

Article

# Advanced Registration and Analysis of MALDI Imaging Mass Spectrometry Measurements through Autofluorescence Microscopy

Nathan Heath Patterson, Michael Tuck, Raf Van de Plas, and Richard M. Caprioli

*Anal. Chem.*, **Just Accepted Manuscript** • DOI: 10.1021/acs.analchem.8b02884 • Publication Date (Web): 01 Oct 2018

Downloaded from <http://pubs.acs.org> on October 3, 2018

## Just Accepted

"Just Accepted" manuscripts have been peer-reviewed and accepted for publication. They are posted online prior to technical editing, formatting for publication and author proofing. The American Chemical Society provides "Just Accepted" as a service to the research community to expedite the dissemination of scientific material as soon as possible after acceptance. "Just Accepted" manuscripts appear in full in PDF format accompanied by an HTML abstract. "Just Accepted" manuscripts have been fully peer reviewed, but should not be considered the official version of record. They are citable by the Digital Object Identifier (DOI®). "Just Accepted" is an optional service offered to authors. Therefore, the "Just Accepted" Web site may not include all articles that will be published in the journal. After a manuscript is technically edited and formatted, it will be removed from the "Just Accepted" Web site and published as an ASAP article. Note that technical editing may introduce minor changes to the manuscript text and/or graphics which could affect content, and all legal disclaimers and ethical guidelines that apply to the journal pertain. ACS cannot be held responsible for errors or consequences arising from the use of information contained in these "Just Accepted" manuscripts.



ACS Publications

is published by the American Chemical Society, 1155 Sixteenth Street N.W.,  
Washington, DC 20036

Published by American Chemical Society. Copyright © American Chemical Society.  
However, no copyright claim is made to original U.S. Government works, or works  
produced by employees of any Commonwealth realm Crown government in the course  
of their duties.

## Title

# Advanced Registration and Analysis of MALDI Imaging Mass Spectrometry Measurements through Autofluorescence Microscopy

## Authors:

Nathan Heath Patterson<sup>1,2</sup>, Michael Tuck<sup>1</sup>, Raf Van de Plas<sup>1,2,5</sup>, Richard M. Caprioli<sup>1,2,3,4</sup>

## Author Affiliations:

<sup>1</sup>Mass Spectrometry Research Center and Departments of <sup>2</sup>Biochemistry, <sup>3</sup>Chemistry,

<sup>4</sup>Pharmacology and Medicine, Vanderbilt University, Nashville, Tennessee, USA.

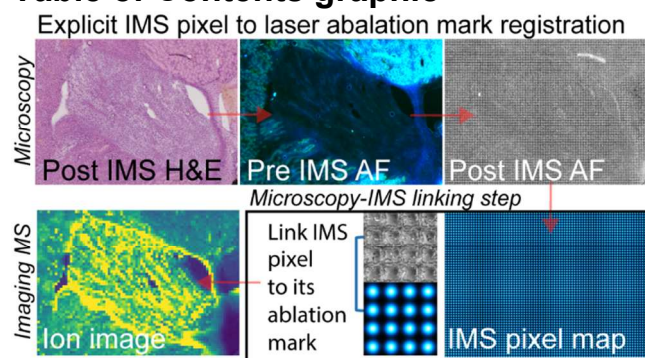
<sup>5</sup>Delft Center for Systems and Control (DCSC), Delft University of Technology, Delft, the Netherlands.

**Contact for corresponding author:** Richard M. Caprioli [richard.m.caprioli@vanderbilt.edu](mailto:richard.m.caprioli@vanderbilt.edu)

## Abstract:

The correlation of imaging mass spectrometry (IMS) with histopathology can help relate novel molecular findings obtained through IMS to the well-characterized and validated histopathology knowledge base. The quality of correlation between these two modalities is limited by the quality of the spatial mapping that is obtained by registration of the two image types. In this work, we develop novel workflows for MALDI IMS-to-microscopy data registration and analysis using non-destructive IMS-compatible wide field autofluorescence (AF) microscopy combined with computational image registration. First, a substantially automated procedure for high accuracy registration between IMS and microscopy data of the same section is described that explicitly links the MALDI laser ablation pattern imaged by microscopy to its corresponding IMS pixel. Subsequent examination of the registered data allows for high confidence co-localization of image features between the two modalities, down to single-cell scales within tissue. Building on this IMS-microscopy spatial mapping, we furthermore demonstrate the automated spatial correlation between IMS measurements from serial sections. This AF-registration-driven inter-section analysis, using a combination of non-linear AF-to-AF and IMS-to-AF image registrations, can be applied to tissue sections that are prepared and imaged with different sample preparations (e.g. lipids vs. proteins) and/or that are measured using different spatial resolutions. Importantly, all registrations, whether within a single section or across serial sections, are entirely independent of the IMS intensity signal content and thus unbiased by it.

## Table of Contents graphic



## Introduction

In histopathology, the examination of thin tissue sections by means of staining and microscopy has yielded a vast amount of information, actively driving prognostic and diagnostic applications in the clinic.<sup>1</sup> Imaging mass spectrometry (IMS) is currently growing in application as a powerful tool that can add a molecular dimension to the analysis of such thin tissue sections.<sup>2</sup> In recent years, there has been substantial progress towards integrating the medically well-characterized histological information, obtained through microscopy-measured stained tissue sections, with the molecular information provided by IMS.<sup>3,4</sup> The presence of both types of information from the same or serial tissue sections holds enormous potential for connecting the body of histopathological knowledge to the molecular building blocks and pathways that drive the underlying patho-mechanisms. However, the spatial linking of the two modalities has largely been achieved through manual or computational approaches that remain difficult to evaluate as no ground truth is available.

Currently, the registration between IMS and microscopy measurements from the same or an adjacent tissue section is often performed on the basis of image intensity patterns and/or fiducials manually provided by a user and almost always without explicit spatial reference points common to both data. While this can be useful, the approach tends to limit the achievable spatial accuracy of the registration to the spatial resolution of the lowest-resolution modality (i.e. the IMS data) or even less. This is often an order of magnitude or more coarser than the spatial resolution of the microscopy image. The manual approach also carries with it the risk of introducing human bias and drift into the registration, effects exacerbated by the often-limited number of fiducials that can be practically supplied by a human operator. These difficulties are further compounded by a lack of strong prior or ground truth knowledge to evaluate registration performance against. The limits on registration accuracy can become particularly influential when focusing on tissue features at microscopic scale (e.g. interpreting IMS data from single-cell or small areas). This means that large patches of tissue, defined by relatively broad annotations such as “normal” or “cancer” tissue, may be correctly aligned. However, using a narrower microscopy image annotation to confidently extract IMS data from relatively small sub-regions or specific groupings of cells tends to be less reliable.

While accounting for the above limitations, recently there have been several useful and important contributions in the area of IMS registration. Abdelmoula *et al.*<sup>5</sup> demonstrated automated non-linear registration between microscopy and IMS using a t-SNE dimensionality reduction of the IMS data as a representative image. Registration has also been implemented as part of automated IMS data analysis using anatomical atlases.<sup>6-9</sup> Verbeeck *et al.*<sup>9</sup> performed manual IMS registration to staining, then computational alignment of staining to a reference stain in the Allen Brain Atlas (ABA), while others demonstrated computational registration of IMS to staining as well as to the ABA. Recently, automated registration between IMS and an MRI atlas was demonstrated using a PCA decomposition IMS image that structurally resembles an image slice from a 3D MRI atlas<sup>8</sup>, further noting a challenge in registration between heterogeneous imaging data types where true spatial cross-references are difficult to obtain. Furthermore, data-driven image fusion, where one imaging modality is used to spatially sharpen or make predictions in a second modality, uses registration as an initial step and its models thus benefit from high spatial accuracy.<sup>10</sup> Many future developments of IMS will continue to employ microscopy images for comparative analysis and interpretation of analytical results, especially as ever finer spatial acquisition resolutions are pursued and achieved.

