Spatial Chemometrics and Comprehensive Chemical Imaging based Molecular Histopathology Delineates Anatomical Heterogeneity at Single Pixel Resolution

Patrick M. Wehrli [1], Wojciech Michno [1], Laurent Guerard [2], Julia Fernandez-Rodriguez [2], Anders Bergh[3], Kaj Blennow [1,4], Henrik Zetterberg [1,4,5,6] and Jörg Hanrieder [1,5]

[1] Department of Psychiatry and Neurochemistry, Sahlgrenska Academy at the University of Gothenburg, Mölndal, Sweden

[2] Center of Cellular Imaging, Core Facilities, Sahlgrenska Academy at the University of Gothenburg, Sweden

[3] Department of Pathology, Umeå University, Umeå, Sweden

[4] Clinical Neurochemistry Laboratory, Sahlgrenska University Hospital, Mölndal, Sweden

[5] Department of Neurodegenerative Disease, Queen Square Institute of Neurology, University College London, London, United Kingdom

[6] UK Dementia Research Institute at UCL, London, United Kingdom

Corresponding author:

Jörg Hanrieder, PhD

Department of Psychiatry and Neurochemistry, Sahlgrenska Academy at the University of Gothenburg, Mölndal Hospital, House V, S-431 80 Mölndal, Sweden

jh@gu.se

Abstract

Imaging mass spectrometry (IMS) is a powerful tool for spatially-resolved chemical analysis and thereby offers novel perspectives for applications in biology and medicine. The understanding of chemically complex systems, such as biological tissues, benefits from the combination of multiple imaging modalities contributing with complementary molecular information. Effective analysis and interpretation of multimodal IMS data is challenging and requires both, precise alignment and combination of the imaging data as well as suitable statistical analysis methods to identify cross-modal correlations. Commonly applied IMS data analysis methods include qualitative comparative analysis where cross-modal interpretation is subject to human judgement; Workflows that incorporate image registration procedures are usually applied for co-representing data rather than to mine data across modalities.

Here, we present an IMS-based, histology-driven strategy for comprehensive interrogation of biological tissues by spatial chemometrics. Our workflow implements a 1+1-evolutionary image registration method enabling direct correlation of chemical information across multiple modalities at single pixel resolution. Comprehensive multimodal imaging data were evaluated using a novel approach based on orthogonal multiblock component analysis (OnPLS). Finally, we present a novel image fusion method by implementing consecutively acquired pathological staining data to enhance histological interpretation.

We demonstrate the method's potential in two biomedical applications where trimodal matrix-assisted laser desorption/ionization (MALDI) IMS delineates pathology associated co-localization patterns of lipids and proteins in (1) a transgenic Alzheimer's disease (AD) mouse model, and in (2) a human xenograft rat model of prostate cancer. The presented image analysis paradigm allows to comprehensively interrogate complex biological systems with single pixel resolution at cellular length scales.

KEYWORDS: Imaging Mass Spectrometry (IMS), matrix assisted laser desorption/ionization (MALDI), spatial chemometrics, multivariate data analysis, correlative imaging, image registration, image fusion

Introduction

Over the last years, imaging mass spectrometry (IMS) has emerged as a powerful tool for chemical imaging to increase understanding of spatial biochemical distribution dynamics in tissue that are associated with histopathological processes.[1,2] The investigation of complex molecular systems in biological tissues benefits from the combination of multiple chemical imaging modalities contributing with complementary molecular information, specifically multimodal IMS, vibrational spectroscopy, magnetic resonance imaging, as well as fluorescence microscopy.[3-7] The acquisition of imaging data in multiple modalities, however, yield datasets that may be misaligned. In order to concatenate misaligned datasets for effective data analysis across modalities, the datasets need to be registered to one another, meaning precisely aligned and distortions corrected.[8] However, image registration of IMS data is particularly challenging due to feature appearance as a result of noise and low contrast. Therefore, commonly applied IMS data analysis methods include qualitative comparative analysis where cross-modal interpretation is subject to human judgement.[9] Workflows that incorporate registration procedures are usually applied for co-representing data rather than to mine data across modalities.[10,11] Moreover, statistical evaluation is then typically performed on averaged data from assigned regions of interest at the expense of spatial information.[12] This highlights the need for advanced bioinformatics tools for both accurate alignment and combination of entire imaging datasets and unbiased data mining strategies for the interpretation of complex multimodal data at the single pixel level.[13,14]

The aim of the present study was, therefore, to develop a spatial chemometrics strategy for comprehensive analysis of multimodal IMS and microscopy data to interrogate molecular histopathology in complex biological tissues. Our workflow implements effective data processing and semi-automated image registration enabling direct correlation of chemical information across multiple modalities while minimizing human bias. We demonstrate the method's potential in two applications using matrix-assisted laser desorption/ionization (MALDI) IMS to investigate pathology associated distribution patterns of lipids and proteins in (1) hippocampus of a transgenic mouse model of Alzheimer's disease (AD), and (2) in a rat model of prostate cancer. Spatial chemometrics analysis revealed spatial organizations and correlations of lipids and peptides pertaining to disease pathology. The presented correlative imaging methodology provides the opportunity to investigate complex biological systems comprehensively and presents a valuable tool for their elucidation.

Materials and Methods

Animals and Tissue Preparation

Transgenic AD mice carrying the Swedish mutation in APP (tgAPP$_{SWE}$) were reared *ad libitum* at an animal facility at Uppsala University under a 12/12 light cycle. Fresh brain tissue samples were obtained from female, 18-month-old C57BL/6 mice. Animals were anesthetized with isoflurane and sacrificed by decapitation. The brains were dissected quickly with less than 3 min post mortem delay and frozen on dry ice. Animal procedures were approved by an ethical committee and performed in compliance with national and local animal care and use guidelines (DNr #C17∕14 at Uppsala University).

For rat model of prostate cancer, samples from adult animals were generated as described before.[15] Briefly, immunocompetent Copenhagen rats were anesthetized, and an incision was

4

made in the lower abdomen to expose the ventral prostate lobes. Then, 1 x 10^3 of highly metastatic Dunning MatLyLu (MLL) cancer cells (suspended in 10 µL RPMI 1640) were carefully injected into one of the ventral prostate lobes using a Hamilton syringe with a 30G needle. Rats were sacrificed 10 days after tumor cell injection, the tumor-containing prostates were removed, weighed, frozen in liquid nitrogen, and stored at -80°C.

Frozen tissue sections (12 µm) were cut in a cryostat microtome (Leica CM 1520, Leica Biosystems, Nussloch, Germany) at -18°C, and collected on indium tin oxide (ITO) conductive glass slides (Bruker Daltonics, Bremen, Germany) and stored at -80°C. Prior to analysis, tissue sections were thawed under vacuum for 1 hour.

H&E staining and brightfield microscopy

The hematoxylin and eosin (H&E) staining of consecutive sections to those analyzed with IMS was executed as followed. The tissue was placed in Mayer´s hematoxylin (Bio-Optica) for 2 min and washed twice with water. Bluing was done in in ammonia water solution for 1 minute. Thereafter, the tissue was washed twice with water, counterstained in 0.2% eosin G for 3 min, and again washed in water. The tissue was then dehydrated in 70% EtOH, 99.5% EtOH, and finally xylene. Image acquisition was done using a wide-field microscope (Zeiss Axio Observer Z1).

MALDI imaging MS

For MALDI imaging of lipids, 1,5 di-amino-naphthalene (1,5-DAN) matrix was applied to the tissue sections using a TM sprayer (HTX Technologies, Carrboro, NC, USA) combined with a HPLC pump (Dionex P-580, Sunnyvale, CA, USA). Before spraying, the solvent pump was purged with 70% acetonitrile (ACN) at 500µL/min for 10 min followed by, manual rinse of matrix loading loop using a syringe. A matrix solution containing 20 mg/mL 1,5-DAN in 70% ACN was sprayed onto the tissue sections with the following instrumental parameters: nitrogen flow (12 psi), spray

temperature (80°C), nozzle height (40 mm), five passes with offsets and rotations, and spray velocity (1250 mm/min), and isocratic flow of 50 µL/min using 70% ACN as pushing solvent. MALDI-IMS was performed on a MALDI TOF/TOF UltrafleXtreme mass spectrometer equipped with SmartBeam II Nd:YAG/355 nm laser. For dual polarity lipid analysis acquisitions were performed in reflective ion mode over a mass range of 200–2500 Da. Here, IMS data was first acquired with 20 laser shots per pixel with a source accelerating voltage of -20 kV (reflector negative mode, RN). This was followed by 50 laser shots per pixel point with source accelerating voltage of 25 kV (reflector positive mode, RP). The mass resolution at m/z 800 was of M/ΔM 20000. External calibration was carried out using peptide calibration standard I (Bruker Daltonics). For protein imaging, we employed a previously validated protocol for robust peptide and protein imaging.[16] In detail, tissue sections were subjected to sequential washes of 99.9% EtOH (Cat # V002075; Sigma Aldrich) (60 s), 70% EtOH (30 s), Carnoy's fluid (6:3:1 EtOH/CHCl$_3$/acetic acid) (90 s), 99.9% EtOH (15 s), H$_2$O with 0.2% trifluoroacetic acid (TFA) (Cat #T6508; Sigma Aldrich) (60 s), and 99.9% EtOH (15 s). For analysis of transgenic mouse model of AD, tissue was exposed to concentrated formic acid vapor in order to enhance Aβ peptide signal, which, however, compromises detection of other intact protein signals.[5,17] 2,5-Dihydroxy-acetophenone (2,5-DHA) was used as matrix compound (Cat.#: D107603, Sigma Aldrich) and applied using a TM Sprayer (HTX Technologies, Chapel Hill, NC, USA). A matrix solution of 15 mg/mL 2,5-DHA in 70% ACN/2%CH$_3$COOH/2%TFA was sprayed onto the tissue sections using the following instrumental parameters: nitrogen flow (10 psi), spray temperature (75°C), nozzle height (40 mm), eight passes with offsets and rotations, and spray velocity (1000 mm/min), and isocratic flow of 100 µL/min using 70% ACN as pushing solvent. Protein MS data of the same region as dual-polarity lipids were acquired over a mass range of 2–20 kDa, running in linear positive mode. The apparent mass resolution was m/Δm=1000 (FWHM) at m/z 4515, which was well in line with the reference values for this mass range provided by the manufacturer. A number of 100 laser shots/raster spot were acquired at 1 kHz laser repetition rate. The laser beam focus was set to "minimum" for AD

mouse model, resulting in a lateral resolution of 15 µm, and to "small" for rat model of prostate cancer, resulting in a lateral resolution of 30 µm. External calibration was performed from calibrant solution spots (Protein Calibration Mix 1, Bruker Daltonics) that were placed next to the tissue slides.

