1038/s41467-024-55306-7

msiFlow: automated workflows for reproducible and scalable multimodal mass spectrometry imaging and microscopy data analysis

Received: 2 August 2024

Accepted: 8 December 2024

Published online: 27 January 2025

Check for updates

Philippa Spangenberg[1], Sebastian Bessler[2], Lars Widera[1], Jenny Bottek[1], Mathis Richter[3], Stephanie Thiebes[1], Devon Siemes[1], Sascha D. Krauß[1], Lukasz G. Migas[4,5], Siva Swapna Kasarla[6], Prasad Phapale[6], Jens Kleesiek[7], Dagmar Führer[8], Lars C. Moeller[8], Heike Heuer[8], Raf Van de Plas[4,5,9], Matthias Gunzer[1,6], Oliver Soehnlein[3], Jens Soltwisch[2], Olga Shevchuk[1], Klaus Dreisewerd[2] & Daniel R. Engel[1]

Multimodal imaging by matrix-assisted laser desorption ionisation mass spectrometry imaging (MALDI MSI) and microscopy holds potential for understanding pathological mechanisms by mapping molecular signatures from the tissue microenvironment to specific cell populations. However, existing software solutions for MALDI MSI data analysis are incomplete, require programming skills and contain laborious manual steps, hindering broadly applicable, reproducible, and high-throughput analysis to generate impactful biological discoveries. Here, we present msiFlow, an accessible open-source, platform-independent and vendor-neutral software for end-to-end, high-throughput, transparent and reproducible analysis of multimodal imaging data. msiFlow integrates all necessary steps from raw data import to analytical visualisation along with state-of-the-art and self-developed algorithms into automated workflows. Using msiFlow, we unravel the molecular heterogeneity of leukocytes in infected tissues by spatial regulation of ether-linked phospholipids containing arachidonic acid. We anticipate that msiFlow will facilitate the broad applicability of MSI in multimodal imaging to uncover context-dependent cellular regulations in disease states.

The cellular heterogeneity in tissues has a reciprocal and decisive influence on the microenvironment and enables a balance between homoeostasis and inflammation. Intra- and intercellular communication is key in both tissue homoeostasis and inflammation, and lipids are emerging as critical regulators and key molecules in these processes[1]. Moreover, the lipid landscape was recently defined as a feature of immune cell identity[2]. Upon inflammation, neutrophils, the most abundant circulating white blood cell subset, readily infiltrate

[1]Institute for Experimental Immunology and Imaging, University Hospital Essen, Essen, Germany. [2]Institute of Hygiene, University of Münster, Münster, Germany. [3]Institute of Experimental Pathology (ExPat), Center of Molecular Biology of Inflammation (ZMBE), University of Münster, Münster, Germany. [4]Mass Spectrometry Research Center, Vanderbilt University, Nashville, TN, USA. [5]Delft Center for Systems and Control, Delft University of Technology, Delft, The Netherlands. [6]Leibniz-Institut für Analytische Wissenschaften – ISAS – e.V., Dortmund, Germany. [7]Institute for AI in Medicine (IKIM), University Hospital Essen, Essen, Germany. [8]Department of Endocrinology, Diabetes and Metabolism, University Hospital Essen, Essen, Germany. [9]Department of Biochemistry, Vanderbilt University, Nashville, TN, USA. e-mail: engel@immunodynamics.de

into and navigate through the tissue. Herein, they respond to a plethora of signals from the microenvironment by synthesis of lipid messengers, such as arachidonic acid (AA) and the oxidised metabolites prostaglandins and leukotrienes, influencing recruitment, phenotype and function of neutrophils[3–5]. In urinary tract infection (UTI), the third most common bacterial infection in humans[6,7] induced by uropathogenic *Escherichia coli* (UPEC)[8], neutrophils and urothelial cells are critical cell populations which provide the first immunological barrier for the containment of infection[9–11]. Data on lipidomic adaptations of urothelial cells and neutrophils in UTI are missing, as algorithms assigning lipids to specific cell populations in tissues are scarce, hindering insights into the decisive role of lipids in regulating mechanisms of inflammation and resolution.

In order to decipher the cellular interplay and the behaviour of specific cell populations in tissues, methods are required enabling in-depth lipidomic profiling with spatial resolution in the micrometre range. Immunofluorescence microscopy (IFM) efficiently determines the distribution of various cell types in tissue niches with high spatial resolution[12]. In contrast, matrix-assisted laser desorption ionisation mass spectrometry imaging (MALDI MSI) with laser-induced postionisation (MALDI-2) provides a label-free technology to investigate the spatial distribution of a large number of lipids and metabolites that are predominantly inaccessible by IFM[13]. Moreover, transmission-mode MALDI-2 (t-MALDI-2) for high-resolution imaging was introduced recently[14,15]. Due to differences in the optical setup, this technique can achieve a pixel size of down to 1 μm whereas standard MALDI MSI can reach 5–20 μm pixel size. For lipids, t-MALDI-2 MSI shows a high signal-to-noise ratio (SNR) while maintaining the depth of analysis (e.g. sensitivity and ion yield) of regular MALDI-2 MSI[14]. Thus, combining IFM and MALDI MSI would enable the assignment of the spatial lipidome to specific cell populations. A recent study integrated multiplex IFM and MSI to map myeloid heterogeneity in its metabolic and cellular context[16]. However, analysis of high-dimensional MALDI MSI data and image co-registration to IFM remains challenging due to the lack of algorithms and complete workflows, that allow reproducible and automated pre-processing, analysis and visualisation of MALDI MSI data[17,18]. As technology advances to achieve higher resolutions, data sizes are increasing, thereby further complicating data handling. Therefore, most commercial software solutions offer a user interface to pre-processed data with reduced size. Although this enables interactive data visualisation and analysis, it offers limited transparency and control over data pre-processing and data quality. In contrast to commercial software, which often remains a black box for users, existing open-source software offers incomplete solutions, as it is mostly designed for specific tasks (e.g. individual pre-processing steps, image registration, analysis or visualisation) and often requires programming skills or contains laborious manual steps (e.g. manually selecting off-/on-tissue regions)[18–24]. As a result, customised data analysis pipelines are constructed from a pool of open-source packages, in-house developed or commercially available software, hindering reproducible and high-throughput analyses and deterring non-expert users.

In this study, we aimed to bridge this gap by developing msiFlow, an open-source software that integrates all steps from import and pre-processing of raw multimodal and multi-vendor imaging data to registration, analysis and visualisation in automated workflows. The workflows operate fully automatically on all major operating systems. By employing msiFlow to a clinically relevant proof-of-concept study, we provide insights into the spatial lipidomic interactome in UTI, revealing a hitherto unknown heterogeneity of neutrophils important for the immune response against invading pathogens.

Results

The msiFlow software

High-dimensional molecular imaging through MALDI MSI holds great potential for comprehensive spatial mapping of the cellular heterogeneity in tissues and deciphering complex molecular interactions within the tissue microenvironment. However, the lack of open-source and easy-to-use software for automated MSI data processing and analysis greatly complicates reproducible and precise mapping of molecular landscapes in situ. To solve this problem, we integrated, optimised and further developed existing bioinformatic methods for data pre-processing, registration, analysis and visualisation into msiFlow, a collection of automated Snakemake workflows enabling reproducible and scalable analyses[25]. We applied msiFlow using a correlative imaging approach consisting of high-resolution (t)-MALDI-2 MSI and IFM in an experimental model of UTI. For this purpose, consecutive bladder sections of 8 μm from female mice were measured by IFM, t-MALDI-2 MSI and MALDI-2 MSI with a pixel size of 0.2 μm, 2 μm and 5 μm, respectively (Fig. 1a). t-MALDI-2 data were measured by orbitrap and MALDI-2 MSI by time-of-flight (TOF). The generated multimodal imaging data were processed and analysed by msiFlow.

msiFlow contains seven Snakemake workflows for pre-processing, registration, segmentation and analysis/visualisation. We have deliberately divided the software into seven main workflows to make the application highly flexible and modular. This modular software design enables easy integration of individual workflows of msiFlow into existing analysis pipelines. It is also possible to combine multiple workflows of msiFlow into one workflow via Snakemake. All workflows are integrated into a Docker[26] image enabling easy-to-use execution on all major operating systems. Each workflow can be run fully automatically through one command in the terminal. Parameters used by msiFlow are defined in one configuration file and can be adjusted by the user depending on the instrument's setting (e.g. mass and spatial resolution) and preferred methods. msiFlow also provides a browser-based interface to adjust the parameters and run the workflows (Supplementary Fig. 1). A detailed description of all parameters and the configurations used in this study for MSI pre-processing is provided in Supplementary Table 1 and on GitHub for all workflows.

msiFlow includes an MSI pre-processing workflow which imports raw MSI files from different vendors (Bruker and Thermo Fisher), processes all files in parallel, and outputs the processed data in the open standard imzML format along with quality control visualisations (Fig. 1b). The workflow contains steps for spectral smoothing, peak picking, peak alignment, matrix removal, peak filtering, normalisation, outlier removal and de-isotoping to generate an endogenous/tissue-origin mono-isotopic peak list (see method details in section "Methods"). We established two UMAP-based clustering approaches to automatically identify and remove off-tissue/matrix pixels and outliers enhancing data quality for subsequent analysis. To this end MALDI-2 MSI data of each sample were reduced to two dimensions by UMAP[27] followed by HDBSCAN clustering[28] (Supplementary Fig. 2a). The cluster connected to most border points of the measured area was considered the off-tissue cluster (Supplementary Fig. 2b). Clusters with high correlation to the off-tissue cluster were combined to an extended off-tissue cluster (Supplementary Fig. 2c) and post-processed to remove isolated objects and fill holes (Supplementary Fig. 2d). With this unsupervised approach we prevent the need of defining known matrix peaks beforehand which are often not consistent throughout all datasets. However, msiFlow also implements a supervised approach based on known matrix peaks. For identification of outliers, data of all samples were reduced to two dimensions by UMAP followed by HDBSCAN clustering (Supplementary Fig. 3a) and sample-specific clusters (SSC) were identified in which 70% of pixels originate from one sample (Supplementary Fig. 3b). Finally, samples in which most pixels were SSC pixels were considered sample outliers (Supplementary Fig. 3c, d).

