Discovering new lipidomic features using cell type specific fluorophore expression to provide spatial and biological specificity in a multimodal workflow with MALDI IMS

Marissa A. Jones†‡, Sung Hoon Cho§, Nathan Heath Patterson‡‖, Raf Van de Plas‡‖⊥, Jeffrey M. Spraggins*‡‖, Mark R. Boothby*§¶∇#□, and Richard M. Caprioli†‡‖¶□

†Department of Chemistry, Vanderbilt University, 7330 Stevenson Center, Station B 351822, Nashville, Tennessee 37235

‡ Mass Spectrometry Research Center and Department of Biochemistry, Vanderbilt University Medical Center, 465 21st Avenue South, MRB III Suite 9160, Nashville, Tennessee 37232.

§ Department of Pathology, Microbiology, and Immunology, 1161 21st Avenue South, MCN AA-4214B, MCN A-5301, Vanderbilt University Medical Center, Nashville, TN 37232.

‖Department of Biochemistry, Vanderbilt University, Nashville, TN 37232.

⊥Delft Center for Systems and Control (DCSC), Delft University of Technology, 2628 CD Delft, The Netherland

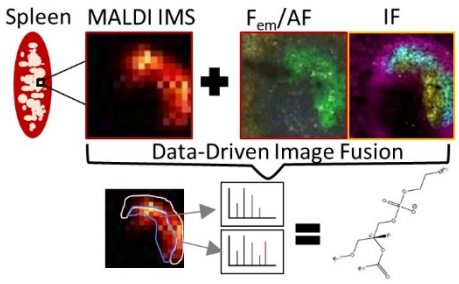
¶Department of Medicine, Vanderbilt University, Nashville, TN

∇ Department of Cancer Biology, Vanderbilt University, Nashville, TN

#Vanderbilt-Ingram Cancer Center, Vanderbilt University, Nashville, TN

□ Pharmacology, Vanderbilt University, Nashville, TN

ABSTRACT: Identifying the spatial distributions of biomolecules in tissue is crucial for understanding integrated function. Imaging Mass Spectrometry (IMS) allows simultaneous mapping of thousands of biosynthetic products such as lipids but has needed a means of identifying specific cell-types or functional states to correlate with molecular localization. We report here advances starting from identity marking with a genetically encoded fluorophore. The fluorescence emission data were integrated with IMS data through multimodal image processing with advanced registration techniques and data-driven image fusion. In an unbiased analysis of spleens, this integrated technology enabled identification of ether lipid species preferentially enriched in germinal centers. We propose that this use of genetic marking for microanatomical regions of interest can be paired with molecular information from IMS for any tissue, cell-type, or activity state for which fluorescence is driven by a gene-tracking allele and ultimately with outputs of other means of spatial mapping.



Simultaneously mapping the spatial localizations of biomolecules enables the formulation of new hypotheses and can test models related to physiology, disease pathogenesis and clinical applications. Although a variety of technologies exist for spatial localization of metabolites, these technologies face barriers in providing full biological context to findings because biosynthesis and steady-state levels of molecular determinants of cell metabolism and function may be regulated post-translationally. Thus, complementary imaging modalities are required for correlation of molecular images with biologically relevant substructures. Matrix-assisted laser desorption/ionization (MALDI) imaging mass spectrometry (IMS) enables the mapping of thousands of unlabeled molecules, including lipids and other metabolic products, directly from tissue sections at high spatial resolution[1]. The challenge of correlating ion localization to unambiguous identification of microanatomical regions of interest (ROIs) is a computational and experimental challenge.

Microscopy images collected from stained tissue (e.g. staining of tissues by hematoxylin and eosin (H&E) or Immunofluorescence (IF)) [2–4] are generally used to provide biological context to IMS data. However, the use of serial sections, the standard method of providing this biological context, limits the discriminant power of scoring cell identity or functional status (e.g., activity of a particular gene) for small regions of interest. Moreover, differences in spatial resolution can make correlating IMS and microscopy images challenging. Routine spatial resolution of most IMS experiments is 10-30 µm but can attain 5 µm resolution using specialized instruments[5,6]. These considerations highlight the need for a multimodal workflow in which biological features can be identified at a microanatomic scale in IMS analyses[7]. The spatial colocalization of a transgenic fluorophore with IMS data provides enhanced biological specificity and advanced data-mining strategies to uncover molecular correlations with ROIs.

Every multimodal analysis has three central processes: registration (alignment of images in 2-D space[8]), data mining (parsing through data for relevant m/z values[9]), and molecular identification (elucidation through MS/MS[10]). Traditionally,

multimodal imaging has relied on manual interpretation of co-registered ion images[11], which is prone to human bias. Other supervised and unsupervised approaches have been used to improve data analysis[12–18]. Each of these approaches still requires an independent benchmark to define cells or structures. Herein we provide evidence of a new approach that enabled the identification of ROIs on the same tissue section using a cell-type specific transgenic fluorophore to provide a biologically specificity and the basis for fluorophore-directed data mining.