Data processing

Data processing was performed in MATLAB R2019a with Bioinformatics Toolbox 4.12 and Image Processing Toolbox 10.4 (MathWorks, Inc.) installed. MALDI imaging data were exported from SCiLS Lab (version 2019c, Bruker Daltonics, Bremen, Germany) in .imzML format and imported into MATLAB using the imzMLConverter by Race et al.[18] MS data were processed by baseline correction and normalization to total ion current (TIC). Variables in MS spectra were reduced using a peak picking routine extracting peak data at full mass resolution. For the extraction of data for region of interest (ROI) analysis, an ROI was selected using the imageSegmenter() function (Image Processing Toolbox) to then subset the dataset to the ROI's boundaries while retaining pixel coordinates.

Image data registration

The workflow for the alignment of IMS modalities involved an affine transformation matrix which was estimated through semi-automated image registration based on PCA scores images as reference images for each modality. The reference images were centered and coarse aligned prior to the automated registration procedure. The automated image registration method utilized an intensity-based optimization approach particularly suited for multimodal applications. The optimization algorithm employed a 1+1-evolutionary optimizer with various settings paired with Mattes Mutual Information metric configuration. Thereby, the geometric transformation is estimated by minimizing the mean square error (MSE) between a fixed reference image (I, with m rows and n columns) and moving image (K) that is to be transformed (eq 1).

7

$$MSE = \frac{1}{m\,n} \sum_{i=1}^{m} \sum_{j=1}^{n} [I(i,j) - K(i,j)]^2$$

(eq 1)

Optimizer parameters involving the search radius (including growth factor, initial and minimal size) were determined via full factorial design (3 x 3) at fixed maximum iterations (n = 300). Optimizer settings and metric configurations including factorial design limits are provided in the supplementary information Table S1. The number of maximum iterations was increased (to n = 5000) for the final rendition of the transformation matrix. In order to register the IMS data, the appropriate transformation matrix was applied directly onto the IMS data cubes using the imwarp function with bicubic interpolation. Manual image registration via fiducial point selection for comparative study was done using the control point selection tool (cpselect() function) of MATLAB's Image Processing Toolbox. Five control points were selected with even spread over the tissue surface at recognizable hallmarks in either image and at variable locations between replicates (n=5).

Multivariate modeling and visualization

Registered data were cropped to common area, reshaped and concatenated into a multimodal dataset prior to chemometrics analysis. Pixels with a total ion current of zero (black pixels) from, for example, off-sample acquisition or pixels from outside irregularly shaped acquisition areas, were omitted from data analysis. Black pixel (bp) and data pixel (dp) matrices were extracted by logical indexing using indices to sums of the multimodal dataset ($x_{i,j}$) that satisfy the equalities in equation 2.

$$\begin{cases} I_{bp} = \sum_{j=1}^{n} X_{i,j} \leq 0 \\ I_{dp} = \sum_{j=1}^{n} X_{i,j} > 0 \end{cases}$$

(eq 2)

8

Multivariate analyses of IMS data were performed in SIMCA 15/16 (Sartorius Stedim Biotech, Umeå, Sweden) and included principal components analysis (PCA), two-way orthogonal partial least squares (O2PLS), and multiblock orthogonal projections to latent structures (OnPLS). OnPLS components describe variation in globally joint, locally joint, and unique parts for each data block as follows:

$$X_i = \underbrace{X_{global}}_{\substack{\text{globally joint} \\ \text{variation}}} + \underbrace{X_{local}}_{\substack{\text{locally joint} \\ \text{variation}}} + \underbrace{X_{unique}}_{\substack{\text{unique} \\ \text{variation}}} + \underbrace{E}_{\substack{\text{residual} \\ \text{noise}}}$$

(eq 3)

Thereby, global variation implies structure that is shared between all blocks, local variation between at least two blocks, and unique variation occurs in only one data block.[19,20] Data were mean centered for data mining and scaled to unit variance for predictive modeling. The number of evaluated components was based on the predictive performance as determined by SIMCA's seven-block cross-validation. Component scores matrices were transferred to MATLAB, where they were reunited with black pixels and reshaped into image dimensions prior to scores image visualization. Respective loadings for the interpretation of the scores images were generated in SIMCA software.

High resolution predictions of scores and single ion images were performed using an Orthogonal Projections to Latent Structures (OPLS) model based on a decomposition matrix of the microscopy image (X-block) and a low resolution image from the IMS domain (Y).[21] Thereby, variables of the microscopy image (RGB, - red, green, blue) were extended with variables from color space transformations and PCA scores of all color bands, a strategy previously used for IMS data fusion.[10] Color space transformation included RGB to YIQ, and RGB to YCbCr using the rgb2ntsc() and rgb2ycbcr() functions, respectively. The decomposition matrix was down-sampled by affine transformation to the IMS resolution. Predictions of the microscopy resolution were performed using the decomposition matrix at its original resolution. In order to display

distribution heterogeneity not contained in the microscopy modality, the prediction image was weighted through multiplication with the squared corresponding IMS image matrix.

Results & Discussion

Multimodal IMS is an emerging, powerful approach in molecular medicine to interrogate complex biochemical patterns associated with anatomical features and histopathology. The presented methodology aligns and combines multimodal IMS data enabling cross-modal multivariate analysis at the measured image resolution. The data processing routine includes the pretreatment of data, alignment of the imaging modalities, combination of the datasets, and advanced spatial chemometrics analysis. We demonstrate the workflow on two different sets of MALDI IMS data to investigate distributions and correlations as well as co-localizations of lipids and proteins related to disease pathology in (1) hippocampus of a transgenic mouse model of Alzheimer's disease, and (2) prostate in a rat model of prostate cancer. Moreover, we present an image fusion strategy based on OPLS prediction to visualize IMS ion distributions at microscopy image resolution.

3.1 Intensity based automated alignment of multimodal IMS data

A simple and commonly applied approach to multimodal IMS data is the combination of data from positive and negative ion mode that was acquired on the same tissue without interruption. In that case, the pixels' coordinate systems of the two datasets correspond exactly and the data can be united into one multimodal dataset without the need for alignment of corresponding pixels by image registration. The combination of data from both ion modes can be done straightforwardly and the added information through multimodal enrichment can benefit tissue delineation by multivariate analysis.[22]

While the combination of IMS data acquired from the same pixel array can be done directly, discontinuously acquired multimodal IMS data are commonly spatially misaligned. Repeated measurement of the same tissue may be done with interruptions, for example, for the application of a different matrix and other treatments. In such cases, the resulting imaging data will likely be misaligned and distorted as a result from, for example, tissue shrinkage during washing procedures. Multimodal data can be obtained from a single tissue section, however, in cases where not all modalities can be acquired on a single section, consecutive sections are required. In such scenarios the imaging data need to be precisely registered to each other (aligned and features matched) in order to achieve alignment of the corresponding pixels in each dataset.[14] An accurate pixel correspondence between datasets is critical for downstream spatial chemometrics analysis of the combined imaging data.

Our first aim was, therefore, to establish an image registration workflow for precise alignment of imaging data acquired in different modalities. Image registration can be done using manual selection of fiducial points. However, the manual selection of control points with high confidence is challenging, particularly on IMS data due to feature appearance. As a consequence, it is difficult to achieve a registration accuracy appropriate for the investigation of small features such as, for example, amyloid beta (Aβ) plaques (50-100 µm) in Alzheimer's disease pathology. An automated method for the registration of IMS data based on a gradient descent algorithm has been presented.[23] This method has been demonstrated to work well for image data with tissue edges present. However, the alignment of small anatomical features without the presence of tissue edges was not successful for our data (no convergence). This highlights the need for an alternative image registration approach. We, therefore, developed a workflow for the registration of multimodal data involving an intensity based optimization algorithm. Our registration procedure starts with the generation of reference images for each IMS modality which are used to estimate the required geometric transformation (Figure 1a). To use single ion images as reference images

would be obvious. However, noise, contrast and variable clarity of features in single ion images pose a great challenge for image registration algorithms to achieve an adequate convergence result. Therefore, we used image PCA to capture the variance in the IMS datasets while separating noise. Resulting PCA scores images that are of high contrast and rich in features, lend themselves as suitable reference images for image registration.