1  
2  
3 In this study, we employ wide field AF microscopy combined with automated computational  
4 image registration to address IMS-to-microscopy spatial registration difficulties and to develop  
5 novel data analysis schemes for the analysis of serial section datasets. In Workflow I, the  
6 registration between modalities is achieved through an AF microscopy intermediate in a way  
7 that is independent of the ion intensity content of the IMS data. This method *explicitly* links an  
8 IMS pixel to its corresponding ablation mark measured in a microscopy image. This measured  
9 ablation mark functions as a ground truth of the ion signal's spatial origin in microscopy  
10 coordinate space. Previous approaches performing direct image registration between modalities  
11 are thus *implicit* linkages of ablation marks to IMS data since the registered IMS image is only  
12 assumed to be registered with the laser ablation mark from which the data originates. After  
13 registration, data can be confidently extracted and analyzed by traditional data analysis  
14 pipelines. In Workflow II, we address the growing need for data analysis workflows that  
15 elegantly and automatically handle IMS datasets acquired from multiple serial sections. This is  
16 accomplished by means of a combination of IMS-AF registration (Workflow I) and AF-to-AF  
17 image registration.  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

## Materials & Methods

### Chemicals and reagents

1,5-diaminonaphthalene (DAN) and dihydroxyacetophenone (DHA) MALDI matrices were acquired from Sigma Aldrich (St. Louis, MO) and recrystallized in hot methanol prior to use.

### Animal & mammalian tissue samples

Rat brain, rat kidney, and mouse brain tissues were purchased from BioreclamationIVT (Hicksville, NY).

### Autofluorescence microscopy

All AF images were captured on a Nikon Eclipse 90i (Nikon Instruments Inc., Melville, NY) microscope equipped with a motorized stage. A 10x objective was used with resulting images with pixel resolutions of 0.92  $\mu\text{m}$  / pixel. DAPI (ex. = 340-380, em. = 435-485), FITC (ex. = 465-495, em. = 515-555), and TRITC (ex. = 528-553, em. = 590-650) (all wavelengths in nm) specific epifluorescence filters were used to cover a range of emission and excitation bands for AF. AF exposure times varied based on tissue type and the objective employed but generally were set around between 100-150 milliseconds with 8x analog gain on the camera.

Brightfield RGB images of stained tissues were acquired using a Leica SCN-400 (Leica Biosystems, Buffalo Grove, IL) slide scanner at 10x (1  $\mu\text{m}$ /pixel).

### Microscopy image preprocessing, registration, and analysis

FIJI ImageJ<sup>11,12</sup> was used for any additional image preprocessing, such as cropping or other relevant operations, and for visualization of microscopy images and data overlays.

Image registration was performed with the *Insight Segmentation and Registration Toolkit (ITK)* and *elastix*<sup>13</sup> libraries wrapped to Python3 using a suite of in-house scripts for SimpleITK<sup>14</sup> and simpleElastix<sup>15</sup> packages (see below for access to the source code). Image transformations were developed using rigid, affine, and B-spline transformation models with various optimized parameters. See the Supporting Information - Supplemental Methods for an in-depth explanation of registration parameters.

Scripts, software, and detailed instructions (for use of the two workflows described in this work) are available at <http://github.com/NHPatterson/regtoolboxmsrc>. The source code of latest version of the software is also included available as supplemental material. The software includes a GUI and command line tools to perform all described registration workflows. The package can also serve as a library to develop custom registration workflows.

### MALDI imaging mass spectrometry

#### Mass spectrometry

MALDI Time-of-Flight (TOF) IMS was performed using a Bruker rapifleX TissueTyper TOF mass spectrometer (Bruker Daltonics, Billerica, MA), equipped with a Gaussian 355 nm Nd:YAG laser. MALDI TOF delayed extraction time, extraction voltages, and laser power settings were optimized for sensitivity in the lipid mass range in both positive and negative ionization mode. MALDI Fourier Transform Ion Cyclotron Resonance (FT-ICR) IMS was performed using a Bruker solariX 9.4T FT-ICR mass spectrometer, equipped with a Gaussian 355 nm Nd:YAG

laser, and using a Bruker solariX 15T FT-ICR mass spectrometer, equipped with a Bruker SmartBeam II<sup>TM</sup> 355 nm Nd:YAG laser.

### *Sample preparation and experiments*

All sections from all organs were cut with a thickness between 6-12  $\mu\text{m}$ . Table 1 lists the tissue types and IMS experiments supporting this work. Matrix sublimation and recrystallization were done following previously published protocols.<sup>16,17</sup>

Sample	Workflow	Section thickness ( $\mu\text{m}$ )	MS instrument	Spatial res. IMS ( $\mu\text{m}$ )	MALDI matrix	No. of pixels	No. of shots per pixel
Rat brain	1	12	RF	10	DAN	73404	300
Rat kidney	1	10	RF	30	DAN	72768	500
Rat kidney	2	6,10,12	RF	15	DAN	varies	300
Mouse brain	2	8	9.4FT-ICR	40	DAN, Au	62833, 62115	750, 750
Rat brain	2	8	9.4 & 15F-TICR	35 (lip.) & 70 (pro.)	DHA, DHA	112490, 26197	500, 1000

**Table 1. Overview of AF and IMS experiments performed.** RF= Bruker rapifleX TissueTyper, 9.4FT-ICR = Bruker 9.4T solariX, 15FT-ICR = Bruker 15T solariX.

Protein imaging using the 15T FTICR instrument was performed based on previously published protocols.<sup>18</sup>

### *MALDI imaging mass spectrometry preprocessing and analysis*

IMS data was exported to the imzML format<sup>19</sup> and loaded into R using the *Cardinal* package.<sup>20</sup>

Peaks with a signal-to-noise ratio greater than 3 were selected and only those peaks which appeared in more than 3% of spectra were retained.

## Results and discussion

The results and discussion section will focus on the data analysis process. To help with understanding the experimental steps involved and their order, step-by-step flowcharts for the workflows are provided in the Supporting Information – Supplemental Methods, along with a guide to important experimental and computational parameters in both workflows.

### Workflow I. High accuracy IMS-to-microscopy registration

A highly automated method for registration of IMS and microscopy data from the same tissue section is demonstrated. This method explicitly links IMS pixels to their corresponding laser ablation marks in microscopy. The use of physical (IMS laser-introduced) ablation marks as a form of ground truth for the spatial mapping allows this method to achieve high spatial accuracy.

#### *Linking microscopy to IMS on the same tissue section through an AF intermediate and explicit pixel-to-laser-ablation-mark registration*

AF microscopy is a non-destructive, IMS compatible microscopy modality that provides high contrast, morphology rich information for computational image registration. Figure 1 outlines the IMS-microscopy registration workflow in case study images from a rat kidney both at low (A) and at high magnification (B). First, an AF image is acquired from the tissue prior to matrix deposition and prior to IMS acquisition (the pre-acquisition AF image). Subsequently, normal MALDI IMS sample preparation and data acquisition is performed. After IMS acquisition, a second AF image is taken (post-acquisition AF image) with the matrix layer still on the tissue. As is apparent from Figure 1 A&B – Post-acquisition AF, residual tissue AF is readily detectable through the matrix layer after acquisition as well as clearly visible MALDI laser ablation marks. Thus in one high-resolution modality, the post-acquisition AF microscopy image empirically measures morphological tissue patterns that tie to other microscopy images as well as the MALDI acquisition trace that ties to the theoretical grid of IMS pixels. This makes it a valuable intermediate modality to spatially map the two sides to each other. Using fluorescence excitation wavelengths past the UV (>350 nm) makes the post-acquisition approach amenable to most matrix preparations and avoids strong fluorescence of the MALDI matrix. Figure S1 shows post-acquisition AF images from DAN, DHA, 9AA, sputtered silver and gold. The post-acquisition AF image is acquired with both the microscope's epifluorescence light source and transmitted light source turned on (the latter set at low voltage; Figure S2-A). The epifluorescence light induces tissue AF, while the transmitted light travels through the partially or fully matrix-depleted laser ablation craters to be detected by the microscope's camera (Figure S2-B). If the overall goal is to spatially map IMS data to classical stained microscopy, such as an H&E, the tissue section can be stained once the post-acquisition AF image is acquired.