In addition to MSI pre-processing, msiFlow provides a workflow for IFM pre-processing and image co-registration to combine MSI and IFM data. In the pre-processing workflow, noise is removed from the images by rolling-ball background subtraction and Gaussian smoothing followed by contrast normalisation (Fig. 1b bottom). Here very high/low intensities are suppressed to enhance image contrast. After

https://doi.org/10.1038/s41467-024-55306-7

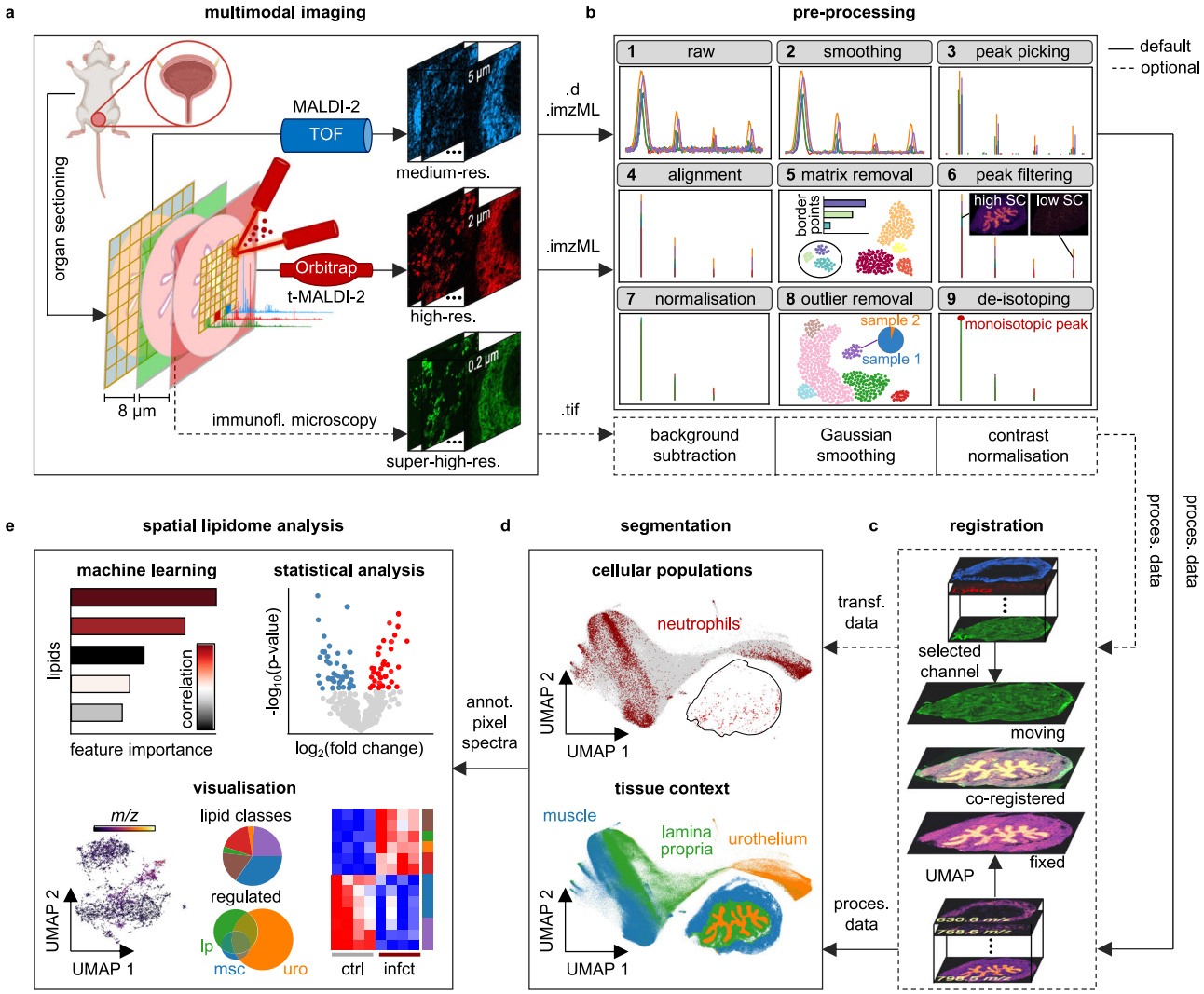


Fig. 1 | Workflow for correlative multimodal mass spectrometry imaging and immunofluorescence microscopy data. a Consecutive mouse bladder sections of 8 µm are measured by immunofluorescence (immunofl.) microscopy (green), matrix-assisted laser desorption ionisation combined with laser-induced post-ionisation mass spectrometry imaging (MALDI-2 MSI) (blue) and transmission-mode MALDI-2 MSI (red) with a lateral pixel size of 0.2 µm, 5 µm and 2 µm, respectively. MALDI-2 MSI data are measured by time-of-flight (TOF) and t-MALDI-2 data by orbitrap. Partially created in BioRender. Engel, D. (2024) https://BioRender.com/a89e002. **b** The pre-processing workflow integrates raw files in the common data formats as input (1) and performs signal smoothing (2), peak picking (3) and peak alignment (4) as first steps. A UMAP-based clustering approach determines matrix/off-tissue clusters which are connected to most pixels of the border of the measured area (5). Ions with low spatial coherence (SC), a measure for ion's informativeness, are removed (6), followed by intensity normalisation (7). UMAP clusters in which most pixels originate from one sample are considered outliers (8). Finally, de-isotoping generates a mono-isotopic peak list (9). The software contains an optional step for IFM pre-processing and image registration to combine MSI and IFM. The IFM pre-processing step removes noise from the microscopy images by background subtraction and Gaussian smoothing, followed by contrast normalisation. **c** In registration, a selected channel from IFM is used as moving image and the UMAP image from MSI is used as fixed image. The learned transformation is applied to all other IFM image channels, which results in transformed data (transf. data). **d** Imaging data are segmented and visualised by UMAP. Segmentation of the Ly6G IFM image indicates the distribution of neutrophils (top). The MSI data are segmented into its main tissue regions (bottom). **e** Lipidomic signatures of segmented regions and cell populations are extracted by machine learning-based classification and correlation (top left) and statistically analysed (top right). Lipidomic changes are visualised by UMAP (bottom left), Venn diagrams (bottom middle), pie charts (bottom middle) and heatmaps (bottom right). Processed data (proces. data), annotated pixel spectra (annot. pixel spectra), lamina propria (lp), muscle (msc), urothelium (uro), control (ctrl), infected (infct).

data pre-processing, MALDI-2 MSI data are combined with the IFM data through image co-registration in which a transformation aligns a moving to a fixed image (Fig. 1c). Several methods have been developed to generate one image out of a MALDI MSI dataset to spatially visualise molecular differences and similarities[29]. Here we used UMAP to reduce the MALDI-2 MSI data to one dimension. Through this approach, we receive one value for each pixel spectrum which can be visualised as a greyscale image. This UMAP image represents the main tissue structure. From IFM we used the autofluorescence (AF) image as it similarly represents the main tissue structure. Registration to the MALDI-2 MSI data was performed using symmetric normalisation implemented in the Advanced Normalisation Tools (ANTs) library[30]. To account for tissue deformations between the consecutive sections, the workflow uses rigid, affine and deformable transformation with mutual information as the optimisation metric.

From the pre-processed (and registered) data, regions of interest (ROIs) are extracted through two segmentation workflows (one for MSI and one for IFM data) (Fig. 1d). The MSI segmentation workflow includes state-of-the-art dimensionality reduction methods (PCA, t-SNE[31], UMAP) and clustering algorithms (k-means, spatial k-means[32],

HDBSCAN, hierarchical, gaussian mixture models) to perform unsupervised segmentation which extracts the main tissue context. The IFM segmentation workflow uses a thresholding-based approach to segment specific markers.

For analysing the ROIs msiFlow contains three analysis workflows. The first analysis workflow is designed to identify and compare molecular changes in different ROIs (e.g. tissue regions) between two groups by applying statistical analysis. The workflow outputs volcano plots, pie charts, Venn diagrams and heatmaps of the regulated lipids (Fig. 1e right). The second analysis workflow unravels molecular signatures of the ROIs (e.g. cell populations) by using machine learning-based classification (e.g. tree-based classifiers such as AdaBoost, LightGBM[33], XGBoost[34]), explainable AI methods (e.g. shapely additive values (SHAP))[35,36] and correlation (e.g. Spearman and Pearson) (Fig. 1e top left). A prior instance of retrieving molecular signatures by SHAP in a multimodal study employing MSI and IFM can be found in Esselman et al.[37]. The third analysis workflow applies a UMAP-based clustering approach to reveal molecular heterogeneity in the ROIs (e.g. cell populations) and plots the heterogeneous lipid signals in UMAPs (Fig. 1e bottom left). In conclusion, msiFlow provides an easy-to-use, vendor-neutral and platform-independent software for automated, end-to-end multimodal image analysis, which significantly complements existing open-source software, as depicted in Supplementary Table 2.

Spatial lipidomic changes in the infected urinary bladder

We applied msiFlow to MALDI MSI data of the clinically relevant model of UTI to reveal niche-specific molecular changes upon infections with UPEC. To segment the urinary bladder into the main tissue compartments we applied dimensionality reduction by UMAP followed by HDBSCAN clustering with manual refinement to the MALDI MSI data. To indicate the molecular changes in the different tissue regions, MS spectra from control and infected samples were visualised and compared by the 2D UMAP embedding (Fig. 2a). Annotation of the spectra to specific tissue areas, i.e. lamina propria (lp), muscle (msc) and urothelium (uro), indicates specific lipidomic signatures in the different tissue regions in control and infected samples (Fig. 2b). In addition, the UMAP visualisation indicates specific tissue layers, such as the urothelium, with strong changes upon infection. Spatial mapping of the UMAP representation visualises the different tissue regions (Fig. 2c). Next we generated tentative lipid annotations for the *m/z* values by using the bulk structure search from the LipidMaps Website (www.lipidmaps.org) and searched for lipid classes expected for MALDI-2 analysis in positive-ion mode (fatty acids/esters [FA], ceramides [Cer], sphingomyelins [SM], hexosyl ceramides [HexCer], triglycerols [TG], diglycerols [DG], glycerophosphocholines [PC], glycerophosphates [PA], glycerophosphoserines [PS], glycerophosphoethanolamines [PE], glycerophosphoglycerols [PG] and sterols [ST]) as $[M+H]^+$, $[M-H_2O+H]^+$, $[M+Na]^+$ and $[M+K]^+$ precursor ions with a mass tolerance of ±0.01 *m/z*. For further analysis, *m/z* signals without potential lipid matches were filtered out in order to eliminate non-lipid peaks (e.g. in-source fragments or chemical background). Then we compared the lipidomic changes separately in the urothelium, lamina propria and muscle. Therefore, we performed statistical analysis of the mean intensities of urothelial, lamina propria and muscle pixels of infected vs. control samples and performed MALDI MSI in data-dependent acquisition (DDA) mode to validate the lipid annotations (Supplementary Table 3). This analysis indicated the strongest lipidomic alterations in the urothelium, the site of bacterial tissue entry, across various lipid classes (Fig. 2d, e). Out of the 12 investigated lipid classes, the main altered classes included TG, DG, PC, ether-linked PC (PC O), PE and ST. Here TG, DG and PC were upregulated and PE and ST were downregulated in the infected urothelium (Fig. 2f, g). Among other lipids, TG 54:4 was exclusively expressed in the urothelium (Fig. 2h). In contrast, PC O-32:0 was not only expressed in the urothelium, but also in the lamina propria close to the urothelial expression of TG 54:4. High-resolution MSI (2 μm pixel size) by t-MALDI-2 revealed the distribution of the UPEC infection by PE 33:1 (704.52 *m/z*), an odd chain fatty acyl PE known to be highly expressed in *E. coli*[38] (Fig. 2h). The t-MALDI-2 measurement further indicated the presence of ramified cells (PC O-32:0) in the bladder tissue and phagocytosis of bacteria in TG 54:4-rich urothelial areas by those highly ramified cells.