We applied an intensity based, semi-automated image registration method to estimate the transformation required to align the corresponding pixels between two reference images and IMS modalities, respectively (Figure 1b). The method solves image registration problems through iterative optimization of a predefined dissimilarity metric using a 1+1-evolutionary optimizer. Thereby, the supplied optimizer parameters determine the success of the convergence. For a greater automation, we implemented a full factorial design to determine the optimal optimizer parameters. The computed transformation matrix is then applied to geometrically transform the IMS dataset (Figure 1c). Thereby, imaging data with different spatial resolution can be combined through interpolation of the lower resolution dataset. Finally, the datasets can be concatenated and analyzed as one multimodal dataset (Figure 1d).
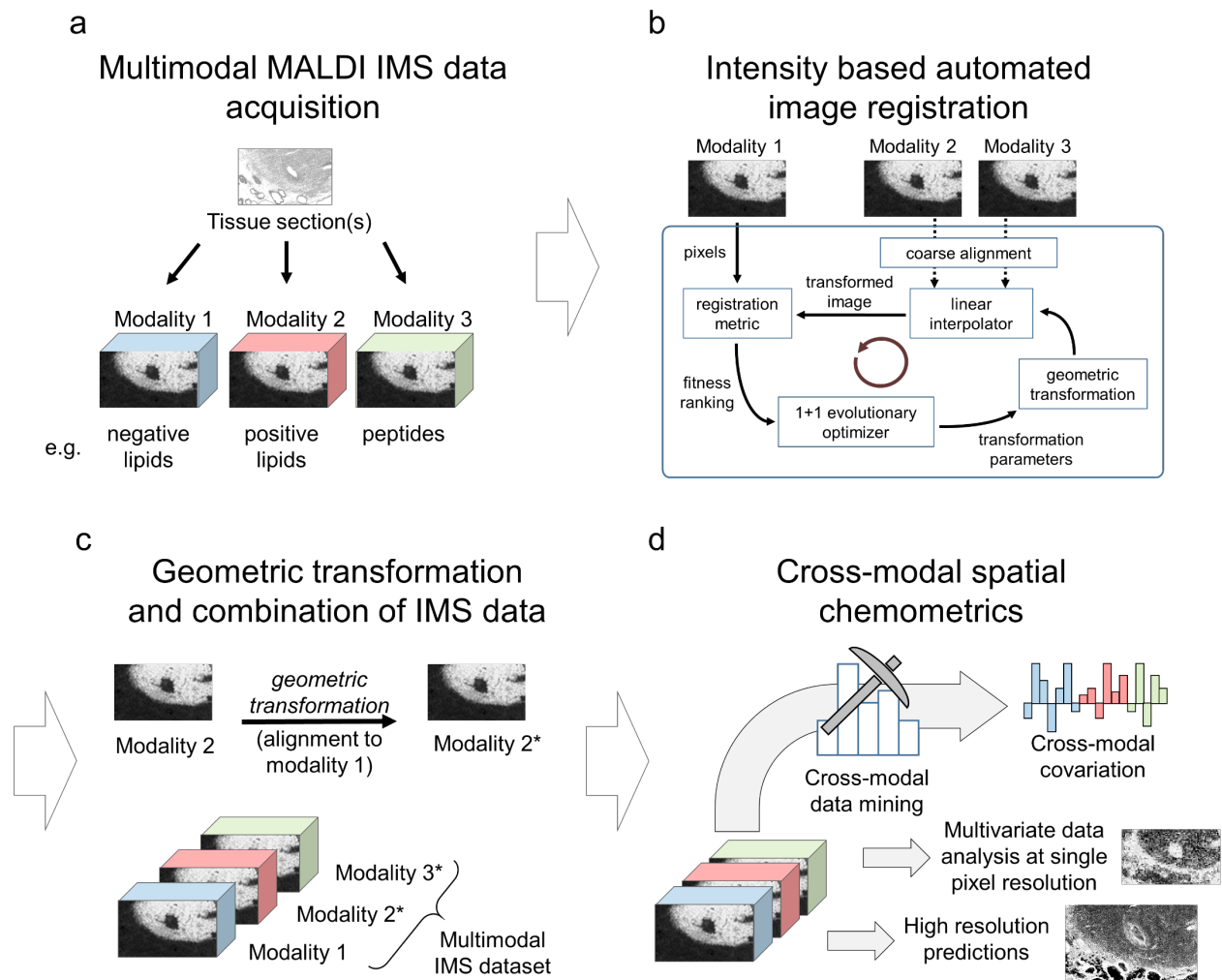
a

Multimodal MALDI IMS data acquisition

Tissue section(s)

e.g.

Modality 1   Modality 2   Modality 3

negative lipids   positive lipids   peptides

b

Intensity based automated image registration

Modality 1   Modality 2   Modality 3

pixels

coarse alignment

transformed image

registration metric   linear interpolator

fitness ranking

1+1 evolutionary optimizer

geometric transformation

transformation parameters

c

Geometric transformation and combination of IMS data

Modality 2

geometric transformation
(alignment to modality 1)

Modality 2*

Modality 3*
Modality 2*
Modality 1   Multimodal IMS dataset

d

Cross-modal spatial chemometrics

Cross-modal data mining

Cross-modal covariation

Multivariate data analysis at single pixel resolution

High resolution predictions

Figure 1. Overall workflow for the multimodal exploration of biological tissues by MALDI IMS and spatial chemometrics. (a) MALDI MS imaging of tissue sections in various modalities provides localizations of different biochemicals including lipids and peptides. (b) In order to combine imaging MS datasets acquired in different modalities, their corresponding pixels need to be precisely aligned. The presented procedure estimates the required geometric transformation using an intensity based semi-automated image registration approach particularly suited for multimodal applications. Feature-rich PCA scores images were used as reference images for image registration. (c) Geometric transformation allows the concatenation of IMS datasets, (d) Advanced spatial chemometrics analysis of the combined imaging data involves multivariate image analysis at the single pixel resolution including multiblock orthogonal component analysis to extract covariations between modalities. High resolution prediction by image fusion can be performed to enhance histological interpretation.

After establishing an automated image registration procedure, we aligned IMS data acquired from both, the same and consecutive tissue sections (Figure 2a). We went on to compare our semi-automated alignment approach to the commonly used manual method for image registration by fiducial point selection (Figure 2b). The accuracy of image registration results is frequently assessed using basic methods such as overlap scores, structural similarity score, and image difference.[24] However, it has been shown that those methods do not provide valid evidence for accurate registrations and should thus be avoided.[24] Therefore, in order to compare the registration accuracy of differentially registered data, we employed O2PLS modeling. O2PLS is a multivariate regression method that allows unsupervised modeling of the joint and unique variations in two blocks of data. The O2PLS method is bidirectional allowing the prediction of each of the data blocks from one another. Thereby, systematic variability in two data blocks is partitioned into a joint predictive variation and into variation that is unique to each block of data (Figure 2c).[25-27] O2PLS has been used in numerous studies including technical applications to model pre-processing effects on spectroscopic data[28], in batch process development[29] but also in systems biology to model variation between sets of "omics" data.[30-33]

Here, we used O2PLS in a first application to assess the pixel-to-pixel correspondence between the two registered IMS datasets taking advantage of their multivariate nature (Figure 2c). A more accurate registration is expected to lead to a better pixel-to-pixel correspondence indicated by stronger correlation between the spectra from each data block and, thus, to an O2PLS model predicting a higher fraction of variation of the Y variables (Q2Y). The predicted fraction of variation of the lipid data is increased with automatically registered peptide data suggesting a higher registration accuracy as compared to manual registration. Inaccurately registered and combined pixels, are expected to increase the O2PLS orthogonal (unrelated) variability in either of the data blocks (X and Y, that is, peptide and lipid data). We found that the orthogonal variability in manually registered data is higher as compared to automatically registered data corroborating a

better registration accuracy of the latter. The automated registration minimized the unrelated variation in either of the datasets resulting in about 9% of orthogonal variability in the negative lipid data and no (0%) orthogonal variation in the peptide data (Figure 2d).