Once all microscopy images are acquired (pre-acquisition AF, post-acquisition AF, and optionally, stained brightfield), the computational registration procedure can bring the different image acquisition grids into a single coordinate system for inter-modality analysis. The IMS-to-stained-microscopy registration is obtained by means of several sub-registrations. Stained or other relevant microscopy is registered to the pre-acquisition AF image. Both modalities typically provide rich tissue morphology at high spatial resolution with relatively little noise. The pre-acquisition AF microscopy is then registered to the post-acquisition AF image. Since this registration between the AF pre-acquisition and residual AF post-acquisition is mono-modal, it can typically be accomplished with relatively straightforward registration procedures. Computationally, since the noise levels are relatively low and the content tends to be very



similar, all microscopy images are registered using a linear transformation model from the *elastix* library with a stochastic gradient descent optimizer (see Supplemental Methods for more detail). After registration between all microscopy modalities, a mask of the theoretical IMS pixel locations is generated at or near the resolution of the corresponding microscopy data, and a 2-D Gaussian intensity pattern is localized to each pixel's acquisition area (Figure 1 A&B - IMS pixel map) to visually identify each pixel's center-point in microscopy space. To achieve pixel-to-ablation-mark registration, corresponding fiducials are manually placed on the IMS pixel map image (at the center of the upsampled IMS pixel) as well as on the same pixels laser ablation crater's center in the post-acquisition AF microscopy image (Figure S3). This selection procedure typically requires 1-5 minutes as only 4-5 corresponding points gave already sufficient registration accuracy in the experiments described here. To further facilitate easy selection of laser ablation landmarks in the sometimes noisy post-acquisition AF image, one can even use the MALDI laser to ablate fiducials off-tissue directly in the MALDI matrix layer. This is an additional step available in those cases where ablation marks are more readily discernable in the matrix layer off-tissue compared to on-tissue (Figure S3 shows the principle and Figure S4 gives a demonstration). In IMS regimes where laser oversampling may occur and IMS pixels' ablation marks on tissue become difficult to individually distinguish under microscopy, it is recommended to e.g. double or triple the step size for the ablation fiducials in the matrix layer allowing very clear delineation of the center point of each fiducial pixel's laser ablation mark (Figure S3).

After landmark selection, all previously aligned microscopy images are registered to the upsampled IMS pixel map using a single linear transformation. In the case studies presented here, we have performed this step in FIJI ImageJ using the "Landmark Correspondences" plugin. However, any software capable of control point registration would suffice. The microscopy-to-IMS registration is performed in this direction to avoid interpolation artifacts of the IMS data during later visualization (e.g., each upscaled IMS pixel is a square without rotation, resizing, or skewing). Importantly, microscopy image resolution is minimally sacrificed during this workflow. For display and overlay of IMS and microscopy data after registration, IMS data simply needs to be scaled up by a factor of IMS resolution / registered microscopy resolution. The display of a 30  $\mu\text{m}$  IMS image onto a 1  $\mu\text{m}$  microscopy image is thus simply a resize/rescale by a factor of 30. Although the microscopy images are interpolated during the registrations, we have remarked minimal qualitative differences. This is to be expected since interpolation-introduced discrepancies will typically only leave a trace at the order of the finest available spatial resolution, which for these microscopy modalities lies more than an order of magnitude away from the IMS resolutions that are relevant here. Videos S1 and S2 show the registration for the rat kidney shown in Figure 1 as a sliding opacity video where each microscopy layer is shown in sequence towards the IMS layers.

### *Evaluation of microscopy-IMS registration*

Where direct IMS-to-microscopy registration is typically calculated on the basis of a limited number of fiducials visible in both modalities intensity content, often difficult to discern or accurately localize, the sheer number and regular-grid nature of the ablation marks in an IMS-to-post-acq.-AF setting provide an abundance of high-quality, easily localizable, and ground truth reference fiducials. This makes it typically possible to achieve a higher spatial precision mapping in IMS-to-post-acq.-AF registration than is commonly possible in generic IMS-to-microscopy registration where registration of the IMS pixel to its origin in microscopy coordinates is *implicit* rather than *explicit*. After registration, manual evaluation of registration

accuracy is possible using overlays of the registered IMS pixel map image and the ablation pattern image (Figure S5, examples of IMS-ablation-pattern overlays). Besides manual qualitative evaluation, quantitative means of accessing global registration accuracy were developed by extracting laser ablation image data from pixels that form the contour of the IMS ROI (visualized in Figure S6). The contour pixels are selected because in many cases they lie matrix layer outside of the tissue section rather than on the tissue section itself. As previously mentioned, clearer delineation of the ablation crater and its center is possible when firing into the more homogenous matrix layer. Of all available IMS contour pixels, four hundred were randomly selected (see Figure 2A). For comparison, a group of contour pixels within the tissue area are highlighted with a blue box in Fig. 2A with a high magnification view shown in Fig. 2 B alongside pixels that are off tissue. For quantitative evaluation, the ablation mark images (from the post-acquisition AF image) of every IMS contour pixel (not only the 400 displayed in Fig. 2A) were summed to obtain a single consensus image of the centering of the ablation over the IMS pixel (Fig. 2C). A selection of the center can then be made on the sum image. Knowing the centroid of the sum ROI and the expected center (1/2 IMS spatial resolution in x and y), the deviation between the empirically determined center and the theoretical center can be quantified. This process was performed for 6 different datasets (Figure S7) and it consistently yielded deviations less than 2.2  $\mu\text{m}$  along each axis. In all case studies, these deviations lie well below the IMS spatial resolution, making this AF-driven registration procedure, to our knowledge, state-of-the-art in terms evaluable accuracy. Together with results for other case studies, an expanded discussion of this evaluation and how it can be used for characterization of stage artifacts and other issues is available in the supporting information (Figure S7 & Figure S8).

### *Driving IMS analysis with high accuracy IMS-to-microscopy registration*

Workflow I forms the basis for subsequent analysis by providing a means to achieve spatially accurate IMS data extraction and high confidence interpretation of combined microscopy and IMS results. Figure 3 shows an example of Workflow I applied to a 10  $\mu\text{m}$  step-size rat brain cerebellum dataset acquired on the rapiflex. The on-tissue laser spot size measured before acquisition was  $\sim 10 \times 10 \mu\text{m}$  using the instrument's "single" laser setting (no beam scanning). Figure 3 A shows the full IMS ROI and ion image  $m/z$  818.5, tentatively identified as PS(38:0) using the LIPIDMAPS database based on mass search, and the corresponding H&E stain at low magnification. Figure 3 B is a higher magnification view and demonstrates drawing ROIs on the microscopy image of 3 individual Purkinje cells at the border of the molecular layer and the grey matter regions of the rat cerebellum. To fully utilize the AF registration in data analysis, a second IMS pixel map was generated where each upsampled IMS pixel's area is assigned a unique integer index that cross references with the corresponding spectrum (Figure S9 shows this for the rat brain cerebellum dataset). ROIs annotated on microscopy are projected onto the indexed pixel map and IMS coordinates under the ROIs are extracted with the degree of ROI-IMS pixel overlap (between 0 and 100% overlap). In the rat brain cerebellum dataset, the spectral data for the 3 Purkinje cells was extracted with the percentage of each pixel's ROI overlap showing a strong spatial correlation of the  $m/z$  818.5's signal compared to the total spectral mean (Figure 3 C). The spatial co-localization of the signal with the structure was probed by Pearson correlation of intensity and IMS pixel-ROI overlap for 40 annotated Purkinje cells as shown in Figure 3 D's scatterplot. This analysis found a statistically significant linear relationship ( $r = 0.763$ ) with a  $p$ -value  $< 0.0001$ . Figure 3D also shows that pixels with less overlap appear to exhibit more heterogeneous signal likely due to the surrounding histology