Identification of lipidomic signatures of neutrophils

Among others, neutrophils are critical immune cells during bacterial infections and efficiently migrate into infected tissue areas to phagocytose bacteria[10]. We detected $Ly6G^+$ neutrophils in the lamina propria around blood vessels and in the infected urothelium suggesting directed migration towards the infection (Supplementary Fig. 4). Microscopy also indicated downregulation of CXCR2 in the urothelium, suggesting niche-specific desensitisation of this chemokine receptor at the site of infection (Supplementary Fig. 4).

To unravel the niche-specific lipidomic signature of neutrophils, we used our multimodal imaging workflow and performed IFM and MSI on the timsTOF fleX instrument on consecutive tissue sections 1 day after infection. Ly6G, a specific marker for neutrophils, was used to localise neutrophils across the tissue and actin, expressed in the muscle tissue, was used to demarcate the organ boundaries to the outside tissue areas and the lamina propria (Fig. 3a). For image co-registration the AF channel from IFM was used as moving and the UMAP representation from MSI as fixed image (Fig. 3b). The learned transformation was applied to all IFM image channels. The registration result was validated by the Jaccard index for the overlap between transformed urothelial mask from IFM and MSI (average 0.8 Jaccard). Registered Ly6G images were segmented to annotate the pixel spectra according to neutrophil-rich areas. Then a binary tree-based classifier was trained with the Ly6G-annotated pixel spectra to extract lipidomic signatures of neutrophils (Fig. 3c). The most important lipids to classify $Ly6G^{+/-}$ pixel spectra were several ether-linked PCs (Fig. 3d). For the classification model, lipids that are anti-correlated with neutrophils or present in very low abundance in neutrophils are as important as lipids which are highly correlated to neutrophils or present in very high abundance. To distinguish between lipids that are highly and lowly abundant in neutrophils, the bar plot is colour-coded according to Pearson's correlation (Fig. 3d). The lipid with the highest feature importance and correlation for neutrophils was PC O-36:4 (Fig. 3e). LC-MS/MS (Supplementary Fig. 5) and in situ MS/MS (Supplementary Fig. 6) indicated that PC O-36:4 contains the esterified AA.

To validate our results, we performed MALDI-2 MSI from bone marrow (BM)-derived neutrophils and applied our data pre-processing workflow (Supplementary Fig. 7a). To extract the main peaks from BM neutrophils, we filtered the peaks based on their spatial coherence and average intensity. The filtered peaks include the top 4 lipids (M + H) for bladder neutrophils. These lipids are expressed at a very high level in BM neutrophils, indicating that our imaging workflow identifies cell population-specific lipids (Supplementary Fig. 7b).

Lipidomic heterogeneity of neutrophils in infection

Neutrophils can adopt specific maturation states in secondary lymphoid organs such as the BM and spleen, thus establishing significant heterogeneity that is important for defence against pathogens. However, little is known about this heterogeneity in infected organs, although certain cellular states may represent specific adaptations to the tissue environment and infectious conditions. To reveal the lipidomic heterogeneity of neutrophils in the infected urinary bladder we used msiFlow. Here registered Ly6G images were first segmented and annotated according to the tissue region (urothelium, lamina propria and muscle) in which they are localised

https://doi.org/10.1038/s41467-024-55306-7

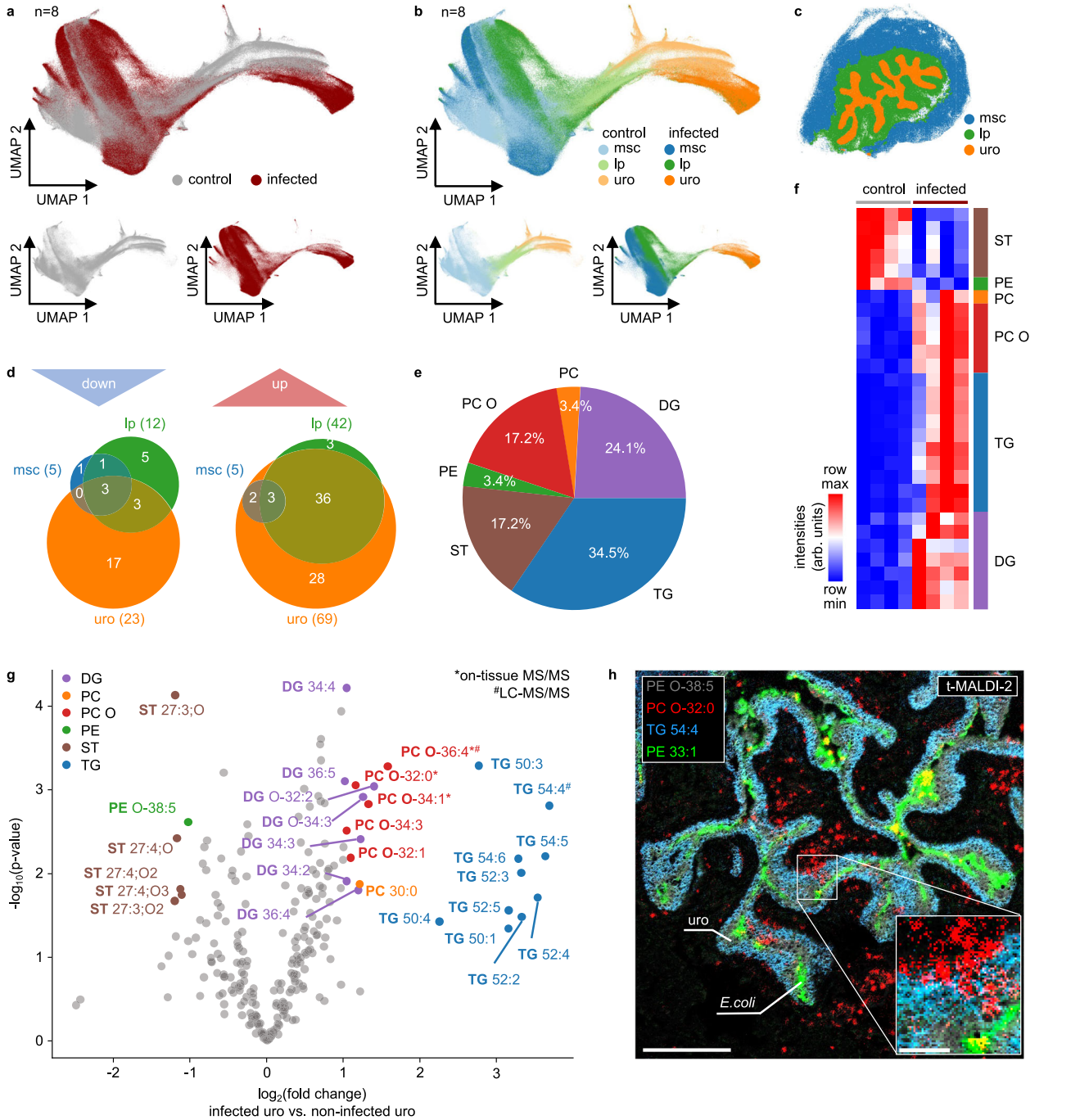


Fig. 2 | Significant lipidomic adaptations in the urothelium upon infection. **a** UMAP embedding of MALDI MSI data of 4 control (grey) and 4 infected (red) bladder sections ($n = 8$) from female mice. **b** UMAP embedding colour-coded according to the muscle (msc), lamina propria (lp) and urothelium (uro) in the control (light-coloured) and infected (dark-coloured) tissue sections. **c** Spatial mapping of the UMAP clusters shown in (**b**) in an infected tissue section. **d** Average spectra of the different tissue regions (msc, lp and uro) were extracted and statistically analysed. Venn diagrams revealing the number of significantly down-regulated (left) and upregulated (right) lipids (p value < 0.05 and $|\log_2(FC)| > 0.5$). **e** Pie chart showing the percentage of lipid classes regulated in the infected urothelium. **f** Heatmap showing the average intensity of regulated lipids (p value < 0.05 and $|\log_2(FC)| > 1$) in the urothelium for each sample. **g** The significantly regulated lipids (p value < 0.05 and $|\log_2(FC)| > 1$) in the urothelium are labelled in the Volcano plot. A full statistical summary including the exact p values is provided in the source data file. Lipids validated by on-tissue MS/MS are marked with *, and by LC-MS/MS are marked with #. **h** One infected bladder was measured by t-MALDI-2 MSI. Spatial distribution of regulated lipids and *E. coli* (PE 33:1) in the infected bladder measured by t-MALDI-2 MSI. Contrast was enhanced by percentile stretching using 0 as min and the 99th percentile as max value. Scale bars indicate 300 μm and 50 μm in the zoom. Urothelium (uro), lamina propria (lp), muscle (msc), diglyceride (DG), phosphatidylcholine (PC), ether-linked phosphatidylcholine (PC O), phosphatidylethanolamine (PE), sterol (ST), triglyceride (TG). Statistical significance was tested by two-sided standard t-test for normally distributed populations with equal variances and two-sided Welch's t-test for normally distributed populations with unequal variances. Non-normally distributed populations were statistically analysed by two-sided Wilcoxon rank-sum test. The Shapiro–Wilk test and Levene test were used to test for normal distribution and equal variances. $n = 8$ (4 control versus 4 infected). Source data are provided as a Source Data file.