Commonly, multimodal data are obtained from consecutive tissue sections (Figure 2a).[3,4] Therefore, we compared protein data from the same and the consecutive section to the corresponding lipid data using O2PLS modeling. The lipid and protein data collected on the same section resulted in a higher Q2Y value and a decreased orthogonal variability in either modality (Figure 2d). This indicates a better pixel-to-pixel correspondence between data blocks of the same section and a worse correspondence to the consecutive section as a result of intrinsic differences between the tissue sections. While concludes that acquisition of multimodal data from the same tissue is preferable, this might not always be possible due to different sample preparation and acquisition parameters that are incompatible. Taken together, the O2PLS modeling strategy allows to compare registration accuracies of both, alignment methods as well as consecutively acquired data from the same or adjacent sections. The here employed O2PLS approach enables relative comparison of registration and acquisition methods. However, it has to be assessed independently whether the different datasets are compatible with respect to tissue distortion and data quality.
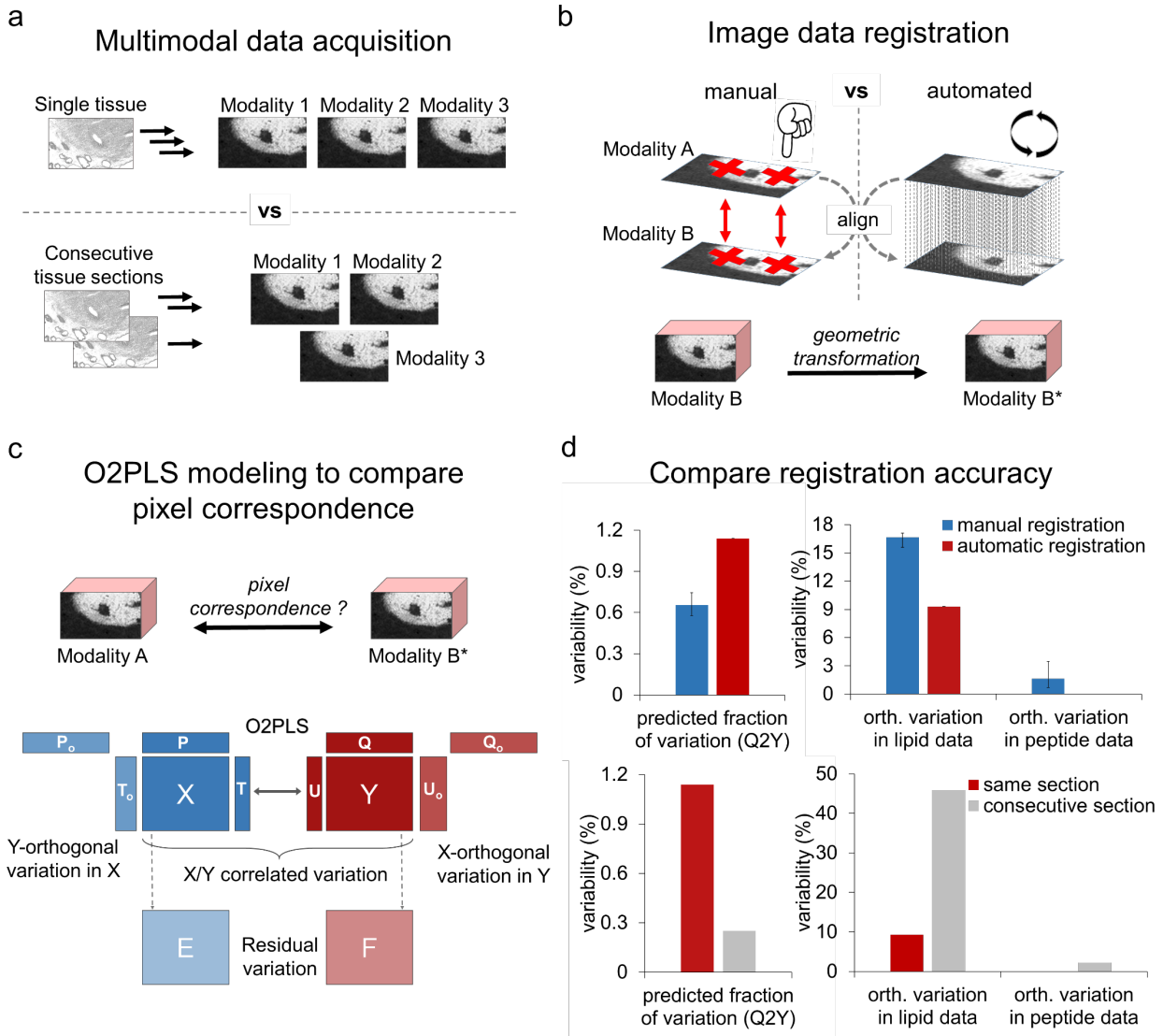
Figure 2. Strategy for the comparison of image registration accuracies using O2PLS modeling. a) Multimodal data can be obtained from a single tissue section or consecutive sections. (b) Commonly, image registration is done via manual fiducial point selection. In here, an intensity-based registration approach is used. (c) Comparison of pixel correspondence and, thereby, registration accuracy, using O2PLS modeling. O2PLS subdivides the systematic variability in the two datasets X and Y into X/Y correlated predictive variation and variation that is orthogonal in each data block. Scores matrices are depicted vertically and their corresponding loading matrices are drawn horizontally. (d) Using the O2PLS pixel correspondence strategy, manual vs automated registration methods are compared, as well as, data acquired on the same vs consecutive tissue section. Error bars show minimum and maximum values of replicates (n=5).

3.2 Classical correlation analysis of trimodal IMS data yields global correlations at the expense of spatial information

Precise alignment of various imaging MS data allows their combination into a multimodal dataset that incorporates comprehensive biochemical information provided by the different imaging modalities. Commonly, statistical analysis of IMS datasets is done using methods that lead to the loss of spatial information as a result of averaging multiple ROI spectral data. This, however, may be confounded by inter- and intra- feature heterogeneity. Therefore, retaining the spatial information at single pixel resolution is essential for the investigation of systems with heterogeneous pathological features.

For the present study, we investigated tissue pathology in the hippocampal brain area in a transgenic mouse model of Alzheimer's disease (tgAPP$_{Swe}$) using correlative multimodal imaging MS. Alignment and combination yielded a trimodal IMS dataset, retaining single pixel resolution without ROI averaging. As an initial approach for interrogating the combined multimodal data, we performed correlation analysis. Here, positive and negative correlations between lipids and peptides can be observed (Figure 3a-b). In order to obtain correlations specific to features of interest, ROI can be selected and analyzed. We selected an amyloid plaque for ROI correlation analysis yielding correlations of analytes pertaining to Alzheimer's disease pathology (Figure 3c-d). The results and interpretation are biased by the selection of the ROI area. Here, correlation analysis of large ROI such as whole tissue areas, result in convolution of individual correlation values due to non-exclusive localization of chemical species to different anatomical features. In turn, selection of smaller ROI focused on individual histological features yield statistical and biological significant chemical co-localizations (Figure 3c-d). This highlights the requirement for chemometrics tools that permit comprehensive interrogation of multimodal imaging data of entire heterogenic tissue areas, while maintaining single pixel resolution to reveal feature specific chemical co-localization patterns.
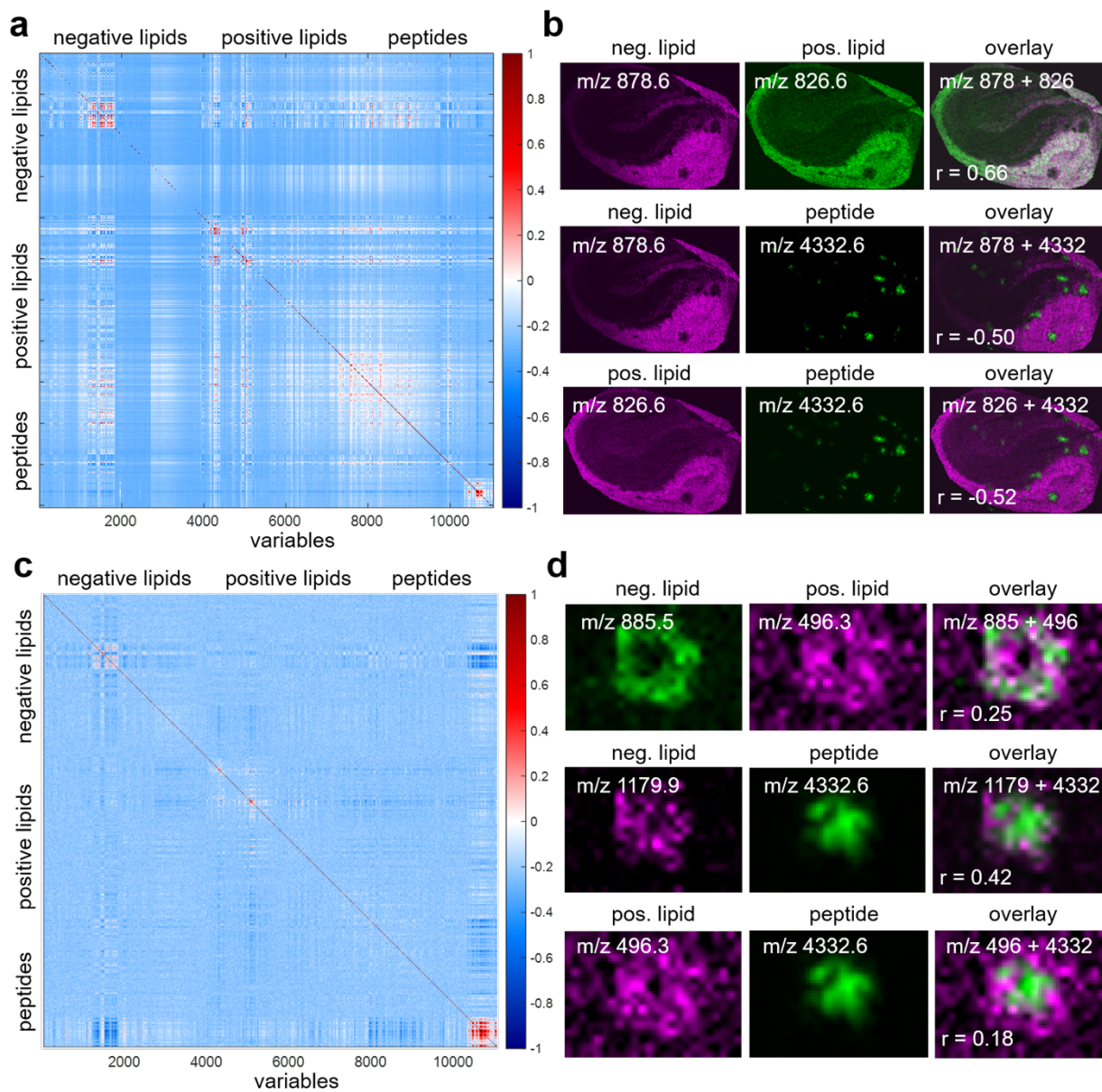
Figure 3. Classical correlation analysis of trimodal IMS data from mouse hippocampus with Alzheimer's disease pathology. a) Correlation plot of trimodal data displays global correlations within and between imaging modes. b) Ion images of correlating ions and overlays with correlation coefficients (Pearson). Area size 2.6 x 1.8 mm c) Correlation plot of the plaque ROI data displays ROI specific correlations. d) Ion images of the ROI and overlays with correlation coefficients. Area size 390 x 255 µm. Observed ions are consistent with following assignments: m/z 496 [PC(16:0/OH)+H]$^+$,[34] m/z 826 [PC(36:1)+K]$^+$,[34,35] m/z 878 [OH-ST d18:1/h22:0 -H]$^-$,[36] m/z 885 [PI(18:0/20:4)-H]$^-$,[35,36] m/z 1179 [GM3-H]$^-$, and m/z 4332 Aβ (1-40).