being within the laser ablation alongside a small part of the Purkinje cell. The variance of this signal is further increased by the variety of cell types in surrounding histologies which can be presumed to have abundance of the displayed signal to varying degrees. Further, this model also assumes that every Purkinje cell body will have the same signal intensity (*i.e.*, every 100% overlap pixel will report the same intensity for the selected signal). This is improbable as the displayed signal may be associated with subcellular regions that are more or less present in different 3D planes of the cell body and one also expects there to be biological heterogeneity amongst the Purkinje cells.

Beyond targeted extraction based on microscopy ROIs, this method allows confident co-localization between microscopy and spatial features on the order of single IMS pixels. Figure S10 shows an instance in the rat brain where hotspots in the ion image of  $m/z$  687.54 within the tissue could be mistaken for noise but upon careful examination of the accurately registered H&E are shown to be neutrophils. In Figure S11, a small cryo-sectioning artifact overlapping primarily a single IMS pixel is shown from a rat kidney dataset acquired at 30  $\mu\text{m}$  IMS step-size. Where the artifact is present, a corresponding depression in signal congruent with the amount of the artifact underneath the pixel is observed ( $r = 0.78$ ,  $p\text{-value} = 0.038$ ).

This workflow greatly diminishes uncertainty in the co-localization of IMS molecular data to microscopy with underlying assumptions that become minor limitations. The method assumes that the IMS ablation grid will be regular along the x and y axes, *e.g.* that the target stage will not encode position imprecisions. While stage imprecisions occur in the grid of IMS pixels as previously discussed, they are usually a single stage step or 2.5  $\mu\text{m}$  on the instruments used in this work. While this is minor at the resolutions explored herein, at higher resolutions such imprecisions could lead to systematic oversampling and should be addressed instrumentally. Second, the individual ablation pattern of one IMS pixel is assumed to be a square while it can vary in shape depending on the laser type, optics, and fluence. The Bruker rapiflex provides a “beam scan” feature that generates roughly square ablation areas using special laser optics, shown several times in Figures S5 and 7 but other instruments may provide ablation shapes that are irregular (like the FTICR laser ablations show in Figure S8). Compounding this is IMS target planarity which can vary the angle at which the laser hits the sample and introduce artifacts into the laser ablation mark. And lastly, the base assumption that drives the data visualization and extraction is that the ions are generated primarily in what is perceived to be the visual center of the ablation crater on microscopy, however, while results presented here point to this being the case, higher accuracy will require a complete characterization of the beam as measured by AF microscopy and subsequent ion profile on the target surface.

## Workflow II. Automated spatial co-localization of ions from serial sections

The accurate linkage of IMS pixels to microscopy data allows new data analysis workflows to be developed using AF-to-AF registration of IMS-linked AF images (Workflow I). Workflow II, outlined in Figure 4, is an AF registration workflow that automates determination of spatially correlated ions from separate IMS datasets acquired from serial sections. This AF-based procedure avoids an actual registration of the noisier, relatively lower spatial resolution IMS data. The overall goal of this computational approach is to improve analysis of separate data sets from serial tissue sections without additional sample preparation (*i.e.* removing matrix and rescanning the same section by IMS) or unconventional data acquisition strategies.

In detail, following accurate linkage of microscopy to IMS (Workflow I), AF images from serial sections are registered using first a linear (affine) then a non-linear (b-spline) transformation

model (Figure 4, step 1). The linear transformation accounts for the coarse translations, rotations, and scaling between serial sections, while the non-linear transformation, initialized by the linear transformation, warps the tissue image locally to account for non-linear, localized discrepancies. Although only briefly mentioned here, details on the mechanics and parameters used for the registrations are discussed in more detail in the Supporting Information. After registration of AF images from serial sections, the linked, indexed IMS pixel maps from each dataset generating during workflow I are overlaid and overlapping indices and degree of overlap are returned (Figure 4, step 2). IMS intensity data matrices are then extracted by their index for cross-analysis. Weighted correlations allow integration of the overlap percentage of the IMS pixels and provides a means of determining the most spatially correlated signals. Altogether an automated analysis of signal correlation between separate datasets is obtained and signals can be queried for cross dataset correlation (Figure 4, step 3). A discussion and demonstration of the fundamentals of this workflow is available in the Supporting Information in Figures S12, S13, and Table S1.

### *Automatic spatial co-localization of IMS datasets at different spatial resolutions and different preparations*

Application of inter-section IMS dataset correlation is shown briefly in the workflow's outline in Figure 4. In that example, sputtered gold and sublimated DAN matrix were used on mouse brain tissue. These particular preparations allow ionization of phospholipids and other metabolites. Figure S14 shows a map the maximum overlap of individual pixels after registration for the correlated datasets. As some IMS pixels in one dataset overlap multiple pixels in the second dataset, only each pixel's maximum overlap per dataset is plotted. The noisy spatial profile in Figure S14 indicates that grid overlap is not spatially biased towards one area of the tissue and enough local and global structure is captured at high overlap to build a representative model. This local and global accuracy is exemplified in Figure 5A and B where correlated ion querying is shown to work for even the single cell layer of the rat brain ventricle. Figure 5C demonstrates the IMS-IMS correlation utility is independent of IMS image resolution using a multi spatial resolution dataset (lipid analysis at 35  $\mu\text{m}$ , protein analysis at 70  $\mu\text{m}$ ). IMS pixel maps are upsampled to near the microscopy resolution (1  $\mu\text{m}$  / pixel) in both datasets, however, each 70  $\mu\text{m}$  IMS pixel is represented by 4900 1  $\mu\text{m}$  pixels, whereas the 35  $\mu\text{m}$  upsampled IMS pixel's contain 1225 1  $\mu\text{m}$  pixels. In consideration of this parameter when calculating overlap threshold, only 35  $\mu\text{m}$  pixels fully contained within a 70  $\mu\text{m}$  pixel were kept for correlative analysis, with the possibility of each 70  $\mu\text{m}$  pixel containing 4 whole 35  $\mu\text{m}$  pixels. The subsequent pixel overlap map for the two datasets is shown in in Figure S15. The ability to handle multi scale IMS datasets for correlation querying addresses comparison of across preparations or instrumentation which do not allow high spatial resolution image acquisition.

In the case studies, the AF image is used to drive registration, but a stained image acquired post-acquisition could also be used to find the registration between serial sections as well. Regardless of the high-resolution image used to drive registration, decoupling mass spectral signal intensity from the registration process provides a robust and IMS data-independent means of determining spatial co-localization between two datasets.

## Conclusion

In this work, a new autofluorescence microscopy-based platform is demonstrated that integrates established techniques from image registration to achieve unprecedented IMS-to-microscopy image registration accuracy and to automate cross-analysis of serial section IMS datasets.