https://doi.org/10.1038/s41467-024-55306-7

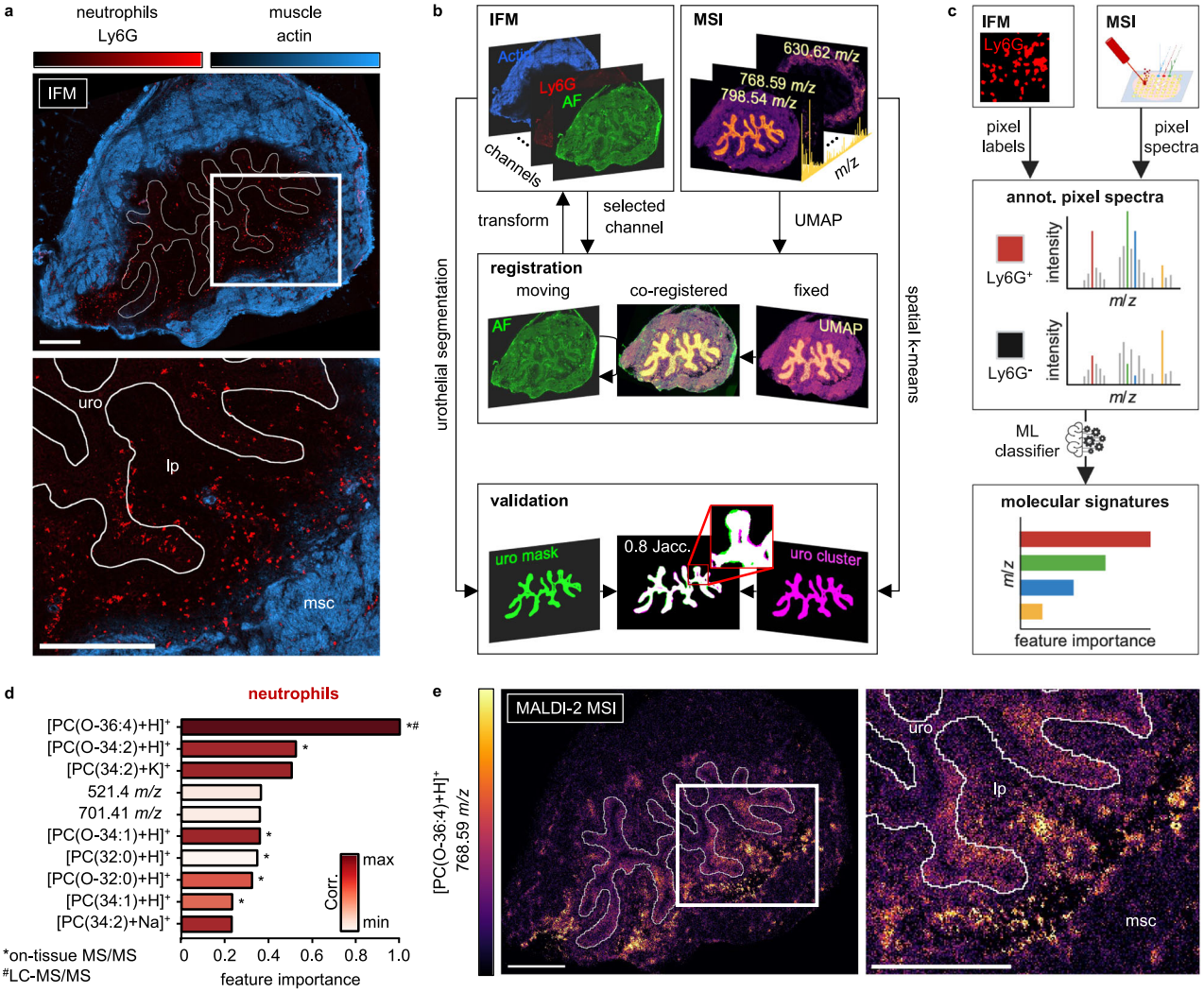
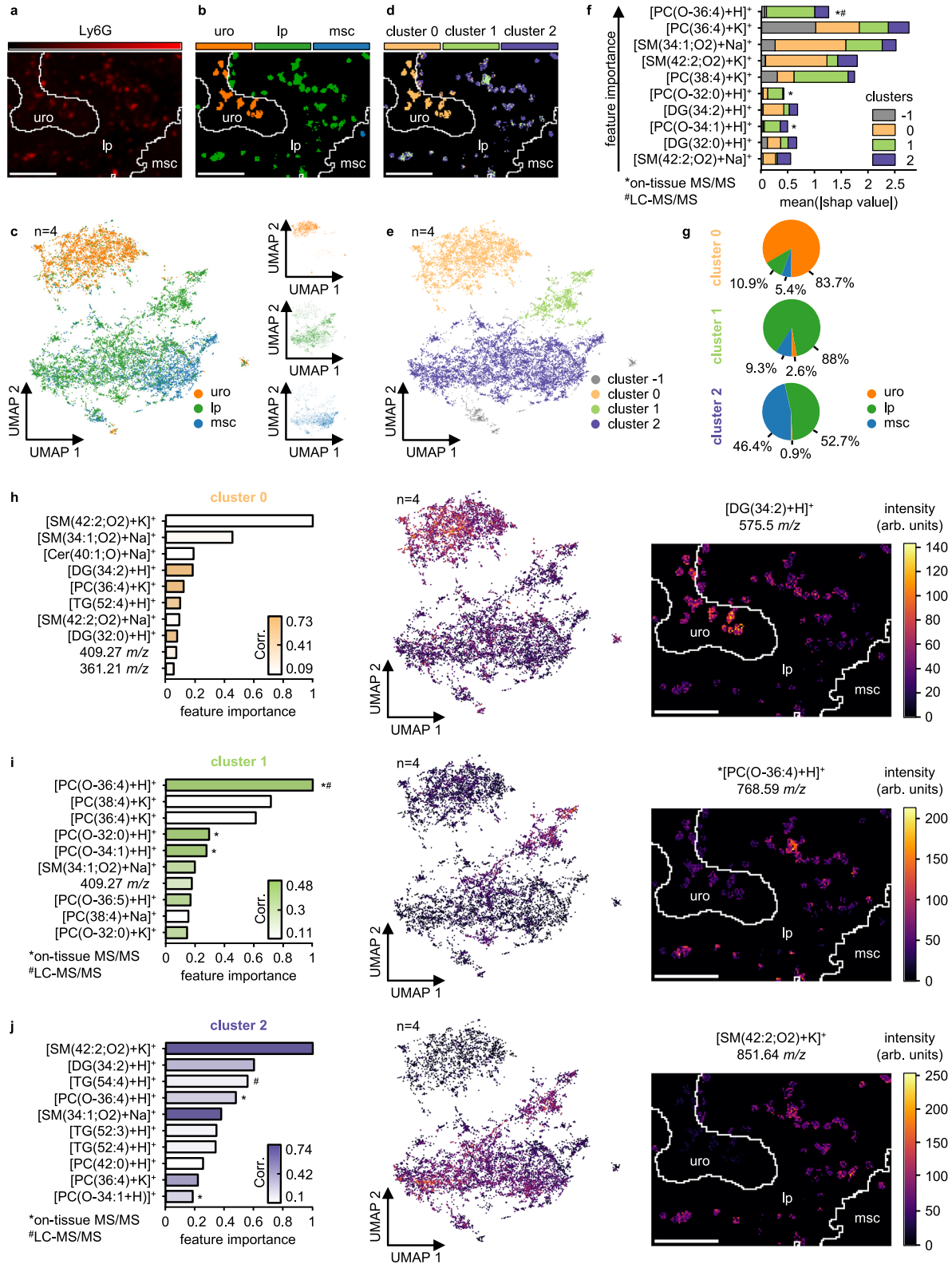


Fig. 3 | Image co-registration and machine learning-based classification extract spatial lipidomic signatures of neutrophils. a Immunofluorescence microscopy (IFM) of Ly6G for neutrophils (red) and actin for smooth muscle cells (blue) was performed on four consecutive bladder sections to the MALDI mass spectrometry imaging (MSI) sections from female mice. Colour bars indicate the scaled intensities (arb. units) ranging from 0 to 1 of Ly6G (red) and actin (blue). Contrast was enhanced by adjusting the min/max value in Fiji. **b** IFM images were registered to the MALDI MSI images. The autofluorescence (AF) image from IFM was used as moving image and the UMAP image from MSI served as fixed image. The learned transformation was applied to all IFM image channels. We validated the registration result by the Jaccard (Jacc.) index for the overlap between transformed urothelial mask from IFM (green) and MSI (magenta). We applied spatial k-means clustering on the MSI data and selected the cluster containing the urothelium. The urothelial mask from microscopy was manually created. **c** Segmentation of the registered Ly6G images was performed to annotate the pixel spectra into neutrophil/non-neutrophil pixels. Then a machine learning (ML)-based classification model was trained with the annotated (annot.) pixel spectra for neutrophils. Created in BioRender. Engel, D. (2024) https://BioRender.com/b17p939. **d** Feature importance ranking revealing the top 10 lipids which are most important for the classifier to distinguish neutrophil and non-neutrophil pixels. The bars are colour-coded according to Pearson's correlation coefficient (corr.) ranging from 0.03 (min) to 0.11 (max). Several ether-linked phospholipids were further characterised by on-tissue MS/MS (marked with *). The top lipid was also validated by LC-MS/MS (marked with #). **e** Ion image of the top lipid identified for neutrophils distribution on the consecutive section to the IFM section shown in (**a**). Colour bar indicates the scaled intensity (arb. units) of 768.59 *m/z* ranging from 0 to 1. Contrast was enhanced by percentile stretching using 0 as min and the 99.9th percentile as max value. Urothelium (uro), lamina propria (lp), muscle (msc). Scale bars indicate 500 μm. Source Data are provided as a Source Data file.