To develop this technology, we analyzed the spleens of unimmunized and immunized mice using a well-characterized tracking allele that encodes green fluorescent protein (GFP) to enable high accuracy image registration and provide biological context[19]. Data mining strategies such as manual interpretation[9,11], standard segmentation[20], and data-driven image fusion[21] were subsequently applied to determine whether lipids could be mapped to a feature of normal microanatomy in immune responses. The analyses show that data-driven image fusion allowed for the most robust mining of multimodal data by leveraging the correlation of F_{em} and IMS to identify previously unknown spatial molecular relationships.

MATERIALS AND METHODS

Materials.

MALDI matrix 1,5-diaminonapthalene (DAN) was purchased from Sigma-Aldrich Chemical Co. (St. Louis, MO, USA). Sheep red blood cells (SRBC), ammonium formate, carboxymethyl cellulose sodium salt, isopropyl alcohol, mass spectrometry grade water, chloroform, and acetonitrile were purchased from Fisher Scientific (Pittsburg, PA, USA); streptavidin-Alexa647 antibody (Ab) and chemically conjugated monoclonal Ab (GL7-FITC, αIgD-PE and αCD35-biotin) were purchased from BD Biosciences (San Jose, CA). C57BL/6-J mice and breeding stock transgenic for a bacterial artificial chromosome that integrates a translational fusion of GFP with AID into the Aicda locus (AID-GFP mice; stock# 018421) were obtained from Jackson Laboratory and bred with C57BL/6-J. All mice were housed in ventilated micro-isolators under Specified-Pathogen-Free conditions in a Vanderbilt mouse facility and used in accordance with protocols approved by the Institutional Animal Care & Use Committee.

Tissue Preparation.

AID-GFP (n=3) and C57BL/6-J (n=3) mice age six to seven weeks were immunized with sheep red blood cells to compare with non-immunized controls (C57BL6-J, n=3) and euthanized eight days post-immunization. Spleens were sectioned at 12 µm and three serial sections were used for H&E, IF, and IMS with F_{em} /AF on all sections prior to a secondary modality (Fig. 1).

Mass Spectrometry Imaging.

IMS sections were washed with ammonium formate and sprayed on a (TM Sprayer, HTX, Chapel Hill, NC, USA) with recrystallized 10 mg/ mL 1,5 DAN in 9:1 (v/v) acetonitrile/deionized water. Negative ion mode IMS data were acquired from m/z 200-2,000 with a raster step of 30 µm with a 9.4T Bruker FT-ICR Solarix mass spectrometer (Bruker Daltonics, Billerica, MA, USA) with a laser power optimized for each sample between 18%-20% with 500 laser shots per pixel. Smart walk of 25 µm was enabled to increase sensitivity. For image fusion analysis, a higher spatial resolution image was generated using the same 9.4T FT-ICR with similar settings except that the raster step was 15 µm without smart walk enabled, and 750 laser shots per pixel were generated at a laser power of 13%. All datasets are available at: https://figshare.com/s/ab2f73880453100e0c2c.