The developed IMS-microscopy registration approach provides a robust and mostly automated means of linking IMS data to microscopy through AF microscopy before and after IMS. Accuracy at this level allowed discovery of molecules expressed in single cells sparsely distributed within the tissue section as well as reduction of uncertainty in microscopy-IMS correlation and data extraction. Both qualitative and quantitative mechanisms for assessment of registration accuracy were developed using the overlay of the registered ablation pattern (ground truth) with theoretical IMS pixel map and examination of the centering of ablation marks within contour pixels of the IMS ROI. With the achieved accuracy, we establish, to our knowledge, the current state-of-the-art in registration accuracy between IMS and microscopy, opening to the door for such datasets being used as benchmarks for direct IMS-microscopy registration.

Linking IMS to microscopy with high accuracy also allowed us to take advantage of the acquired microscopy images to detect correlations between different IMS datasets. Here, non-linear registration between the high-resolution, morphology rich microscopy data was used to achieve registration between serial sections, vastly improving the spatial mapping accuracy that can be reached with IMS alone. This method addresses the growing need for computational approaches to handle serially acquired IMS datasets from one tissue. As MALDI IMS sample preparations are not one size fits all, the expectation going forward is sample preparations of ever finer molecular specificity which in turn requires data analysis methods for multiple sample preparations. Our platform does so without additional or unusual IMS sample preparation or acquisition strategies. We demonstrate this potential using correlated protein and lipids with multi-scale IMS datasets, showing correlations at the single cell layer. However, the list of possible experiments to which the methodology could be relevant expands to the number of different IMS sample preparations, conditions, and instrumentations that are developed.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

**Supporting Information**

Supplemental methods, experimental walkthrough, accuracy evaluation scheme for Workflow I and evaluation of fundamentals for Workflow II (PDF).

Video S1 & S2: Low (S1) and high (S2) magnification videos of the sequence of microscopy layers explicitly linking IMS pixels to their microscopy location in 30 µm rat kidney MALDI IMS dataset (zip).

Source code for the MSRC Registration Toolbox to perform experiments (zip).

## Figure Captions

**Figure 1. Microscopy to IMS single tissue section registration workflow shown using a rat brain cerebellum example ROI.** (A) Low magnification view of the tissue images. (B) High magnification ROI of the images for registration. Registration proceeds from the bottom towards the top. Registration I: Registration between secondary microscopy (H&E staining in example) and pre-acquisition AF image. Registration II: registration between pre-acquisition AF and post-acquisition AF. Registration III: Manual registration between post-acquisition (laser ablation marks) and corresponding IMS pixels. Scale bar 150  $\mu\text{m}$ . Pre-acq: pre-acquisition. Post-acq: post-acquisition.

**Figure 2. Quantitative and qualitative analysis of the IMS ROI's contour pixels to determine IMS to microscopy registration accuracy applied to a 30  $\mu\text{m}$  spatial resolution rat kidney dataset acquired on the Bruker rapifleX.** (A) Images of 400 randomly selected individual IMS contour pixels out of 946 of the underlying post-acquisition AF with laser ablation craters visible. The blue rectangle indicates 3 rows where the contour is over the tissue and not in the matrix surrounding the tissue. The image's contrast was increased for display purposes. (B) Single pixel examples where top row is off tissue and bottom row is on tissue, blue here indicates a pixel on the tissue surface versus in surrounding matrix. (C) Sum image of all contour pixels from the dataset with a red outline indicating the center of the ablation where error is calculated. In A-C, each tile is 30  $\mu\text{m}$ .

**Figure 3. Analysis of signal from Purkinje cells following AF based IMS registration to H&E of the same section.** (A) Low magnification of ion image, H&E and their overlay. (B) High magnification images 3 annotated Purkinje cells in the rat brain section. (C) Scatter plot of signal intensity and IMS pixel – ROI overlap from 40 annotated Purkinje cells with a linear regression line.  $r$  = Pearson correlation coefficient. P-value rejects null hypothesis that  $r = 0$ , no correlation between pixel overlap and signal of  $m/z$  818.5 exists.

**Figure 4. Workflow for automated discovery of correlated ions from serial sections using AF microscopy registration.** *Step 1:* First, microscopy images are registered to IMS following the method detailed in Fig. 1. Then microscopy layer images from the serial sections are registered together using non-linear models with the resulting transformation applied to the upscaled IMS pixel mask. *Step 2:* This step finds the IMS pixels that overlap as well as quantifies any overlap after registration. The overlapping IMS pixel indices are collated with the percentage of overlap as a table. *Step 3:* Data analysis is carried out by first extracting data matrices from the two IMS datasets where observations are overlapping pixels. A second vector contains the degree of overlap for each observation that is used to weigh the observations for correlation (left). The correlation matrix is then queried with an ion of interest from one dataset and the highest correlated ions from the second dataset are returned (right).

**Figure 5. Automated spatial co-localization of ions from different datasets by high-resolution microscopy image registration.** All ion images are scaled from 0 to their max

intensity individually. (A) Full size images from two separate datasets, showing a queried ion from the sputtered gold sample acquisition and the species with the highest spatial co-localization of it from the DAN acquisition. (B) Zoom in areas from the highlighted red boxes showing the fine structure at the brain ventricles. (C) A second dataset for co-localization of lipids and proteins from serial sections using different IMS spatial resolutions (DHA - Protein = 75µm, DHA - Lipid = 35µm).



## References

- (1) Digital pathology in clinical use: where are we now and what is holding us back?, Griffin, J.; Treanor, D. *Histopathology* **2017**, *70*, 134-145.
- (2) Molecular imaging of biological samples: localization of peptides and proteins using MALDI-TOF MS, Caprioli, R. M.; Farmer, T. B.; Gile, J. *Anal. Chem.* **1997**, *69*, 4751-4760.
- (3) Integrating histology and imaging mass spectrometry, Chaurand, P.; Schwartz, S. A.; Billheimer, D.; Xu, B. J.; Crecelius, A.; Caprioli, R. M. *Anal. Chem.* **2004**, *76*, 1145-1155.
- (4) MALDI-imaging segmentation is a powerful tool for spatial functional proteomic analysis of human larynx carcinoma, Alexandrov, T.; Becker, M.; Guntinas-Lichius, O.; Ernst, G.; von Eggeling, F. *J. Cancer Res. Clin. Oncol.* **2013**, *139*, 85-95.
- (5) Automatic generic registration of mass spectrometry imaging data to histology using nonlinear stochastic embedding, Abdelmoula, W. M.; Skraskova, K.; Balluff, B.; Carreira, R. J.; Tolner, E. A.; Lelieveldt, B. P.; van der Maaten, L.; Morreau, H.; van den Maagdenberg, A. M.; Heeren, R. M.; McDonnell, L. A.; Dijkstra, J. *Anal. Chem.* **2014**, *86*, 9204-9211.
- (6) Automatic registration of mass spectrometry imaging data sets to the Allen brain atlas, Abdelmoula, W. M.; Carreira, R. J.; Shyti, R.; Balluff, B.; van Zeijl, R. J.; Tolner, E. A.; Lelieveldt, B. F.; van den Maagdenberg, A. M.; McDonnell, L. A.; Dijkstra, J. *Anal. Chem.* **2014**, *86*, 3947-3954.
- (7) Precise Anatomic Localization of Accumulated Lipids in Mfp2 Deficient Murine Brains Through Automated Registration of SIMS Images to the Allen Brain Atlas, Skraskova, K.; Khmelinskii, A.; Abdelmoula, W. M.; De Munter, S.; Baes, M.; McDonnell, L.; Dijkstra, J.; Heeren, R. M. *J. Am. Soc. Mass Spectrom.* **2015**, *26*, 948-957.
- (8) Connecting imaging mass spectrometry and magnetic resonance imaging-based anatomical atlases for automated anatomical interpretation and differential analysis, Verbeeck, N.; Spraggins, J. M.; Murphy, M. J. M.; Wang, H. D.; Deutch, A. Y.; Caprioli, R. M.; Van de Plas, R. *Biochim. Biophys. Acta* **2017**, *1865*, 967-977.
- (9) Automated anatomical interpretation of ion distributions in tissue: linking imaging mass spectrometry to curated atlases, Verbeeck, N.; Yang, J.; De Moor, B.; Caprioli, R. M.; Waelkens, E.; Van de Plas, R. *Anal. Chem.* **2014**, *86*, 8974-8982.
- (10) Image fusion of mass spectrometry and microscopy: a multimodality paradigm for molecular tissue mapping, Van de Plas, R.; Yang, J.; Spraggins, J.; Caprioli, R. M. *Nat. Methods* **2015**, *12*, 366-372.
- (11) Fiji: an open-source platform for biological-image analysis, Schindelin, J.; Arganda-Carreras, I.; Frise, E.; Kaynig, V.; Longair, M.; Pietzsch, T.; Preibisch, S.; Rueden, C.; Saalfeld, S.; Schmid, B.; Tinevez, J. Y.; White, D. J.; Hartenstein, V.; Eliceiri, K.; Tomancak, P.; Cardona, A. *Nat. Methods* **2012**, *9*, 676-682.
- (12) The ImageJ ecosystem: An open platform for biomedical image analysis, Schindelin, J.; Rueden, C. T.; Hiner, M. C.; Eliceiri, K. W. *Mol. Reprod. Dev.* **2015**, *82*, 518-529.
- (13) elastix: a toolbox for intensity-based medical image registration, Klein, S.; Staring, M.; Murphy, K.; Viergever, M. A.; Pluim, J. P. *IEEE Trans. Med. Imaging* **2010**, *29*, 196-205.
- (14) The Design of SimpleITK, Lowekamp, B. C.; Chen, D. T.; Ibanez, L.; Blezek, D. *Front. Neuroinform.* **2013**, *7*, 45.
- (15) SimpleElastix: A user-friendly, multi-lingual library for medical image registration, Marstal, K.; Berendsen, F.; Staring, M.; Klein, S. In *WBIR Conference*: Las Vegas, Nevada, USA, 2016.
- (16) Sublimation of new matrix candidates for high spatial resolution imaging mass spectrometry of lipids: enhanced information in both positive and negative polarities after 1,5-diaminonaphthalene deposition, Thomas, A.; Charbonneau, J. L.; Fournaise, E.; Chaurand, P. *Anal. Chem.* **2012**, *84*, 2048-2054.