3.3 O2PLS modeling reveals the heterogeneity of pathological features and corresponding chemical signatures

Spatial chemometrics involves multivariate analysis of chemical data with spatial information to provide maximum relevant chemical insight. In multivariate analysis of IMS data, each pixel of the imaging dataset is treated as individual observations while retaining its spatial information (coordinates) during data analysis to be able to later reconstruct scores results into component scores images. Image PCA has previously been used to extract biochemical characteristics of tissues and cell cultures from IMS data.[37-40]

While PCA allows the extraction of features based on maximum variance, certain covariances particularly in between entire blocks of data require more advanced methods. Therefore, special attention should be paid to the unidirectional prediction methods O2PLS, and multiblock orthogonal component analysis based on the OnPLS algorithm.[19] As described above, O2PLS models the joint and unique variation between two blocks of data. Therefore, we applied O2PLS modeling to understand the relationship between lipids and peptides specifically and, further, the heterogeneity between and within pathological features captured by the cross-modal molecular imaging approach (Figure 4a).

In the present study, O2PLS modeling revealed amyloid plaque heterogeneity in the hippocampus of a mouse model of Alzheimer's disease (Figure 4b). Compositional variation of two plaques in particular, but also of certain core structures of other plaques, are consistent with reduced amounts of a number of amyloid peptides, most prominently Aβ1-38 (m/z 4135) and Aβ (m/z 4332) as well as increased levels of Aβ1-40 (m/z 4332). Inspection of single ion images of the high loading variables confirm these observations. There, the two plaques appear Aβ1-38 deficient while exhibiting increased levels of Aβ1-40 (m/z 4332) as compared to the majority of plaques (Figure 4c). These variations had not been recognized on PCA scores images of the same datasets (Supplementary Figure 1). While the profiles of amyloid peptides vary between

different plaques features (Figure 4b-c), lipid variations correlated to peptide data (u1 scores, Figure 4d) appear consistent among plaques, indicating that lipid profiles are independent of plaque peptide composition within the plaque structure. Furthermore, a relatively large fraction (51%) of unique variation in the lipid dataset indicates lipid profiles without covariance to the peptide data. In fact, orthogonal components describe lipid localizations in the periphery of plaque structures (u_o6, u_o7, u_o10) and also in the dentate gyrus (u_o10). Interestingly, the two large plaques deficient in Aβ1-38 do not exhibit these peripheral lipid localizations (u_o10 scores, Figure 4d and Supplementary Figure 2). These observations demonstrate the potential of O2PLS and highlight the importance of retaining spatial information over the whole tissue area at the single pixel resolution for the investigation of systems with heterogeneous features. Indeed, O2PLS has previously been used for applications in systems biology to model transcript, protein and metabolite data,[30] lipidomics[31] and for NMR metabolomics and DIGE proteomics data from xenograft prostate cancer mouse models.[32]
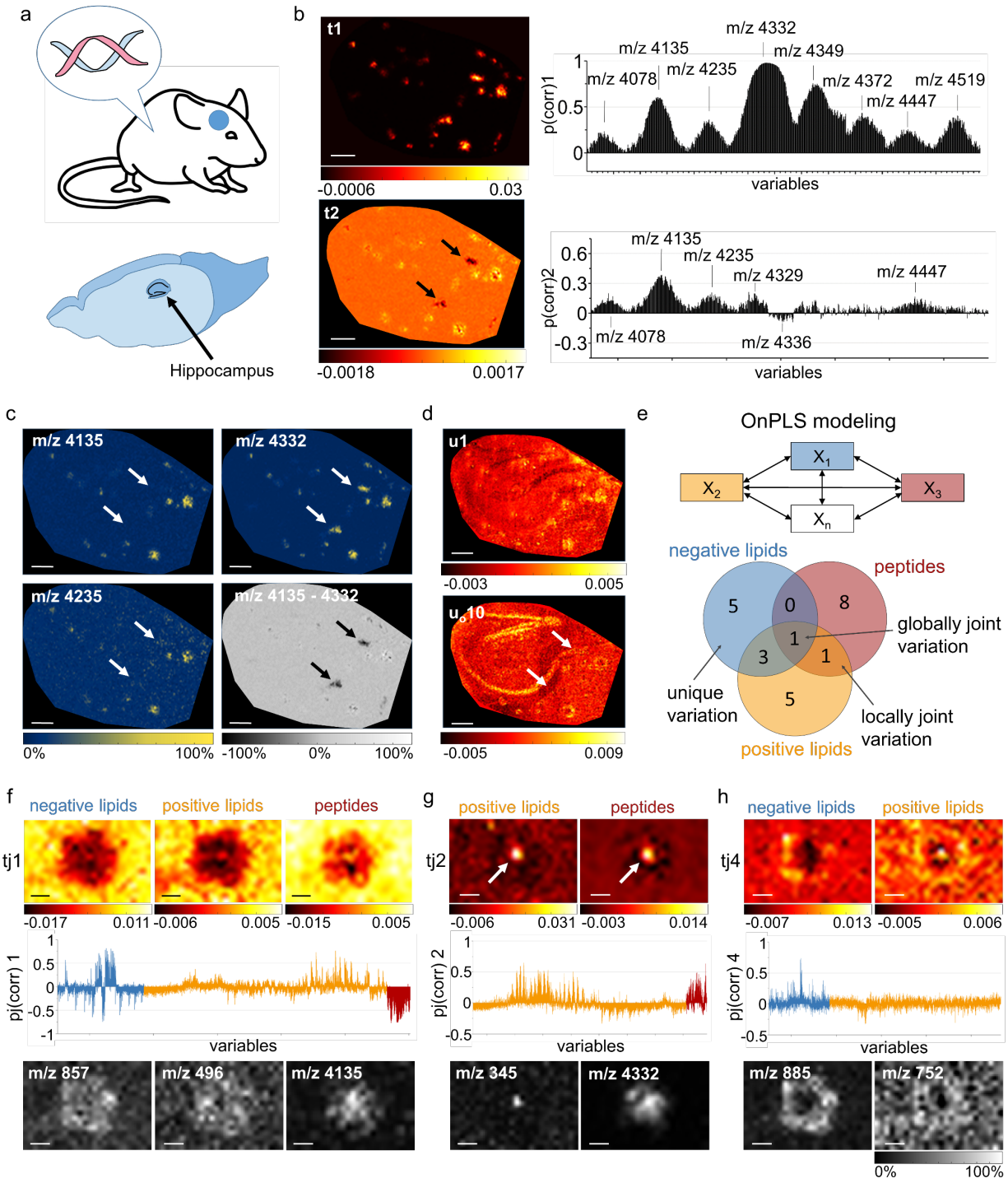
Figure 4. Spatial chemometrics analysis of multimodal MALDI IMS data obtained from mouse hippocampus with Alzheimer's disease pathology. a) Transgenic mouse model of Alzheimer's disease (tgSwe). b) O2PLS analysis of combined peptide and lipid data reveals plaque heterogeneity (t2 scores image) and corresponding chemical profiles (p2 loadings), R^2 = 0.857. c) Single ion images of high loading variables (O2PLS) confirm that two plaques in particular contain reduced levels of specific Aβ peptides, most prominently Aβ1-38 (m/z 4135),

whereas the most abundant peptide, Aβ1-42 (m/z 4332), is observed at similar levels. d) Scores images of the lipid data that correlated to peptide data show consistent lipid profiles between plaques (u1). Lipids localizing to plaque periphery and dentate gyrus are described in an orthogonal component (u_o10). There, the two large Aβ1-38 deficient plaques do not exhibit peripheral lipid localizations. e) Schematic representation of potential relationships between all matrices modeled by OnPLS. Number of OnPLS components from a region of interest analysis (single plaque) based on trimodal data. (f) Globally joint component scores image and loadings describing covariance among all three modalities. (g) Locally joint component of positive ion mode lipid data and peptide data displays a distinct feature in the core of the plaque. (h) Locally joint component of positive and negative ion mode lipid data appears to account for peripheral localizations. Normalized to total ion current, scale bars entire hippocampus = 300 μm, scale bars plaque ROI = 60 μm.