(Fig. 4a, b) followed by dimensionality reduction by UMAP (Fig. 4c). Next HDBSCAN clustering of MSI spectra of Ly6G$^+$ pixels was performed, generating three clusters (Fig. 4d, e). A classification model was trained on the clustered Ly6G$^+$ pixel data to identify the lipids with the strongest feature importance and SHAP values. We found that PC O-36:4 was the most important lipid for the classification model based on the feature importance (Fig. 4f). Mapping of the clusters to the different tissue compartments revealed that neutrophils in cluster 0 are mainly located in the urothelium by 83%, neutrophils in cluster 1 are mainly located in the lamina propria by 88% and neutrophils in cluster 2 are equally distributed in the lamina propria (52.7%) and muscle (46.4%) (Fig. 4g). In addition, binary classifiers were trained for each class which revealed class-specific lipidomic signatures (Fig. 4h–j). The distribution of the top lipids showed strong expression of DG 34:2 (575.5 *m/z*) in cluster 0 (Fig. 4h), PC O-36:4 (768.59 *m/z*) in cluster 1 (Fig. 4j) and SM 42:2;O2 (851.54 *m/z*) in cluster 2 (Fig. 4i). These data demonstrate the spatial expression of PC O-36:4, which contains the immunologically important metabolite AA, in neutrophils in different tissue niches in the urinary bladder. msiFlow resolved the lipidomic heterogeneity of neutrophils in infected tissue areas, shedding light on the tissue-specific lipidomic adaptations of neutrophils in infected tissues and providing possibilities for the adaptations of an immunological response in tissues.

https://doi.org/10.1038/s41467-024-55306-7



Discussion

In this study, msiFlow, a powerful open-source end-to-end software for automated pre-processing, analysis, visualisation and registration of MALDI MSI and IFM data was generated. Using msiFlow, we clarified the cell-specific lipidomic adaptations of urothelial cells and neutrophils in UTI at a high spatial resolution and identified specific ether-linked PCs and the AA metabolism by neutrophils in certain tissue niches.

With msiFlow, we address the current lack of complete, automated and open-source MSI software by integrating all necessary steps from raw data import to multivariate analysis and visual output as well as registration into msiFlow. All steps are automated in Snakemake workflows, which enables parallel data processing for high-throughput analyses. All workflows can be run via a single command in the terminal on all major operating systems and do not require complicated

Fig. 4 | Lipidomic heterogeneity of neutrophils. a Registered Ly6G image showing the distribution of neutrophils in the uro, lp and msc. Colour bar indicates the scaled intensity (arb. units) of Ly6G ranging from 0 to 1. **b** Segmented Ly6G image, colour-coded according to the different tissue regions. **c** UMAP of $Ly6G^+$ pixel spectra in four infected bladder samples ($n = 4$) from female mice colour-coded according to the different tissue regions. **d** HDBSCAN clustering was performed on the UMAP embedding of $Ly6G^+$ pixel spectra. The spatial distribution of the three main clusters is shown in the image of an infected bladder region. **e** HDBSCAN clusters shown in the UMAP embedding of $Ly6G^+$ pixels. Cluster −1 indicates noise pixels. **f** A XGBoost multi-class classifier was trained on the annotated clustered pixel spectra to extract the mean SHAP values of the 10 most important lipids based on the feature importance. The top 10 lipids are shown in ascending order according to their feature importance. **g** Distribution of cluster pixels in the different tissue regions. **h**–**j** A XGBoost binary classifier was trained for each cluster to reveal cluster-specific lipidomic signatures. The resulting top 10 lipids based on their feature importance are shown in the bar plots (left) and coloured according to their Pearson's correlation (corr.). Signal intensities of lipids with the highest feature importance and positive Pearson's correlation for cluster 0 (**h**), cluster 1 (**i**) and cluster 2 (**j**) shown in the UMAP (middle) and a spatial heatmap (right). Urothelium (uro), lamina propria (lp), muscle (msc). Scale bars indicate 200 μm. $n = 4$ (UPEC-infected). Source data are provided as a Source Data file.

package installations, as we incorporated all workflows into a Docker image, making it broadly applicable and suitable for non-programmers compared to common software which only runs on Windows (e.g. SCiLS Lab, LipoStar and msiQuant[39]) or requires programming skills (e.g. Cardinal). Nevertheless, msiFlow provides full flexibility to define all desired steps and preferred methods to be executed by specifying a manageable number of parameters in one configuration file. In addition, msiFlow is fully developed in the open-source programming language Python, as opposed to the widely used proprietary programming language MATLAB. Thus, msiFlow is more affordable and accessible than commercial or proprietary software developed in MATLAB[16,22,40]. Unlike most software that only accepts data in imzML format[21,22,39–42], msiFlow accepts raw timsTOF data in addition to imzML, preventing the need for additional software for data parsing and the dependence on manufacturer-specific software.

We have made special efforts to develop a complete pre-processing pipeline to filter the data into an endogenous/tissue-origin mono-isotopic peak list for downstream analysis. Simple peak picking is often not sufficient to filter out non-informative or exogenous signals (e.g. from matrix, isotopes and chemical noise) from raw spectra[43]. Such unfiltered data can lead to prediction models learning from non-biological meaningful signals. Therefore, our workflow includes a peak filtering step in addition to standard peak picking. Our workflow performs pixel-wise peak picking and utilises the spatial information for peak filtering by computing the spatial coherence metric[44], presenting a significant improvement over common MS software that uses single mass spectra without considering spatial information[45]. In addition, we have developed two UMAP-based clustering approaches that sufficiently determine matrix/off-tissue pixels and outliers in datasets fully automatically, which was previously often done manually (e.g. manual selection of specific regions in SCiLS Lab). The automated matrix detection approach is applicable for whole-sample measurements due to its reliance on tissue architecture and borders and the outlier detection approach is effective for multi-sample datasets as it relies on spectral similarity across multiple samples. Our workflow also provides visual results after each pre-processing step, offering transparency and the possibility of quality control compared to commercial software.

msiFlow not only enables end-to-end analysis of MSI data, but also registration with IFM for multimodal imaging, complementing existing open-source software which either does not include registration[42] or focuses on registration without data pre-processing and analysis[18]. In this study, we registered IFM to high-resolution MALDI MSI to account for the emerging interest in multimodal analyses. Typically, immunohistochemistry of tissue sections is registered to molecular data from MSI and annotated by pathologists[46,47]. However, registration with IFM enables the precise localisation of specific cell populations to characterise the phenotypical and molecular heterogeneity of cells in the tissue context. We deliberately used consecutive sections as the washing steps of the staining procedure of cell-specific molecules strongly reduce the signal-to-noise of the spatial lipidome. Thus, our approach avoids this problem by registering imaging data from consecutive tissue sections. However, we acknowledge limitations in achieving complete cell-specificity due to the usage of consecutive tissue sections and lateral diffusion of lipids from surrounding cells during MSI sample preparation. While our automated registration approach might be less accurate than the popular landmark-based registration selecting features present in both modalities[16,48] or laser ablation marks from MSI[18], it does not require manual steps and is not dependent on a post-acquisition pattern. Furthermore, our MSI data of BM-derived neutrophils revealed a substantial abundance of the identified lipids, validating the adequacy of our image registration, segmentation and feature extraction approach for identifying molecular signals associated with particular cells within tissue sections to a certain degree.

So far msiFlow does not include an annotation or interpretation step, as we see the main aim of the software in identifying interesting molecular patterns and preliminary candidates which can be validated by MS/MS in the next step. However, as msiFlow provides a mono-isotopic peak list and contains a mass alignment procedure, tentative lipid annotations for interesting candidates can be made through mass matching to databases such as LipidMaps, as presented in this study. Due to the modular software design, an annotation routine as well as available software for biological interpretation[49] can easily be added in the future. Moreover, we anticipate that msiFlow can be scaled to other MSI modalities, such as metabolomics and proteomics.

To demonstrate the strength of our multimodal imaging approach and msiFlow to biological research, we applied msiFlow to lipidomic MSI and IFM data of UPEC-infected bladders. This revealed high abundance of ether-linked PCs in neutrophils which is in line with previous studies[2,50,51]. Polyunsaturated fatty acids (PUFAs) containing lipids, such as PC O-36:4, strongly influence cellular function via effects on membrane properties, and by acting as a precursor pool for lipid mediators, such as AA. Among others, these mediators are enzymatically converted by cyclooxygenase and lipoxygenase to generate eicosanoids, such as prostaglandins and leukotrienes[4]. These metabolites are critical for neutrophil migration and swarming in the three-dimensional tissue context[52]. Recently, changes in the composition of PCs were observed during neutrophil differentiation stages, indicating reduced abundance of this PUFA in mature stages of neutrophils[2]. Similarly, we found that PC O-36:4 was most abundant in neutrophils after extravasation in the lamina propria, but not detectable at the site of infection in the urothelium, suggesting metabolic oxidation of AA into leukotrienes to facilitate neutrophil swarming towards the local infection in the urothelium. Besides AA oxidation, the reduced abundance of PC O-36:4 in the urothelium could also indicate neutrophil apoptosis as it is assumed that ether-linked PCs are important membrane components required for neutrophil survival[53]. In addition, the location of UPEC, visualised by PE 33:1 in our t-MALDI-2 data (Fig. 2h), may play a crucial role for the function of uroepithelial cells and neutrophils. In particular, lipopolysaccharides from UPEC are recognised by Toll-like receptor (TLR) 4 expressed by both neutrophils and uroepithelial cells[54]. Upon sensing through TLRs, uroepithelial cells secrete a wide range of soluble factors, including the antimicrobial peptides cathelicidin and β-defensin, as well as chemokines, such as CXC-chemokine ligand 1 (CXCL1), the latter regulating neutrophil

migration into the urothelium[10]. Thus, migration of neutrophils, phagocytosis of UPEC and the crosstalk of uroepithelial cells with neutrophils could result in significant changes in the lipid composition, which we have unravelled by msiFlow. Taken together, msiFlow revealed a heterogeneous distribution of lipids, such as PC O-36:4 in neutrophils, indicating a critical and spatial role of certain lipid species for neutrophil state and function.

With the advent of our established software and the high-resolution t-MALDI-2 MSI technologies, we found an increased expression of molecules in the population of urothelial cells important for mucus production and modulation of the bacterial colonisation, such as PCs, DGs and TGs[55,56]. Our t-MALDI-2 MSI data revealed an accumulation of TGs in the urothelium close to the bacteria. Interestingly, TGs were also among the lipids with the highest feature importance and correlation to neutrophils in the urothelium. TGs are stored in lipid droplets (LDs), which have been described as sites of eicosanoid synthesis[57]. It is also known that defence molecules, such as Cathelicidin (CAMP) and histones, accumulate on the LDs of challenged cells which makes them more resistant to bacterial species including *E. coli*[58], potentially through the mechanism of neutrophil extracellular traps (NETs)[58]. Hence our findings suggest an important role of LDs in supporting the immune defence against UPECs in UTI through eicosanoid synthesis and NET formation.

In conclusion, we established an open-source, platform-independent and vendor-neutral software for automated, end-to-end, transparent, reproducible and scalable multimodal MSI and IFM image analysis with a low barrier to entry. Using msiFlow, we not only identified lipidomic signatures of neutrophils, validated by MSI of BM neutrophils and in line with previous studies, but also unravelled a hitherto unknown lipidomic heterogeneity of neutrophils in UTI. This is an essential step towards uncovering the context-dependent regulation of leucocytes in inflammatory conditions. The easy usability and completeness of our software will facilitate the applicability of MSI in the emerging field of multimodal imaging.