- (17) Matrix sublimation/recrystallization for imaging proteins by mass spectrometry at high spatial resolution, Yang, J.; Caprioli, R. M. *Anal. Chem.* **2011**, *83*, 5728-5734.
- (18) Enhanced Ion Transmission Efficiency up to  $m/z$  24000 for MALDI Protein Imaging Mass Spectrometry, Prentice, B. M.; Ryan, D. J.; Van de Plas, R.; Caprioli, R. M.; Spraggins, J. M. *Anal. Chem.* **2018**, *90*, 5090-5099.
- (19) imzML--a common data format for the flexible exchange and processing of mass spectrometry imaging data, Schramm, T.; Hester, A.; Klinkert, I.; Both, J. P.; Heeren, R. M.; Brunelle, A.; Laprevote, O.; Desbenoit, N.; Robbe, M. F.; Stoeckli, M.; Spengler, B.; Rompp, A. *J. Proteomics* **2012**, *75*, 5106-5110.
- (20) Cardinal: an R package for statistical analysis of mass spectrometry-based imaging experiments, Bemis, K. D.; Harry, A.; Eberlin, L. S.; Ferreira, C.; van de Ven, S. M.; Mallick, P.; Stolowitz, M.; Vitek, O. *Bioinformatics* **2015**, *31*, 2418-2420.

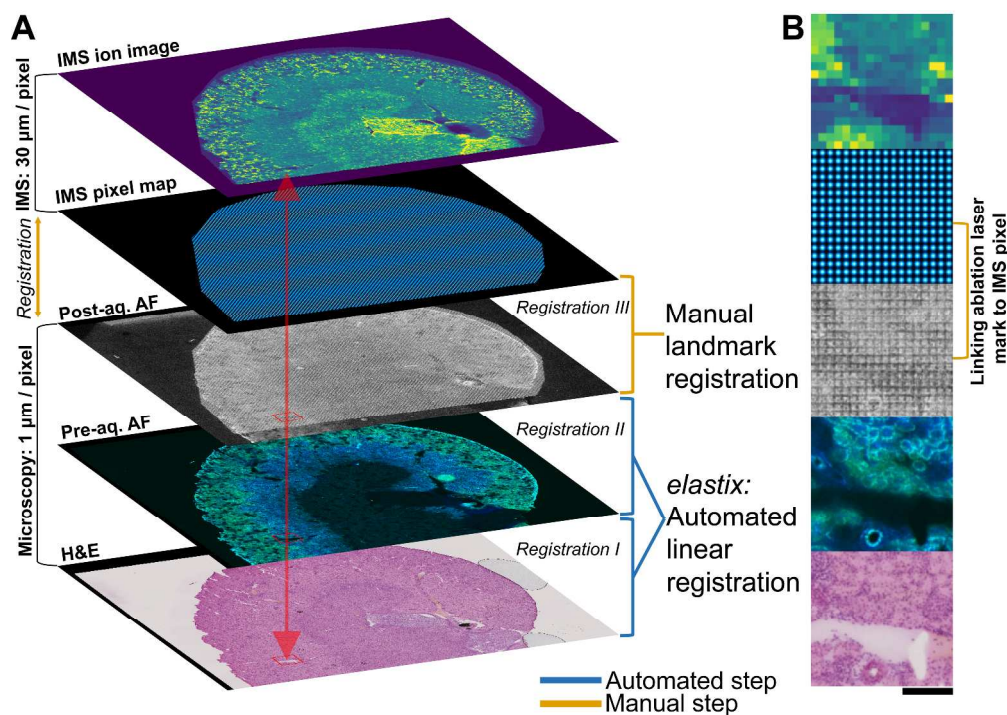


Figure 1. Microscopy to IMS single tissue section registration workflow shown using a rat brain cerebellum example ROI. (A) Low magnification view of the tissue images. (B) High magnification ROI of the images for registration. Registration proceeds from the bottom towards the top. Registration I: Registration between secondary microscopy (H&E staining in example) and pre-acquisition AF image. Registration II: registration between pre-acquisition AF and post-acquisition AF. Registration III: Manual registration between post-acquisition (laser ablation marks) and corresponding IMS pixels. Scale bar 150  $\mu\text{m}$ . Pre-acq: pre-acquisition. Post-acq: post-acquisition.

671x475mm (150 x 150 DPI)