3.4 Intrinsic heterogeneity of histological features captured in small regions of interest

Following initial O2PLS modeling of the amyloid plaque pathology in the entire hippocampal tissue area we continued with a region of interest analysis of a single plaque feature. While the O2PLS model included lipid and peptide data as two blocks, further efforts were directed towards exploiting the trimodality of this dataset. Therefore, we employed multiblock component analysis, based on the OnPLS algorithm, to further explore the relationships between all three modalities in an unsupervised manner (Figure 4e).[19] While O2PLS models the joint and unique variation between two blocks of data, OnPLS allows the analysis of multiple data blocks. In the two-block case the two algorithms generate very similar results. OnPLS partitions the total variation into globally joint, locally joint, and unique parts. Global variation is shared between all data blocks of a multiblock dataset, local variation is shared between at least two blocks, and unique variation occurs in only one data block. OnPLS is a descriptive modeling technique with a purpose to reveal the relationships between multiple blocks of data and an aim to enhance interpretability of the results.[19,20,41]

From the single plaque ROI data, OnPLS generated one predictive component of the global variation and multiple non-global components (Figure 4e). Again, we applied this in the context of spatial chemometrics to visualize scores results as images. The globally joint component images

22

displayed an overall structure of the plaque as outlined by PI(36:4) (m/z 857), LPC (m/z 496), and Aβ1:38 (m/z 4135) (Figure 4f). The first locally joint component between positive lipid data and peptide data revealed a pronounced localization in the plaque's core including Aβ1:40 (m/z 4332) (Figure 4g). These spatially focused localizations are confirmed on single ion images and occur within a number of other plaques but had previously not received the attention as they potentially could be regarded as noise during visual inspection of the single ion images (Supplementary Figure 3). Further, locally joint variation between lipid modalities visualized lipid profiles localizing to the periphery of the plaque including PI(38:4) (m/z 885) and PE(38:6) (m/z 752) (Figure 4h). In conclusion, features of the scale as these core structures (30-40 μm) require the modalities to be precisely aligned for the extraction of feature specific cross-modal covariance to be successful. This emphasizes the importance a method for accurate image alignment and further illustrates the notable benefit of analyzing imaging data at single pixel resolution. Moreover, this demonstrates the discovery of biochemical co-localizations through multimodal enrichment that might have been mistaken for noise or go unnoticed if only one modality had been considered.

3.5 Spatial chemometrics exploration of rat prostate cancer

MALDI IMS has found important application in cancer research to investigate the high chemical complexity and heterogeneity of cancer tissue.[42,43] Therefore, to further demonstrate the potential of our workflow for correlative imaging, we investigated molecular histopathology in a human xenograft rat model of prostate cancer (Figure 5a). Multimodal imaging including MALDI IMS and H&E microscopy data were acquired from three consecutive tissue sections. While the tumor appearance was most conserved between sections, benign tissue in the tumor-bearing prostate showed distortions and intrinsic differences that were only partially matched (Figure 5b). Therefore, we focused our study on the tumor micro- and macro-environments with limitation to proximal surroundings. Spatial chemometrics analysis was then used to visualize the tumor heterogeneity in respect to covariance of the underlying chemical signatures. Image PCA of the

combined IMS data generated a component that appear to describe differences between cancer and benign tissue (t2), components that account for cancer heterogeneity (t6, t8), and also components that describe the surrounding tissue (t3, t5, t7) (Figure 5c). OnPLS modeling of the dataset generated components for joint variation and each case of locally joint variation, as well as unique variation (Figure 5d). The primary component (tj1), describing covariance between all three modalities, appears to account for molecular signatures (pj1) that distinguish cancerous and normal prostate tissue (Figure 5e). Here, specific loadings characteristic of the tumor were, for example, PI(38:4) (m/z 885), phosphatidyl choline PC(30:0) (m/z 706), and apolipoprotein C1 (m/z 6640). This observation demonstrates the potential of OnPLS, being an unsupervised method, as a valuable alternative to supervised modeling that require either manual or automated ROI annotation.

In conclusion, the visualization of scores from multivariate analyses, have the potential to reveal spatially defined chemical signatures that can provide valuable new insight in the heterogeneity of cancer and the interaction between the invasive tumor front and the surrounding tissue. Furthermore, spatial chemometrics of multimodal data can extract chemical profiles that may lead to the discovery of candidate biomarkers and create links between lipids and proteins that otherwise could be missed.
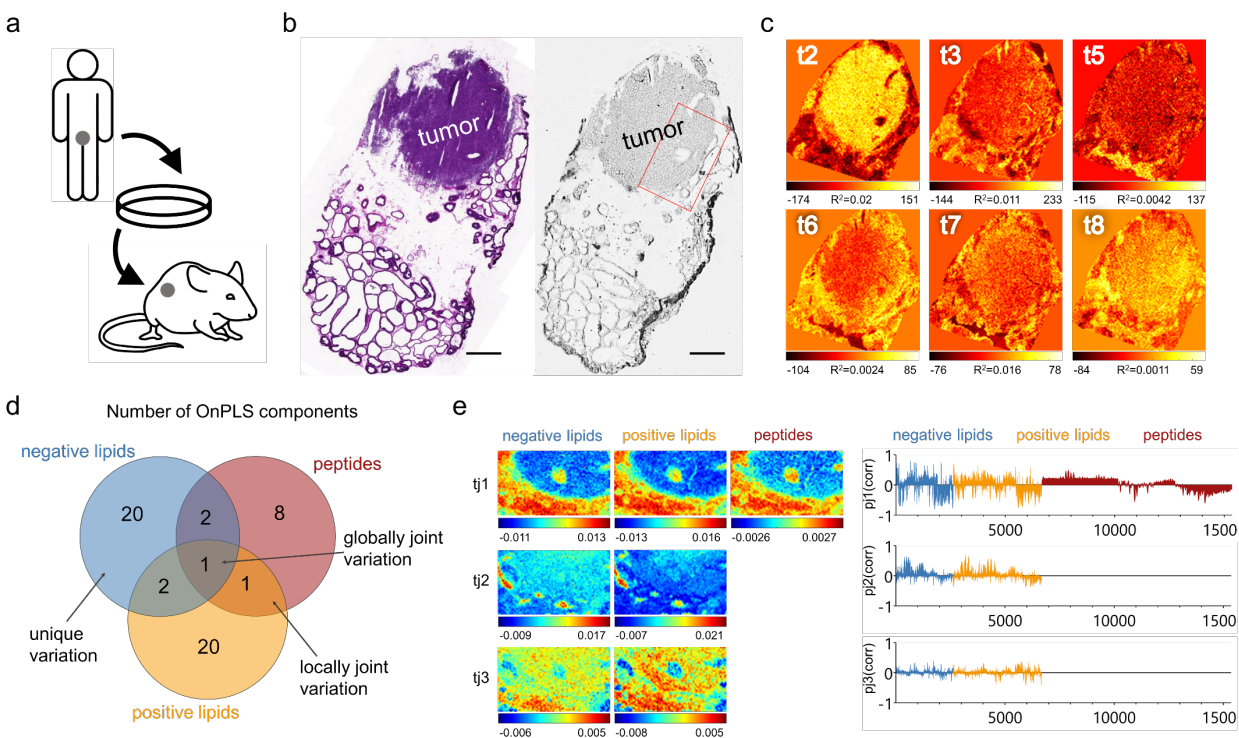
Figure 5. Spatial chemometrics exploration of trimodal IMS data of rat prostate cancer. (a) Highly metastatic (Dunning MLL) human prostate cancer xenograft rat model, (b) bright field microscopy image of H&E stained and (consecutive) native rat prostate tissue heavily affected by the tumor, scale bar = 800 µm, (c) PCA scores images generated from the trimodal MALDI IMS data (negative and positive ion mode lipids, and peptides, log transformed), (d) number of OnPLS components of joint and unique variation, (e) top three OnPLS joint t-scores images and loadings, whereby tj1 appears to account for cancer specific joint variation between all three modalities (R^2(negative lipids) 0.22, R^2(positive lipids) 0.11, R^2(peptides) 0.10.

3.6 Multimodal image fusion predicts correlated chemical information at high resolution

Information extracted by advanced data analyses can lead interpretation for improved understanding of the disease pathology. Pathological evaluation of tissue sections is commonly done using histological staining and brightfield microscopy at sub-micron resolution. However, information from MALDI IMS is usually generated at a much inferior resolution, typically 10-100 µm. As a result, integrative interpretation across the resolution domains can be challenging and, therefore, an approximation of the IMS information at the microscopy resolution may benefit data interpretation. To achieve this, we used OPLS to first model the relationships between

variables of the microscopy image data (X-block) and the IMS image (Y-block) at the IMS resolution, to then predict a high resolution image using that model (Figure 6a), a strategy similar to that for image fusion previously presented.[10] Thereby, a multivariate decomposition matrix of the microscopy image is generated in an effort to provide the feature diversity in distinct variables. The decomposition matrix consisted of color channels including color space transformations and their PCA scores. Heterogeneous distribution that is not contained in the microscopy modality was finally approximated by post-prediction weighting. The approach for high resolution approximation was demonstrated on an OnPLS scores image and on single ion images of the prostate cancer dataset (Figure 6b-c). While the examples in this case were purposely selected for their correlations between microscopy and IMS images, not all cross-modality relationships can be described accurately by a predictive model. Therefore, model evaluation by, for example, cross-validation to assess prediction reliability is important. Since, variation that is not represented within the microscopy modality can theoretically not be accurately predicted, we weighted the high resolution prediction with the original IMS image to approximate the analyte's distribution that is not described by the microscopy modality.

Moreover, variables with a strong contribution to OnPLS results might be of particular interest to predict at a high resolution. However, only variables within a certain correlation threshold between the modalities are suitable for predictive modeling. Therefore, variables should be selected with both, a high loading value in a particular component and strong correlation between IMS and microscopy modalities. An OnPLS model of IMS data and H&E data in two blocks provides the covariations between the modalities. These values plotted against the loadings of a component of interest will guide the selection of loadings suitable for prediction at the extremes of the scatter plot (Supplementary Figure 4).