Methods

Ethics statement

The animal study was approved by LANUV, Recklinghausen, Germany. The study was conducted in accordance with the local legislation and institutional requirements.

Animal studies

Six- to eight-week-old female C57BL/6J mice were used throughout the experiments. We used female animals for this study, as UTI mostly affects women. Animals were purchased from Charles River Laboratories and maintained under specific-pathogen-free conditions in the central animal facility at the University Hospital Essen. The mice were kept at 20–24 °C and 45–65% humidity in a 12-h dark-light cycle. They were housed in stable social groups and provided with species-appropriate food and water, suitable bedding and nesting material as well as a shelter for rodents. The local review board (Bezirksregierung Köln, Landesamt für Natur, Umwelt und Verbraucherschutz NRW in Recklinghausen, Germany) approved the animal experiments.

Urinary tract infection model

UPEC strain 536 (O6:K15:H31) was cultured for 3 h at 37 °C in LB medium. The bacteria were harvested by centrifugation at 1500 × *g* for 20 min and then the OD(600) was measured. The bacteria were resuspended at a concentration of 10^{10} bacteria/mL in sterile PBS. A mixture of ketamine and xylazine 80/10 mg/kg body weight in 150 µl PBS was injected intraperitoneally to anesthetise the mice. The infection of the animals was induced by transurethral inoculation of 5×10^8 UPEC in 0.05 mL PBS using a soft polyethylene catheter.

IFM

UPEC-infected urinary bladders ($n = 4$) were fixed overnight in PLP buffer [pH 7.4, 0.05 M phosphate buffer containing 0.1 M L-lysine, 2 mg/mL sodium periodate and paraformaldehyde with a final w/v concentration of 1%], equilibrated in 30% sucrose for 24 h and stored at −80 °C. Bladder tissue was cut into 8 µm thick sections at −20 °C using a cryostat. Unspecific binding was blocked by incubation of the sections with PBS containing 1% BSA and 0.05% Triton X-100 for 1 h. Bladder sections were mounted into MACSwell sample carriers, blocked using a blocking buffer containing 10% BSA and 2% goat serum for 1 h at RT before nuclei were counterstained using DAPI-staining solution (Miltenyi Biotec) according to the manufacturer's recommendations before being placed into a MACSima imaging system. Sections were then incubated with directly conjugated antibodies against Ly6G (1A8, Miltenyi Biotec, PE,1:50), EpCAM (REA977, Miltenyi Biotec, APC,1:50), SMA (REAL650, Miltenyi Biotec, FITC, 1:300) and CXCR2 (SA044G4, BioLegend, PE, 1:50). Acquired pictures were stitched using the pre-processing pipeline in MACS iQ View Analysis Software (Miltenyi Biotec) for downstream analysis.

MALDI MSI

Unfixed tissue sections of infected and control bladders ($n = 10$, 5 infected and 5 control) were thawed under a gentle stream of dry N_2 gas. For t-MALDI measurements, the tissue sections were washed with 250 µL of a 150 mM ammonium acetate solution to remove alkali metal salts and dried again under a gentle stream of dry N_2 gas. 2,5-DHAP matrix was applied onto the tissue section by sublimation in a home-build sublimation device described earlier by Bien et al.[59]. 1.5 mL of a 20 mg/mL solution of 2,5-DHAP in acetone was filled into the matrix reservoir and heated to about 120 °C which caused the acetone to evaporate. The sample was mounted to the cold side of the sublimation device and kept at around 4 °C at an approximate vertical distance of 6 cm above the matrix reservoir. The sublimation device was evacuated to a pressure of about 5×10^{-3} mbar and the sublimation was conducted for 10 min. The samples were measured immediately after sublimation.

All MALDI MSI measurements were carried out in positive-ion mode. MALDI-2 MSI data with a pixel size of 5 µm was acquired with a timsTOF fleX MALDI-2 (QTOF) instrument with microGRID extension (Bruker Daltonics, Bremen, Germany). The MALDI and postionisation laser were operated with a pulse repetition rate of 1 kHz and a delay of 10 µs. The ablation laser power was set to 80% with 25 shots per pixel. The ion detection range was set to m/z 300–1500.

The TIMS separated MS/MS measurement was acquired with a pixel size of 50 µm with 250 laser shots, a $1/k_0$ range from 1.4 to 1.8 with N_2 as collision gas, a ramp time of 250 ms, an isolation window of 1 Da and 30 eV collision energy.

The t-MALDI-2 MSI data with a pixel size of 2 µm was acquired with a setup which has been described in detail previously[14,60]. For this, a Q Exactive Plus Orbitrap mass spectrometer (Thermo Fisher Scientific, Bremen, Germany) equipped with a modified dual-ion funnel source (Spectroglyph, Kennewick, WA, USA), which enabled transmission-mode illumination and MALDI-2 postionisation, was used. Both lasers were operated with a pulse repetition rate of 100 Hz and a delay of 10 µs. A mass resolving power of 70,000 (defined for m/z 200) with a fixed injection time of 250 ms and an ion detection range of m/z 350–1500 were used. MALDI-DDA-MSI was used to confirm some of the tentatively annotated lipids with on-tissue MS/MS. For this a similar approach to the one used by Ellis et al.[61]. was chosen, which alternates full-scan and MS/MS pixel. For this front side illumination was used with a step size in x- and y-direction of 10 µm and 20 µm, respectively. For the full-scan pixels, an ion detection range of m/z 550–1500 was used. For DDA MS/MS measurements an isolation window of 1 Da and a fixed first mass of m/z 100 with an NCE of 25 were used. The exclusion time was set to 30 s with isotope exclusion enabled. The mass resolving power for both full-scan and MS/MS was set to 70,000 and the

injection time was fixed at 250 ms. The MALDI-DDA-MSI data were analysed using Lipostar MSI (vs. 1.3, Molecular Horizon, Bettona, Italy) and annotated using the Lipid Maps Structure Database (https://www.lipidmaps.org/databases/lmsd). The import and annotation settings for Lipostar MSI are shown in Supplementary Fig. 8.

Isolation of bone marrow-derived neutrophils

BM-derived neutrophils were obtained using a mouse Neutrophil Isolation Kit (Miltenyi Biotec, 130-097-658) following the manufacturer's instructions. In brief, the BM of a femur was flushed, and erythrocytes lysed using RBC Lysis Buffer (BioLegend, 420302; 1 min on ice). Cells were then sequentially incubated with Neutrophil Biotin-Antibody Cocktail and Anti-Biotin Microbeads before neutrophils were isolated using a negative selection on a magnetic column. The isolated neutrophils were then centrifuged onto slides using a cytospin centrifuge, fixed for 5 min using 4% formaldehyde, washed two times with 500 µL PBS followed by two washes with 500 µL of a 150 mM ammonium acetate solution to remove the PBS. For MALDI MSI matrix application was carried out as described above and the cells were measured with 50 µm pixel size on the timsTOF fleX MALDI-2 (QTOF).

LC-MS/MS

Lipids were extracted using methyl tert-butyl ether (MTBE) extraction and analysed with liquid chromatography-tandem mass spectrometry (LC-MS/MS) as reported earlier[62]. In brief, snap-frozen urinary bladder tissues were homogenised in ice-cold IPA:H2O (1:1 v/v; 150 µL), spiked with 10 µL of SPLASH lipid standard mix, and then subjected to liquid–liquid extraction using MeOH, MTBE and water solvent mixture with a final ratio of 1:3:1 (v/v/v). The upper organic phase was collected, dried out under a nitrogen gas stream, and reconstituted in 100 µL of ACN:IPA:H2O buffer (65:30:5 v/v/v). Resuspended lipid extracts (10 µL) were loaded on a reversed-phase ACQUITY UPLC HSS T3 (1.8 µm, 100 × 2.1 mm, Waters Corporation) column and separated using a Vanquish Duo UHPLC-system (Thermo Fisher Scientific, Waltham, MA, USA) with a flow rate of 250 µL/min. The mobile phases consisted of eluent A (H2O:ACN, 40:60 v/v) and eluent B (IPA:ACN, 90:10 v/v) both with 10 mM ammonium formate and 0.1% formic acid. All the datasets were acquired independently in negative- and positive-ion mode in a data-dependent manner using Orbitrap Fusion Lumos Tribrid Mass Spectrometer (Thermo Fisher Scientific, Waltham, MA, USA) equipped with a heated electrospray ionisation source. Lipids were annotated and validated with lipid class and molecular-species specific diagnostic fragment ions[63].

MSI data pre-processing

MALDI MSI data were pre-processed by custom-designed Python scripts which were automated in a Snakemake[64] pipeline. The pipeline takes raw timsTOF files as input, processes all files in parallel and outputs the processed data in imzML format along with various quality control visualisations for each step. All parameters used within this study are listed in Supplementary Table 1. In the first step, Savitzky–Golay smoothing[65] was used to reduce spectral random noise. Centroid spectra were extracted by using the find_peaks function from the SciPy signal processing library with default parameters. Peaks with a SNR of at least three were selected. The noise was calculated by the median absolute deviation. To eliminate mass drifts, the pipeline contains an alignment procedure using a kernel-based clustering approach adapted from pyBASIS[20]. Here peaks which were present in at least 3% of all pixels across all samples formed the common m/z vector to which the peaks were aligned by nearest-neighbour mapping. Then off-tissue/matrix pixels were identified and removed for each dataset (Supplementary Fig. 2). To identify matrix pixels, data of each sample were first reduced to two dimensions by using UMAP followed by HDBSCAN clustering. The cluster which was connected to most pixels of the border of the measured area was considered the matrix cluster. Clusters with a Spearman correlation above 0.7 to the matrix cluster were combined to an extended matrix cluster. In the third step, the binary image of the extended matrix cluster was post-processed by removing isolated objects of up to 5 pixels followed by a binary closing operation using a 5 × 5 pixel square structuring element for the dilation and a 2 × 2 pixel square structuring element for the erosion. For the post-processing, the scikit-image and SciPy libraries were used. After matrix removal, the spatial coherence[44] was calculated for each ion which measures its informativeness. Ions with low spatial coherence were removed. Then the data were normalised by median-fold-change normalisation to account for intra-sample and inter-sample variation. To indicate spectra variations among the samples, a UMAP-based outlier detection method was applied (Supplementary Fig. 3). Here the data of all samples were first reduced to two dimensions by using UMAP followed by HDBSCAN clustering. Then SSCs were identified. SSCs are clusters in which most pixels originate from one sample. Finally, samples in which most pixels were SSC pixels were considered sample outliers. De-isotoping was performed as the last step. In an iterative approach, isotopes were identified based on their theoretical m/z value within a predefined tolerance range and their theoretical intensity pattern.