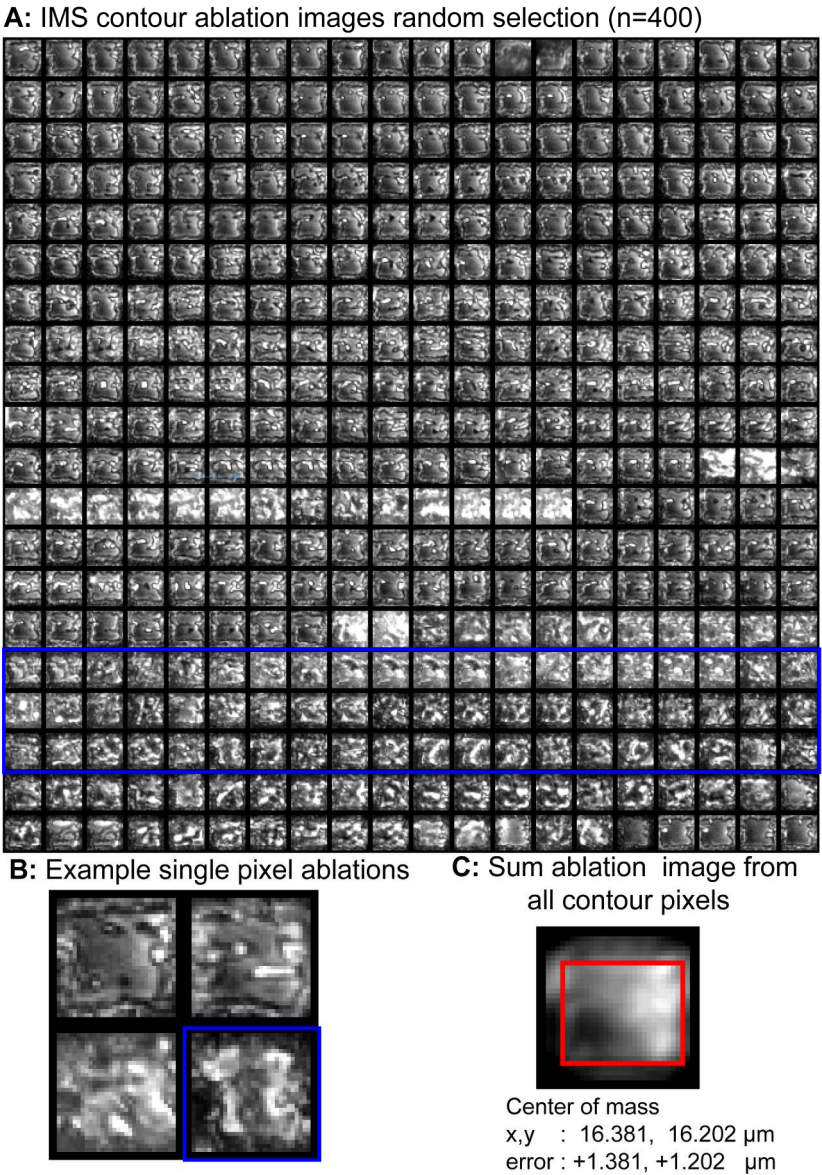


Figure 2. Quantitative and qualitative analysis of the IMS ROI's contour pixels to determine IMS to microscopy registration accuracy applied to a 30  $\mu\text{m}$  spatial resolution rat kidney dataset acquired on the Bruker rapifleX. (A) Images of 400 randomly selected individual IMS contour pixels out of 946 of the underlying post-acquisition AF with laser ablation craters visible. The blue rectangle indicates 3 rows where the contour is over the tissue and not in the matrix surrounding the tissue. The image's contrast was increased for display purposes. (B) Single pixel examples where top row is off tissue and bottom row is on tissue, blue here indicates a pixel on the tissue surface versus in surrounding matrix. (C) Sum image of all contour pixels from the dataset with a red outline indicating the center of the ablation where error is calculated. In A-C, each tile is 30  $\mu\text{m}$ .

318x466mm (150 x 150 DPI)



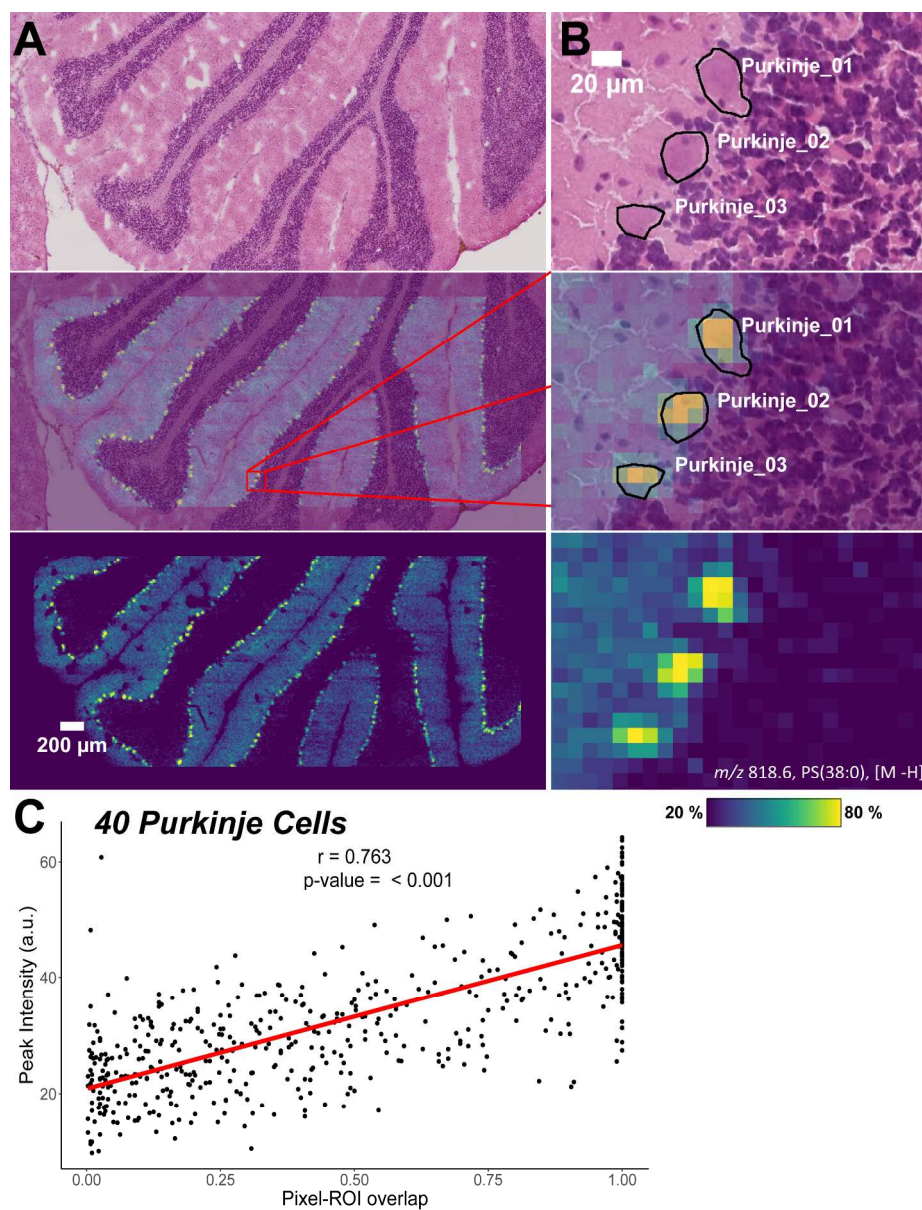


Figure 3. Analysis of signal from Purkinje cells following AF based IMS registration to H&E of the same section. (A) Low magnification of ion image, H&E and their overlay. (B) High magnification images 3 annotated Purkinje cells in the rat brain section. (C) Scatter plot of signal intensity and IMS pixel - ROI overlap from 40 annotated Purkinje cells with a linear regression line.  $r$  = Pearson correlation coefficient. P-value rejects null hypothesis that  $r = 0$ , no correlation between pixel overlap and signal of  $m/z$  818.5 exists.