Image fusion by OPLS prediction and distribution weighting projects the IMS image information onto a higher resolution microscopy image and presents yet another chemometrics strategy in the toolbox of spatial chemometrics.
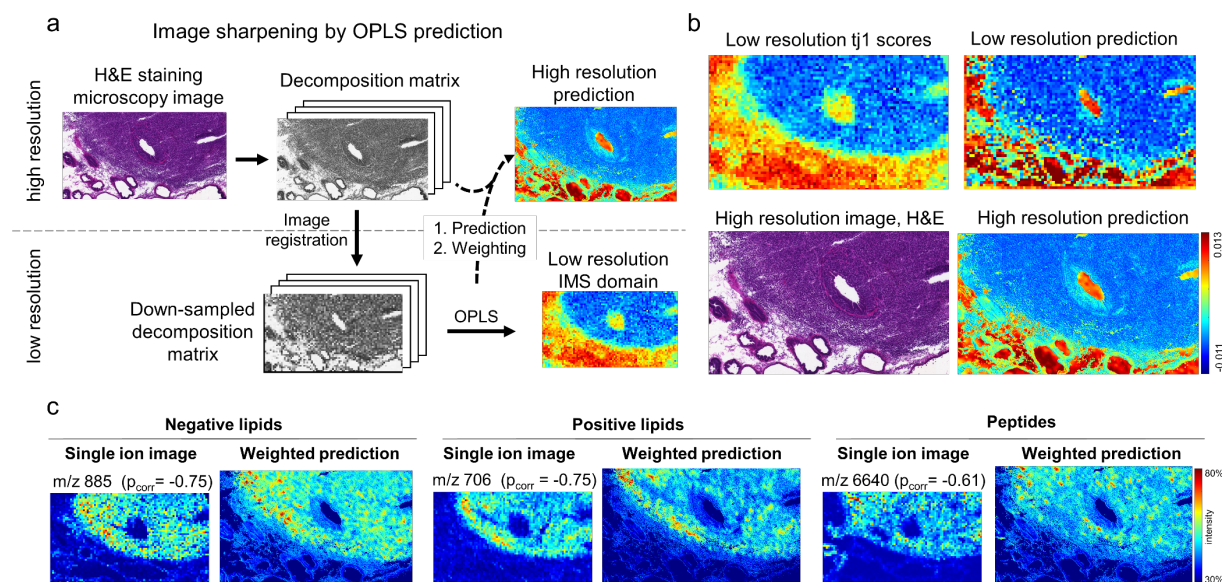

Figure 6. Multimodal IMS image fusion strategy by OPLS predictive modeling. (a) A low resolution IMS image and a corresponding high resolution microscopy image are used to predict the IMS image information at the microscopy resolution, (b) OnPLS tj1 scores image of prostate cancer IMS data (30 µm spatial resolution), its low resolution and high resolution predictions, and corresponding H&E stained microscopy image (1.25 µm spatial resolution), (c) high resolution predictions for high OnPLS loading ions for each modality; correlated pj1 loading value indicated. Signals correspond to PI(38:4) (m/z 885), PC(30:0) (m/z 706), and apolipoprotein C1 (m/z 6640).

Conclusions

In conclusion, we demonstrated that multimodal molecular imaging in conjunction with advanced multivariate statistical modeling reveals remarkable insight into disease pathology. The combination of multiple IMS modalities was enabled through precise image data alignment accomplished by an intensity-based semi-automated image registration procedure. Cross-modality chemometrics modeling of the combined data, using multiblock orthogonal component analysis, allowed the visualization of covariance structures and unique variations at the single pixel resolution that otherwise could not be achieved. Hence, the presented strategy lays the

groundwork for correlative molecular imaging to further understand the interplay of the underlying biochemistry with a multimodality aspect. Spatial chemometrics creates avenues to new biological understanding through integration of molecular information from multiple imaging modalities, and contributes to unlocking the full potential of multimodal imaging studies.

Acknowledgments

Conflict of interest disclosure: HZ has served at scientific advisory boards for Roche Diagnostics, Samumed, CogRx and Wave, has given lectures in symposia sponsored by Biogen and Alzecure, and is a co-founder of Brain Biomarker Solutions in Gothenburg AB, a GU Ventures-based platform company at the University of Gothenburg (all unrelated to the submitted work). KB has served as a consultant or at advisory boards for Abcam, Axon, Biogen, Lilly, MagQu, Novartis and Roche Diagnostics, and is a co-founder of Brain Biomarker Solutions in

Gothenburg AB, a GU Ventures-based platform company at the University of Gothenburg, all

unrelated to the work presented in this paper. The other authors declare no conflict of interest.

References

(1)   McDonnell, L. A.; Heeren, R. M. A. Imaging mass spectrometry. *Mass Spectrometry Reviews* **2007,** *26* (4), 606-643. 10.1002/mas.20124

(2)   Michno, W.; Wehrli, P. M.; Blennow, K.; Zetterberg, H.; Hanrieder, J. Molecular imaging mass spectrometry for probing protein dynamics in neurodegenerative disease pathology. *Journal of Neurochemistry* **2019,** *151* (4), 488-506. 10.1111/jnc.14559

(3)   Kaya, I.; Brinet, D.; Michno, W.; Başkurt, M.; Zetterberg, H.; Blenow, K.; Hanrieder, J. Novel Trimodal MALDI Imaging Mass Spectrometry (IMS3) at 10 μm Reveals Spatial Lipid and Peptide Correlates Implicated in Aβ Plaque Pathology in Alzheimer's Disease. *ACS Chemical Neuroscience* **2017,** *8* (12), 2778-2790. 10.1021/acschemneuro.7b00314

(4)   Kaya, I.; Michno, W.; Brinet, D.; Iacone, Y.; Zanni, G.; Blennow, K.; Zetterberg, H.; Hanrieder, J. Histology-Compatible MALDI Mass Spectrometry Based Imaging of Neuronal Lipids for Subsequent Immunofluorescent Staining. *Analytical Chemistry* **2017,** *89* (8), 4685-4694. 10.1021/acs.analchem.7b00313

(5)   Michno, W.; Kaya, I.; Nyström, S.; Guerard, L.; Nilsson, K. P. R.; Hammarström, P.; Blennow, K.; Zetterberg, H.; Hanrieder, J. Multimodal Chemical Imaging of Amyloid Plaque Polymorphism Reveals Aβ Aggregation Dependent Anionic Lipid Accumulations and Metabolism. *Analytical Chemistry* **2018,** *90* (13), 8130-8138. 10.1021/acs.analchem.8b01361

(6)   Tian, X.; Xie, B.; Zou, Z.; Jiao, Y.; Lin, L.-E.; Chen, C.-L.; Hsu, C.-C.; Peng, J.; Yang, Z. Multimodal Imaging of Amyloid Plaques: Fusion of the Single-Probe Mass Spectrometry Image and Fluorescence Microscopy Image. *Analytical Chemistry* **2019,** *91* (20), 12882-12889. 10.1021/acs.analchem.9b02792

(7)   Goubran, M.; Leuze, C.; Hsueh, B.; Aswendt, M.; Ye, L.; Tian, Q.; Cheng, M. Y.; Crow, A.; Steinberg, G. K.; McNab, J. A.et al. Multimodal image registration and connectivity analysis for integration of connectomic data from microscopy to MRI. *Nature Communications* **2019,** *10* (1), 5504. 10.1038/s41467-019-13374-0

(8)   Patterson, N. H.; Yang, E.; Kranjec, E.-A.; Chaurand, P. Co-registration and analysis of multiple imaging mass spectrometry datasets targeting different analytes. *Bioinformatics* **2018,** *35* (7), 1261-1262. 10.1093/bioinformatics/bty780

(9)   McDonnell, L. A.; van Remoortere, A.; van Zeijl, R. J. M.; Deelder, A. M. Mass Spectrometry Image Correlation: Quantifying Colocalization. *Journal of Proteome Research* **2008,** *7* (8), 3619-3627. 10.1021/pr800214d

(10) Van de Plas, R.; Yang, J.; Spraggins, J.; Caprioli, R. M. Fusion of mass spectrometry and microscopy: a multi-modality paradigm for molecular tissue mapping. *Nature methods* **2015,** *12* (4), 366-372. 10.1038/nmeth.3296

(11) Tortorella, S.; Tiberi, P.; Bowman, A. P.; Claes, B. S. R.; Ščupáková, K.; Heeren, R. M. A.; Ellis, S. R.; Cruciani, G. LipostarMSI: Comprehensive, Vendor-Neutral Software for Visualization, Data Analysis, and Automated Molecular Identification in Mass Spectrometry Imaging. *Journal of the American Society for Mass Spectrometry* **2019**, DOI:10.1021/jasms.9b00034 10.1021/jasms.9b00034. 10.1021/jasms.9b00034

(12) Alexandrov, T. MALDI imaging mass spectrometry: statistical data analysis and current computational challenges. *BMC Bioinformatics* **2012,** *13* (16), S11. 10.1186/1471-2105-13-s16-s11

(13) Cassat, J. E.; Moore, J. L.; Wilson, K. J.; Stark, Z.; Prentice, B. M.; Van de Plas, R.; Perry, W. J.; Zhang, Y.; Virostko, J.; Colvin, D. C.et al. Integrated molecular imaging reveals tissue heterogeneity driving host-pathogen interactions. *Science Translational Medicine* **2018,** *10* (432), eaan6361. 10.1126/scitranslmed.aan6361

(14) Masyuko, R.; Lanni, E.; Sweedler, J. V.; Bohn, P. W. Correlated Imaging – A Grand Challenge in Chemical Analysis. *The Analyst* **2013,** *138* (7), 1924-1939. 10.1039/c3an36416j

(15) Strömvall, K.; Thysell, E.; Halin Bergström, S.; Bergh, A. Aggressive rat prostate tumors reprogram the benign parts of the prostate and regional lymph nodes prior to metastasis. *PLoS One* **2017,** *12* (5), e0176679-e0176679. 10.1371/journal.pone.0176679