Lipid annotation

In the first step, a list of tentative annotations of m/z values was made by using the bulk structure search provided on the LipidMaps website (www.lipidmaps.org). We searched for matches between all expected lipid classes (fatty acids/esters [FA], ceramides [Cer], sphingomyelins [SM], hexosyl ceramides [HexCer], triglycerols [TG], diglycerols [DG], glycerophosphocholines [PC], glycerophosphates [PA], glycerophosphoserines [PS], glycerophosphoethanolamines [PE], glycerophosphoglycerols [PG], sterols [ST]) and $[M+H]^+$, $[M-H_2O+H]^+$, $[M+Na]^+$ and $[M+K]^+$ precursor ions with a mass tolerance of ±0.01 m/z. This list was then manually curated considering the biological context and expanded with the annotations gained from MALDI-DDA-MSI as well as other on-tissue MALDI MS/MS experiments. In addition, some lipids were validated by LC-MS/MS, as described in LC-MS/MS validation strategy for lipid annotations[66]. If no further information from MS/MS was gained, the annotation is only based on the lipid species level, which means in the context of ether lipids, that 1-O-alkyl lipids with at least one double bond could also be interpreted as 1-O-alkenyl ethers, as described in lipid nomenclature guidelines[67].

Data analysis, statistics and visualisation

Tissue segmentation was performed on the full m/z spectrum of all control and infected urinary bladder sections by UMAP (n_neighbors = 10, min_dist = 0.0, dist_metric = 'cosine') followed by HDBSCAN clustering (min_samples = 50, min_cluster_size = 10,000). The clusters were manually merged into super-clusters which represent the main tissue context (urothelium, lamina propria and muscle). Then mean spectra of the pixel clusters were extracted and statistically analysed by two-sided statistical tests using the Python package SciPy. The standard t-test was used for normally distributed populations with equal variances and the Welch's t-test for normally distributed populations with unequal variances. Non-normally distributed populations were statistically analysed by the Wilcoxon rank-sum test. The Shapiro–Wilk test and Levene test were used to test for normal distribution and equal variances. The ratio between the means of two populations (here control and infected) was determined as $\log_2$(fold change). A full statistical summary (means, standard deviations, test statistics, degrees of freedom, p values and fold changes) is provided in the source data file.

For the lipidomic analysis of neutrophils, we reduced the m/z spectrum by filtering out m/z values for which no lipid in the database could be matched based on our m/z tolerance, as well as tissue-specific m/z values with a Pearson's correlation (SciPy Python package) above 0.5 to one of the tissue regions extracted from the previous tissue

segmentation (urothelium, lamina propria, muscle). To differentiate between $Ly6G^+$ and $Ly6G^-$ pixels, we segmented the registered Ly6G images using the Python package Scikit-image. First images were Gaussian smoothed with sigma = 1 followed by Otsu thresholding. Finally, small objects (below 10 pixels size) were removed. Unsupervised clustering of $Ly6G^+$ pixels was performed by UMAP (n_neighbors = 3, min_dist = 0.0, dist_metric = 'cosine') followed by HDBSCAN clustering (min_samples = 30, min_cluster_size = 500).

Important lipids for the ROIs (here tissue regions and neutrophils) were extracted by using Extreme Gradient Boosting (XGBoost)-based classification with class weights, Pearson's correlation and SHAP values. Therefore, the Python packages XGBoost, SciPy and SHAP were used. Class weights were calculated by #samples / (#classes * #occurences) using the Python package scikit-learn.

The Python packages Pandas, Seaborn, Matplotlib and matplotlib-venn were used to generate all visualisations (volcano plots, scatterplots, pie charts, Venn diagrams, segmented images and bar plots).

Microscopy data pre-processing

Images were pre-processed by rolling-ball background subtraction, Gaussian smoothing with sigma = 3 followed by percentile stretching using the Python library scikit-image.

Registration of microscopy and MALDI MSI

Initial transformation of the images was performed in Fiji for microscopy (rotation and cropping) to match the imaged tissue region and orientation of both modalities. For precise registration we used symmetric normalisation (SyNRA) as transformation, consisting of rigid, affine and deformable transformation, and Mattes mutual information as optimisation metric implemented in the ANTs library[30]. We down-sampled the microscopy image, using linear interpolation, to the same spatial dimensions as the MALDI image using the Python package ImgAug (https://github.com/aleju/imgaug). The down-sampled AF from microscopy was used as moving image and the UMAP image from MSI, which visualises the main tissue structure similar to the AF image, was used as fixed image. We validated the registration result by the Jaccard index for the overlap between registered urothelial mask from microscopy and MSI mask. We applied spatial k-means clustering on the MSI data and selected the cluster containing the urothelium. The urothelial mask from microscopy was manually created.

Statistics and reproducibility

All experiments were conducted once. No measures were taken to verify the reproducibility of the experimental findings, as the primary objective was to demonstrate the software's applicability by a proof-of-concept study.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All MALDI MSI and IFM data (raw and results) generated in this study have been deposited in Zenodo under accession code https://doi.org/10.5281/zenodo.14184882. The intermediate pre-processing results of the MALDI-2 MSI bladder data are available upon request from the corresponding author D.R.E. The data are not publicly available due to data size. Requests will be fulfilled within 2 weeks. The LC-MS/MS data generated in this study have been deposited in MassIVE under identifier MSV000096465. Source data are provided with this paper.

Code availability

The msiFlow source code is publicly available on GitHub (https://github.com/Immunodynamics-Engel-Lab/msiflow)[68].

References

1. Hornburg, D. et al. Dynamic lipidome alterations associated with human health, disease and ageing. *Nat. Metab.* **5**, 1578–1594 (2023).
2. Morgan, P. K. et al. A lipid atlas of human and mouse immune cells provides insights into ferroptosis susceptibility. *Nat. Cell Biol.* **26**, 645–659 (2024).
3. Serhan, C. N. Pro-resolving lipid mediators are leads for resolution physiology. *Nature* **510**, 92–101 (2014).
4. Sheppe, A. E. F. & Edelmann, M. J. Roles of eicosanoids in regulating inflammation and neutrophil migration as an innate host response to bacterial infections. *Infect. Immun.* **89**, e0009521 (2021).
5. Dennis, E. A. & Norris, P. C. Eicosanoid storm in infection and inflammation. *Nat. Rev. Immunol.* **15**, 511–523 (2015).
6. Foxman, B. Epidemiology of urinary tract infections: incidence, morbidity, and economic costs. *Dis. Mon.* **49**, 53–70 (2003).
7. Ulett, G. C. et al. Uropathogenic Escherichia coli virulence and innate immune responses during urinary tract infection. *Curr. Opin. Microbiol* **16**, 100–107 (2013).
8. Kuhn, H. W., Hreha, T. N. & Hunstad, D. A. Immune defenses in the urinary tract. *Trends Immunol.* **44**, 701–711 (2023).
9. Kolaczkowska, E. & Kubes, P. Neutrophil recruitment and function in health and inflammation. *Nat. Rev. Immunol.* **13**, 159–175 (2013).
10. Schiwon, M. et al. Crosstalk between sentinel and helper macrophages permits neutrophil migration into infected uroepithelium. *Cell* **156**, 456–468 (2014).
11. Shin, K. et al. Hedgehog/Wnt feedback supports regenerative proliferation of epithelial stem cells in bladder. *Nature* **472**, 110–114 (2011).
12. Kinkhabwala, A. et al. MACSima imaging cyclic staining (MICS) technology reveals combinatorial target pairs for CAR T cell treatment of solid tumors. *Sci. Rep.* **12**, 1911 (2022).
13. Soltwisch, J. et al. Mass spectrometry imaging with laser-induced postionization. *Science* **348**, 211–215 (2015).
14. Niehaus, M., Soltwisch, J., Belov, M. E. & Dreisewerd, K. Transmission-mode MALDI-2 mass spectrometry imaging of cells and tissues at subcellular resolution. *Nat. Methods* **16**, 925–931 (2019).
15. Dreisewerd, K., Bien, T. & Soltwisch, J. MALDI-2 and t-MALDI-2 Mass Spectrometry Imaging. *Methods Mol. Biol.* **2437**, 21–40 (2022).
16. Goossens, P. et al. Integrating multiplex immunofluorescent and mass spectrometry imaging to map myeloid heterogeneity in its metabolic and cellular context. *Cell Metab.* **34**, 1214–1225.e1216 (2022).
17. Scupakova, K., Dewez, F., Walch, A. K., Heeren, R. M. A. & Balluff, B. Morphometric cell classification for single-cell MALDI-mass spectrometry imaging. *Angew. Chem. Int. Ed.* **59**, 17447–17450 (2020).
18. Patterson, N. H., Tuck, M., Van de Plas, R. & Caprioli, R. M. Advanced registration and analysis of MALDI imaging mass spectrometry measurements through autofluorescence microscopy. *Anal. Chem.* **90**, 12395–12403 (2018).
19. Patterson, N. H., Yang, E., Kranjec, E. A. & Chaurand, P. Co-registration and analysis of multiple imaging mass spectrometry datasets targeting different analytes. *Bioinformatics* **35**, 1261–1262 (2019).
20. Veselkov, K. et al. BASIS: high-performance bioinformatics platform for processing of large-scale mass spectrometry imaging data in chemically augmented histology. *Sci. Rep.* **8**, 4053 (2018).
21. Bemis, K. D. et al. Cardinal: an R package for statistical analysis of mass spectrometry-based imaging experiments. *Bioinformatics* **31**, 2418–2420 (2015).
22. Bokhart, M. T., Nazari, M., Garrard, K. P. & Muddiman, D. C. MSiReader v1.0: evolving open-source mass spectrometry imaging software for targeted and untargeted analyses. *J. Am. Soc. Mass Spectrom.* **29**, 8–16 (2018).
23. Wullems, K. et al. Fast visual exploration of mass spectrometry images with interactive dynamic spectral similarity pseudocoloring. *Sci. Rep.* **11**, 4606 (2021).