421x548mm (150 x 150 DPI)

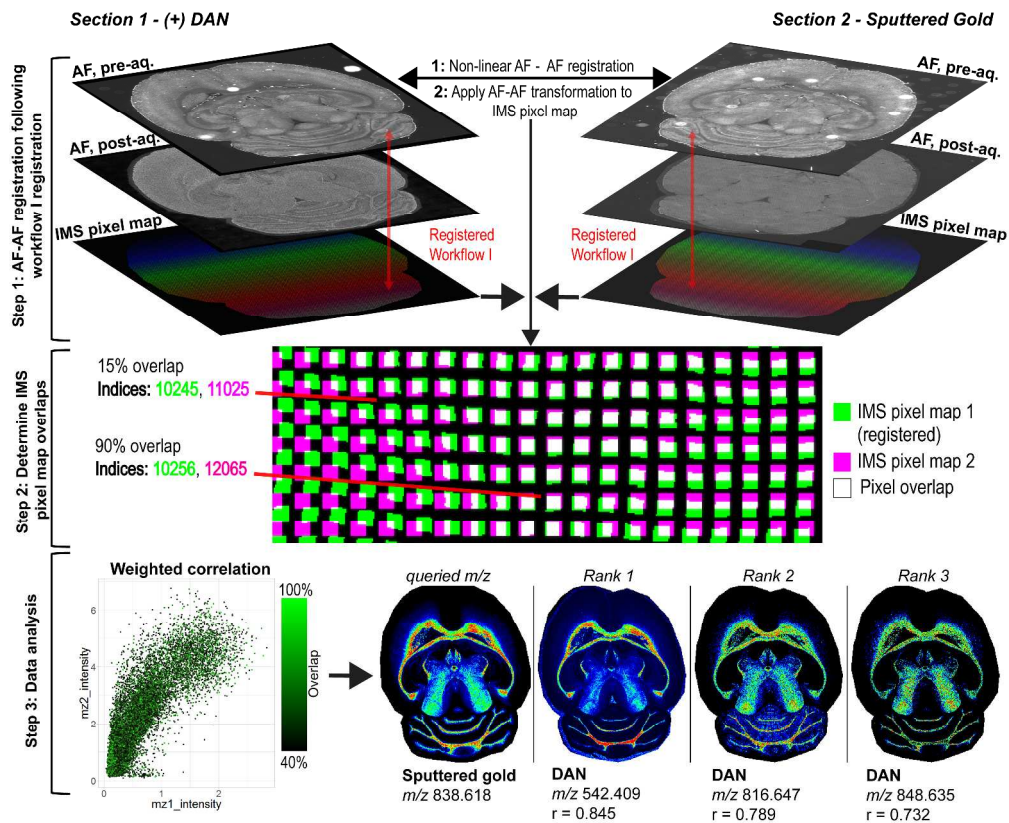


Figure 4. Workflow for automated discovery of correlated ions from serial sections using AF microscopy registration. Step 1: First, microscopy images are registered to IMS following the method detailed in Fig. 1. Then microscopy layer images from the serial sections are registered together using non-linear models with the resulting transformation applied to the upscaled IMS pixel mask. Step 2: This step finds the IMS pixels that overlap as well as quantifies any overlap after registration. The overlapping IMS pixel indices are collated with the percentage of overlap as a table. Step 3: Data analysis is carried out by first extracting data matrices from the two IMS datasets where observations are overlapping pixels. A second vector contains the degree of overlap for each observation that is used to weigh the observations for correlation (left). The correlation matrix is then queried with an ion of interest from one dataset and the highest correlated ions from the second dataset are returned (right).

501x408mm (300 x 300 DPI)

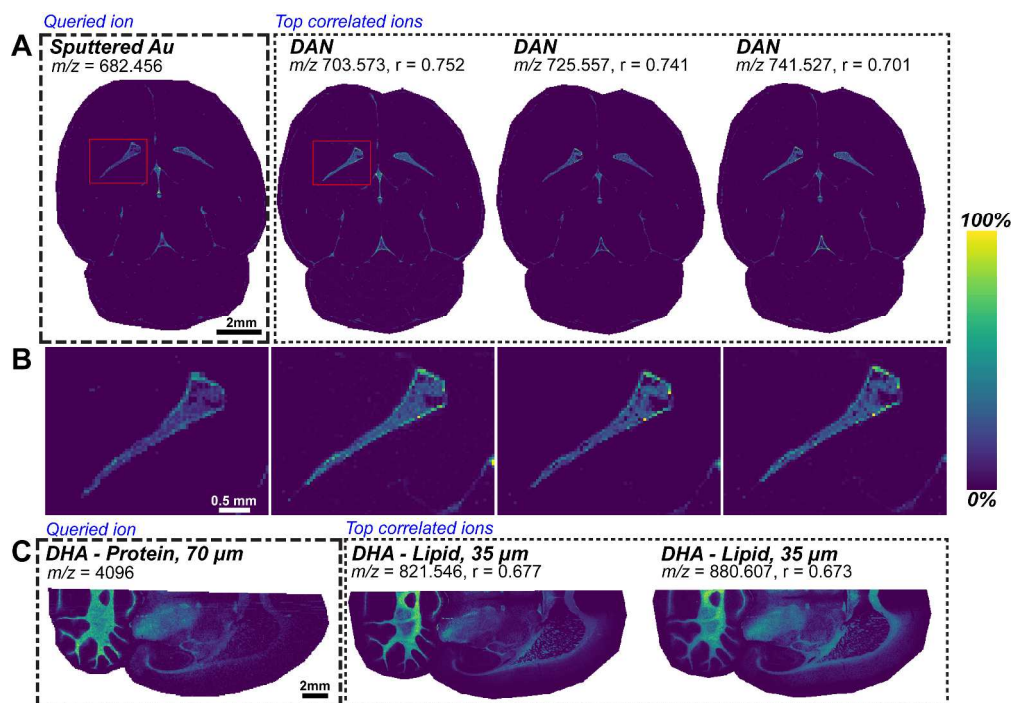


Figure 5. Automated spatial co-localization of ions from different datasets by high-resolution microscopy image registration. All ion images are scaled from 0 to their max intensity individually. (A) Full size images from two separate datasets, showing a queried ion from the sputtered gold sample acquisition and the species with the highest spatial co-localization of it from the DAN acquisition. (B) Zoom in areas from the highlighted red boxes showing the fine structure at the brain ventricles. (C) A second dataset for co-localization of lipids and proteins from serial sections using different IMS spatial resolutions (DHA - Protein =  $75 \mu\text{m}$ , DHA - Lipid =  $35 \mu\text{m}$ ).

422x298mm (300 x 300 DPI)



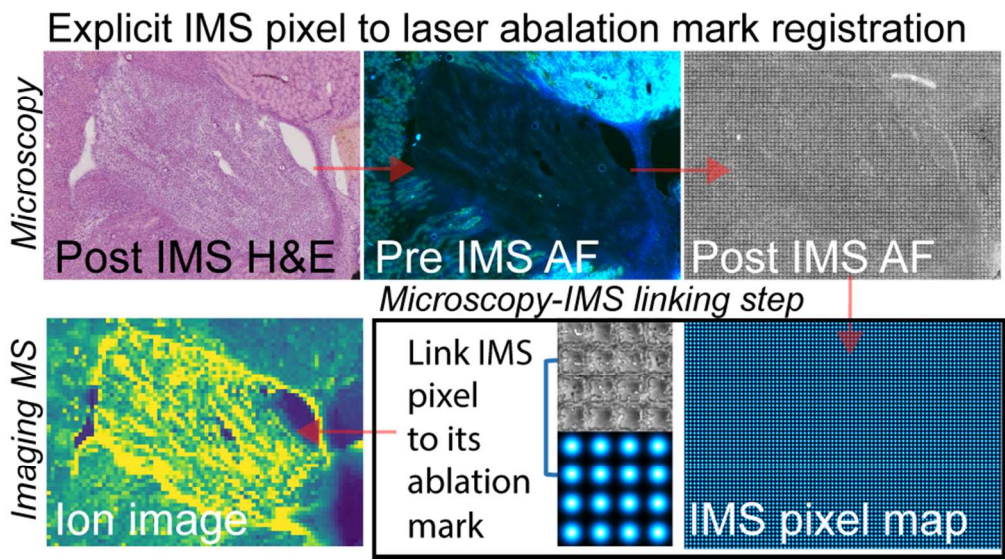


Table of contents graphic

84x46mm (300 x 300 DPI)