(16) Michno, W.; Wehrli, P. M.; Zetterberg, H.; Blennow, K.; Hanrieder, J. GM1 locates to mature amyloid structures implicating a prominent role for glycolipid-protein interactions in Alzheimer pathology. *Biochimica et biophysica acta. Proteins and proteomics* **2018**, DOI:10.1016/j.bbapap.2018.09.010 10.1016/j.bbapap.2018.09.010. 10.1016/j.bbapap.2018.09.010

(17) Michno, W.; Nyström, S.; Wehrli, P.; Lashley, T.; Brinkmalm, G.; Guerard, L.; Syvänen, S.; Sehlin, D.; Kaya, I.; Brinet, D.et al. Pyroglutamation of amyloid-βx-42 (Aβx-42) followed by Aβ1-40 deposition underlies plaque polymorphism in progressing Alzheimer's disease pathology. *Journal of Biological Chemistry* **2019**, DOI:10.1074/jbc.RA118.006604 10.1074/jbc.RA118.006604. 10.1074/jbc.RA118.006604

(18) Race, A. M.; Styles, I. B.; Bunch, J. Inclusive sharing of mass spectrometry imaging data requires a converter for all. *Journal of Proteomics* **2012,** *75* (16), 5111-5112. 10.1016/j.jprot.2012.05.035

(19) Löfstedt, T.; Trygg, J. OnPLS—a novel multiblock method for the modelling of predictive and orthogonal variation. *Journal of Chemometrics* **2011,** *25* (8), 441-455. 10.1002/cem.1388

(20) Löfstedt, T.; Hoffman, D.; Trygg, J. Global, local and unique decompositions in OnPLS for multiblock data analysis. *Analytica Chimica Acta* **2013,** *791*, 13-24. 10.1016/j.aca.2013.06.026

(21) Trygg, J.; Wold, S. Orthogonal projections to latent structures (O-PLS). *Journal of Chemometrics* **2002,** *16* (3), 119-128. doi:10.1002/cem.695

(22) Wehrli, P. M.; Lindberg, E.; Angerer, T. B.; Wold, A. E.; Gottfries, J.; Fletcher, J. S. Maximising the potential for bacterial phenotyping using time-of-flight secondary ion mass spectrometry with multivariate analysis and Tandem Mass Spectrometry. *Surface and Interface Analysis* **2014,** *46* (S1), 173-176. 10.1002/sia.5505

(23) Veselkov, K. A.; Mirnezami, R.; Strittmatter, N.; Goldin, R. D.; Kinross, J.; Speller, A. V. M.; Abramov, T.; Jones, E. A.; Darzi, A.; Holmes, E.et al. Chemo-informatic strategy for imaging mass spectrometry-based hyperspectral profiling of lipid signatures in colorectal cancer. *Proceedings of the National Academy of Sciences* **2014,** *111* (3), 1216-1221. 10.1073/pnas.1310524111

(24) Rohlfing, T. Image similarity and tissue overlaps as surrogates for image registration accuracy: widely used but unreliable. *IEEE transactions on medical imaging* **2012,** *31* (2), 153-163. 10.1109/TMI.2011.2163944

(25) Trygg, J. O2-PLS for qualitative and quantitative analysis in multivariate calibration. *Journal of Chemometrics* **2002,** *16* (6), 283-293. 10.1002/cem.724

(26) Trygg, J.; Wold, S. O2-PLS, a two-block (X–Y) latent variable regression (LVR) method with an integral OSC filter. *Journal of Chemometrics* **2003,** *17* (1), 53-64. 10.1002/cem.775

(27) Galindo-Prieto, B.; Trygg, J.; Geladi, P. A new approach for variable influence on projection (VIP) in O2PLS models. *Chemometrics and Intelligent Laboratory Systems* **2017,** *160*, 110-124. 10.1016/j.chemolab.2016.11.005

(28) Gabrielsson, J.; Jonsson, H.; Airiau, C.; Schmidt, B.; Escott, R.; Trygg, J. OPLS methodology for analysis of pre-processing effects on spectroscopic data. *Chemometrics and Intelligent Laboratory Systems* **2006,** *84* (1), 153-158. 10.1016/j.chemolab.2006.03.013

(29) Gabrielsson, J.; Jonsson, H.; Airiau, C.; Schmidt, B.; Escott, R.; Trygg, J. The OPLS methodology for analysis of multi-block batch process data. *Journal of Chemometrics* **2006,** *20* (8-10), 362-369. 10.1002/cem.1009

(30) Bylesjö, M.; Nilsson, R.; Srivastava, V.; Grönlund, A.; Johansson, A. I.; Jansson, S.; Karlsson, J.; Moritz, T.; Wingsle, G.; Trygg, J. Integrated Analysis of Transcript, Protein and Metabolite Data To Study Lignin Biosynthesis in Hybrid Aspen. *Journal of Proteome Research* **2009,** *8* (1), 199-210. 10.1021/pr800298s

(31) Kirwan, G. M.; Johansson, E.; Kleemann, R.; Verheij, E. R.; Wheelock, Å. M.; Goto, S.; Trygg, J.; Wheelock, C. E. Building Multivariate Systems Biology Models. *Analytical Chemistry* **2012,** *84* (16), 7064-7071. 10.1021/ac301269r

(32) Rantalainen, M.; Cloarec, O.; Beckonert, O.; Wilson, I. D.; Jackson, D.; Tonge, R.; Rowlinson, R.; Rayner, S.; Nickson, J.; Wilkinson, R. W.et al. Statistically Integrated Metabonomic−Proteomic Studies on a Human Prostate Cancer Xenograft Model in Mice. *Journal of Proteome Research* **2006,** *5* (10), 2642-2655. 10.1021/pr060124w

(33) Bouhaddani, S. e.; Houwing-Duistermaat, J.; Salo, P.; Perola, M.; Jongbloed, G.; Uh, H.-W. Evaluation of O2PLS in Omics data integration. *BMC Bioinformatics* **2016,** *17* (2), S11. 10.1186/s12859-015-0854-z

(34) Berry, K. A. Z.; Hankin, J. A.; Barkley, R. M.; Spraggins, J. M.; Caprioli, R. M.; Murphy, R. C. MALDI imaging of lipid biochemistry in tissues by mass spectrometry. *Chem Rev* **2011,** *111* (10), 6491-6512. 10.1021/cr200280p

(35) Wang, H.-Y. J.; Jackson, S. N.; Post, J.; Woods, A. S. A Minimalist Approach to MALDI Imaging of Glycerophospholipids and Sphingolipids in Rat Brain Sections. *Int J Mass Spectrom* **2008,** *278* (2-3), 143-149. 10.1016/j.ijms.2008.04.005

(36) Hong, J. H.; Kang, J. W.; Kim, D. K.; Baik, S. H.; Kim, K. H.; Shanta, S. R.; Jung, J. H.; Mook-Jung, I.; Kim, K. P. Global changes of phospholipids identified by MALDI imaging mass spectrometry in a mouse model of Alzheimer's disease. *Journal of lipid research* **2016,** *57* (1), 36-45. 10.1194/jlr.M057869

(37) Deininger, S.-O.; Ebert, M. P.; Fütterer, A.; Gerhard, M.; Röcken, C. MALDI Imaging Combined with Hierarchical Clustering as a New Tool for the Interpretation of Complex Human Cancers. *Journal of Proteome Research* **2008,** *7* (12), 5230-5236. 10.1021/pr8005777

(38) Wehrli, P. M.; Michno, W.; Blennow, K.; Zetterberg, H.; Hanrieder, J. Chemometric Strategies for Sensitive Annotation and Validation of Anatomical Regions of Interest in Complex Imaging Mass Spectrometry Data. *Journal of The American Society for Mass Spectrometry* **2019**, DOI:10.1007/s13361-019-02327-y 10.1007/s13361-019-02327-y. 10.1007/s13361-019-02327-y

(39) Lauer, F.; Diehn, S.; Seifert, S.; Kneipp, J.; Sauerland, V.; Barahona, C.; Weidner, S. Multivariate Analysis of MALDI Imaging Mass Spectrometry Data of Mixtures of Single Pollen Grains. *Journal of The American Society for Mass Spectrometry* **2018,** *29* (11), 2237-2247. 10.1007/s13361-018-2036-5

(40) Weaver, E. M.; Hummon, A. B.; Keithley, R. B. Chemometric analysis of MALDI mass spectrometric images of three-dimensional cell culture systems. *Analytical Methods* **2015,** *7* (17), 7208-7219. 10.1039/C5AY00293A

(41)   Reinke, S. N.; Galindo-Prieto, B.; Skotare, T.; Broadhurst, D. I.; Singhania, A.; Horowitz, D.; Djukanović, R.; Hinks, T. S. C.; Geladi, P.; Trygg, J.et al. OnPLS-Based Multi-Block Data Integration: A Multivariate Approach to Interrogating Biological Interactions in Asthma. *Analytical chemistry* **2018,** *90* (22), 13400-13408. 10.1021/acs.analchem.8b03205

(42)   McDonnell, L. A.; Corthals, G. L.; Willems, S. M.; van Remoortere, A.; van Zeijl, R. J. M.; Deelder, A. M. Peptide and protein imaging mass spectrometry in cancer research. *Journal of Proteomics* **2010,** *73* (10), 1921-1944. 10.1016/j.jprot.2010.05.007

(43)   Randall, E. C.; Emdal, K. B.; Laramy, J. K.; Kim, M.; Roos, A.; Calligaris, D.; Regan, M. S.; Gupta, S. K.; Mladek, A. C.; Carlson, B. L.et al. Integrated mapping of pharmacokinetics and pharmacodynamics in a patient-derived xenograft model of glioblastoma. *Nature Communications* **2018,** *9* (1), 4904. 10.1038/s41467-018-07334-3