24. Foll, M. C. et al. Accessible and reproducible mass spectrometry imaging data analysis in Galaxy. *Gigascience* **8**, https://doi.org/10.1093/gigascience/giz143 (2019).
25. Koster, J. & Rahmann, S. Snakemake-a scalable bioinformatics workflow engine. *Bioinformatics* **34**, 3600 (2018).
26. Merkel, D. Docker: lightweight Linux containers for consistent development and deployment. *Linux J.* **2014**, 2 (2014).
27. McInnes, L., Healy, J. & Melville, J. Umap: uniform manifold approximation and projection. *J. Open Source Softw.* 861 (2018).
28. McInnes, L., Healy, J. & Astels, S. hdbscan: hierarchical density based clustering. *J. Open Source Softw.* **2**, 205 (2017).
29. Verbeeck, N., Caprioli, R. M. & Van de Plas, R. Unsupervised machine learning for exploratory data analysis in imaging mass spectrometry. *Mass Spectrom. Rev.* **39**, 245–291 (2020).
30. Avants, B. B., Epstein, C. L., Grossman, M. & Gee, J. C. Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. *Med. Image Anal.* **12**, 26–41 (2008).
31. Laurens van der Maaten, G. H. Visualizing data using t-SNE. *J. Mach. Learn. Res.* **9**, 2579–2605 (2008).
32. Alexandrov, T. & Kobarg, J. H. Efficient spatial segmentation of large imaging mass spectrometry datasets with spatially aware clustering. *Bioinformatics* **27**, i230–i238 (2011).
33. Ke, G. et al. LightGBM: a highly efficient gradient boosting decision tree. *Adv. Neural Inf. Process. Syst.* **30**, 3149–3157 (2017).
34. Chen, T. & Guestrin, C. In *Proc. 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 785–794 (ACM, 2016).
35. Lundberg, S. M. & Lee, S.-I. A unified approach to interpreting model predictions. *Adv. Neural Inf. Process. Syst.* **30** (2017).
36. Tideman, L. E. et al. Automated biomarker candidate discovery in imaging mass spectrometry data through spatially localized Shapley additive explanations. *Anal. Chim. Acta* **1177**, 338522 (2021).
37. Esselman, A. B. et al. In situ molecular profiles of glomerular cells by integrated imaging mass spectrometry and multiplexed immunofluorescence microscopy. *Kidney Int.* **107**, 332–337 (2025).
38. Oursel, D. et al. Lipid composition of membranes of Escherichia coli by liquid chromatography/tandem mass spectrometry using negative electrospray ionization. *Rapid Commun. Mass Spectrom.* **21**, 1721–1728 (2007).
39. Kallback, P., Nilsson, A., Shariatgorji, M. & Andren, P. E. msIQuant-Quantitation Software for mass spectrometry imaging enabling fast access, visualization, and analysis of large data sets. *Anal. Chem.* **88**, 4346–4353 (2016).
40. Robichaud, G., Garrard, K. P., Barry, J. A. & Muddiman, D. C. MSiReader: an open-source interface to view and analyze high resolving power MS imaging files on Matlab platform. *J. Am. Soc. Mass Spectrom.* **24**, 718–721 (2013).
41. Guo, G. et al. Automated annotation and visualisation of high-resolution spatial proteomic mass spectrometry imaging data using HIT-MAP. *Nat. Commun.* **12**, 3241 (2021).
42. Bond, N. J., Koulman, A., Griffin, J. L. & Hall, Z. massPix: an R package for annotation and interpretation of mass spectrometry imaging data for lipidomics. *Metabolomics* **13**, 128 (2017).
43. Baquer, G. et al. What are we imaging? Software tools and experimental strategies for annotation and identification of small molecules in mass spectrometry imaging. *Mass Spectrom. Rev.* **42**, 1927–1964 (2023).
44. Grelard, F. et al. Esmraldi: efficient methods for the fusion of mass spectrometry and magnetic resonance images. *BMC Bioinform.* **22**, 56 (2021).
45. Strohalm, M., Kavan, D., Novak, P., Volny, M. & Havlicek, V. mMass 3: a cross-platform software environment for precise analysis of mass spectrometric data. *Anal. Chem.* **82**, 4648–4651 (2010).
46. Bien, T. et al. MALDI-2 mass spectrometry and immunohistochemistry imaging of Gb3Cer, Gb4Cer, and further glycosphingolipids in human colorectal cancer tissue. *Anal. Chem.* **92**, 7096–7105 (2020).
47. Chaurand, P. et al. Integrating histology and imaging mass spectrometry. *Anal. Chem.* **76**, 1145–1155 (2004).
48. Verbeeck, N. et al. Automated anatomical interpretation of ion distributions in tissue: linking imaging mass spectrometry to curated atlases. *Anal. Chem.* **86**, 8974–8982 (2014).
49. Molenaar, M. R. et al. LION/web: a web-based ontology enrichment tool for lipidomic data analysis. *Gigascience* **8**, giz061 (2019).
50. Mueller, H. W., O'Flaherty, J. T. & Wykle, R. L. Ether lipid content and fatty acid distribution in rabbit polymorphonuclear neutrophil phospholipids. *Lipids* **17**, 72–77 (1982).
51. Mueller, H. W., O'Flaherty, J. T., Greene, D. G., Samuel, M. P. & Wykle, R. L. 1-O-alkyl-linked glycerophospholipids of human neutrophils: distribution of arachidonate and other acyl residues in the ether-linked and diacyl species. *J. Lipid Res.* **25**, 383–388 (1984).
52. Lammermann, T. et al. Neutrophil swarms require LTB4 and integrins at sites of cell death in vivo. *Nature* **498**, 371–375 (2013).
53. Dean, J. M. & Lodhi, I. J. Structural and functional roles of ether lipids. *Protein Cell* **9**, 196–206 (2018).
54. Abraham, S. N. & Miao, Y. The nature of immune responses to urinary tract infections. *Nat. Rev. Immunol.* **15**, 655–663 (2015).
55. Scott, A. J., Ellis, S. R., Hofstaedter, C. E., Heeren, R. M. A. & Ernst, R. K. Spatial lipidomics reveals biased phospholipid remodeling in acute Pseudomonas lung infection. *iScience* **26**, 107700 (2023).
56. Roussel, C. et al. Short-term supplementation with omega-3 polyunsaturated fatty acids modulates primarily mucolytic species from the gut luminal mucin niche in a human fermentation system. *Gut Microbes* **14**, 2120344 (2022).
57. Welte, M. A. Expanding roles for lipid droplets. *Curr. Biol.* **25**, R470–R481 (2015).
58. Bosch, M., Sweet, M. J., Parton, R. G. & Pol, A. Lipid droplets and the host-pathogen dynamic: FATal attraction? *J. Cell Biol.* **220**, https://doi.org/10.1083/jcb.202104005 (2021).
59. Bien, T., Hambleton, E. A., Dreisewerd, K. & Soltwisch, J. Molecular insights into symbiosis-mapping sterols in a marine flatworm-algae-system using high spatial resolution MALDI-2-MS imaging with ion mobility separation. *Anal. Bioanal. Chem.* **413**, 2767–2777 (2021).
60. Bien, T., Bessler, S., Dreisewerd, K. & Soltwisch, J. Transmission-mode MALDI mass spectrometry imaging of single cells: optimizing sample preparation protocols. *Anal. Chem.* **93**, 4513–4520 (2021).
61. Ellis, S. R. et al. Automated, parallel mass spectrometry imaging and structural identification of lipids. *Nat. Methods* **15**, 515–518 (2018).
62. Gautam, J. et al. Characterization of lipid signatures in the plasma and insulin-sensitive tissues of the C57BL/6J mice fed on obesogenic diets. *Biochim. Biophys. Acta Mol. Cell Biol. Lipids* **1868**, 159348 (2023).
63. Garikapati, V., Colasante, C., Baumgart-Vogt, E. & Spengler, B. Sequential lipidomic, metabolomic, and proteomic analyses of serum, liver, and heart tissue specimens from peroxisomal biogenesis factor 11alpha knockout mice. *Anal. Bioanal. Chem.* **414**, 2235–2250 (2022).
64. Molder, F. et al. Sustainable data analysis with Snakemake. *F1000Research* **10**, 33 (2021).
65. Savitzky, A. & Golay, M. J. E. Smoothing and differentiation of data by simplified least squares procedures. *Anal. Chem.* **36**, 1627–1639 (1964).

66. Palmer, A. et al. FDR-controlled metabolite annotation for high-resolution imaging mass spectrometry. *Nat. Methods* **14**, 57–60 (2017).
67. Kopczynski, D. et al. Goslin 2.0 implements the recent lipid shorthand nomenclature for MS-derived lipid structures. *Anal. Chem.* **94**, 6097–6101 (2022).
68. Spangenberg, P. et al. msiFlow: automated workflows for reproducible and scalable multimodal mass spectrometry imaging and microscopy data analysis. msiFlow. https://doi.org/10.5281/zenodo.14056181 (2024).

Acknowledgements

We acknowledge support by the Open Access Publication Fund of the University of Duisburg-Essen, the Central Animal Facilities of the Medical Faculty Essen, the Imaging Center Essen (Alexandra Brenzel and Dr. Anthony Squire) and the Immunoproteomics group (Dr. Olga Shevchuk, Stephanie Tautges-Schaefer, Stephanie Thiebes and Jenny Dick), supported by INST 20876/486-1 to DRE. The following authors receive funding from the German Research Foundation: FOR5427 SP1 (O.Sh.); TRR296 P09 (D.R.E.); TRR332 A3 (D.R.E); FOR5427 SP4 (D.R.E.); EN984/15-1,16-1 and 18-1 (D.R.E.); INST 20876/486-1 (D.R.E); DR416/13-1 (K.D.); SO976/4-1 (J.S.); TRR332 Z1 (D.R.E., K.D., O.So.); SFB1009 A13 (O.So.); SFB1123 A6 (O.So.); TRR332 A2 (O.So.). O.So. receives funding from Novo Nordisk, the Leducq Foundation, the IMF and the IZKF of the Medical Faculty of the University of Münster.

Author contributions

P.S. designed the software. P.S. developed and implemented the algorithms and workflows. D.S. implemented the browser-based user interface. L.G.M. implemented the raw MSI data reader. S.B. acquired MSI data. L.W., J.B., M.R. and S.T. acquired IFM data. L.W., J.B. and S.T. performed animal experiments and sample preparation. M.R. performed isolation of neutrophils. S.S.K. performed LC-MS/MS measurements. P.S., S.B., L.W., J.S., O.Sh., K.D., P.P., S.S.K. and D.R.E. interpreted the data. S.D.K, J.K., H.H., D.F., R.V.P., L.C.M., M.G. and O.So. contributed to the project development. P.S. and D.R.E. wrote the manuscript. D.R.E. supervised and coordinated the study. All authors revised the manuscript and approved the submitted version.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-024-55306-7.

Correspondence and requests for materials should be addressed to Daniel R. Engel.

Peer review information *Nature Communications* thanks Janet Sherman, who co-reviewed with Behnaz Akbari, and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

© The Author(s) 2025