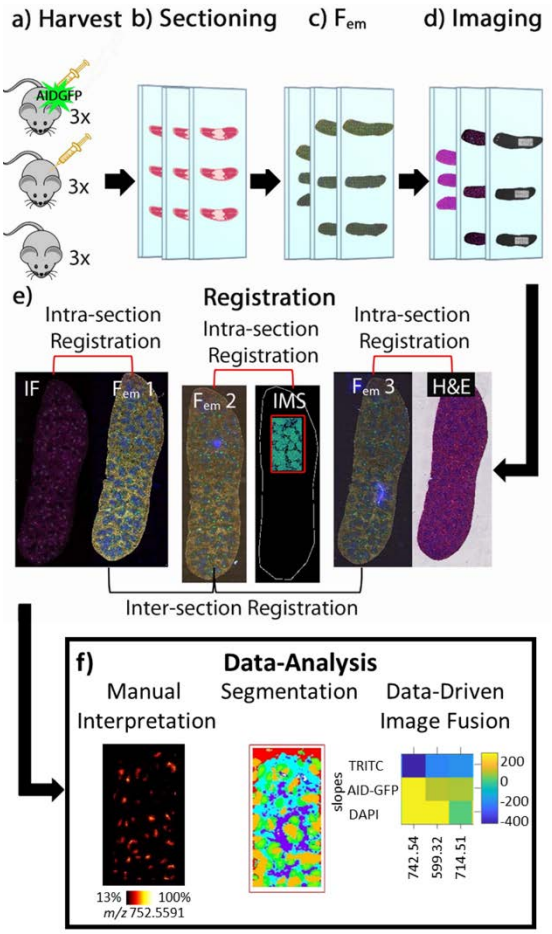


Figure 1. Workflow for multimodal analysis and data extraction. Shown are a schematic (a) and representative data (b-f) to illustrate the initial IMS analyses. a) Mice of the indicated genotypes (bearing or lacking an Aicda BAC transgene engineered to express AID-GFP translational fusion protein) and immunization status were used starting at 6-7 weeks of age. b-d) Spleens harvested 8 d post-immunization were used to generate triads of serial tissue sections (12 μ m thickness) (b), followed by fluorescence emission (F_{em}) and other imaging modalities (c). After processing, immunofluorescence (IF), IMS, and hematoxylin and eosin staining (H&E) were each performed with one of the three sections (d). (e) F_{em} data from sections 1-3, as indicated, are shown adjacent to the IF, one m/z from negative ion mode IMS, and H&E images from the same section as the F_{em} . Intra- and intersection registrations were then performed using a published method in which IMS data are aligned with the post IMS laser ablation marks, and all other modalities were aligned to IMS data through F_{em} on each section9 . f) Manual interpretation, segmentation, and data-driven image fusion were performed with publicly available software to map ions of interest, as detailed in the Methods.

Image Registration.

Image registration techniques were performed according to previously published methods[19] , however, rather than using AF images for registration we used F_{em} / AF images.

IMS Data Analysis.

All data was RMS normalized and further analyses were performed: manual interpretation analyses were performed in SCiLS, spatially shrunken centroid segmentation analysis was performed in R with the package Cardinal. Image fusion analyses were performed according to previously published methods[21], but utilizing the partial least squares regression correlation to compare image pairs of IMS and F_{em} data. Localization to germinal centers was determined using QuPath software for annotation and an R script for data extraction. Weighted averages were tested for significance with ratio T tests.

Identification of lipid Species

LC-MS/MS of total splenocytes was performed on a Q Exactive HF mass spectrometer from m/z 375-1650 in PRM mode with an isolation window of 2 Da for each ion of interest using eluates from a Vanquish UHPLC (Thermo Scientific, Waltham, MA, USA). MS/MS resolving power was 15,000 at m/z 200, while full scan resolving power was at m/z 200 was 30,000. Complementary analyses were performed using MS/MS based imaging experiments using a 15T Bruker FT-ICR solariX mass spectrometer (Bruker Daltonics, Billerica, MA, USA). Data was collected in negative ion mode from m/z 250-2,000 with 1,000 laser shots per pixel and a raster step was set to 60-120 µm. Ions were isolated with a 2-6 Da mass window and fragmented using collision induced dissociation (CID) with a collision energy of 17-27 V.

RESULTS AND DISSCUSSION

Overview. We used multiple microscopy-based imaging modalities (H&E, IF, and F_{em}) registered to IMS measurements to elucidate the lipidomic differences between GC and splenic white pulp or lymphoid follicles (Fig. 1). Specifically, the AID-GFP (Activation-Induced Deaminase-Green Fluorescent Protein) transgene provided a region-specific fluorophore[22]. This tracking allele highlights a micro-anatomical feature that forms within lymphoid follicles during the course of humoral immunity due to a large increase in Aicda gene expression in GC B lymphocytes, which diversifies and improves qualities of antibody responses [22]. F_{em} provided a non-destructive means of identifying GC via co-localization with AID-GFP, while AF from endogenous molecules provided histological images of other splenic tissue structures (e.g., red pulp and white pulp surrounding GC). This F_{em}/AF modality also provided a single image type that could be collected from every tissue section prior to other modalities (i.e. H&E stained microscopy, IF microscopy, and IMS) (Fig. 1 b-d), enabling high accuracy image registration (Fig. 1e). In H&E stains, the most traditional means of providing biological context to IMS data, red pulp can be differentiated from white pulp but GC are less conclusively differentiated, IF microscopy allowed for the identification of GC and their s (LZs) and dark zones (DZs) substructures. Because AID expression is similar in LZ and DZ, both H&E and IF after immunostaining were performed on serial sections. As this results in plane-of-section differences from sections used for IMS, advanced registration approaches were needed. By integrating these modalities into a single multi-planar dataset, we enabled a full integration of imaging modalities to provide a unique combination of molecular coverage, spatial resolution, and biological specificity.

Registration. We first tested whether this method allowed incorporation of F_{em} as an additional modality within each section to enable a high degree of spatially localized biological information. Sections were analyzed for GCs in spleens of mice, immunized or not, and bearing or lacking an AID-GFP transgenic fluorophore. The same tissue sections were then used for IMS, while serial sections were used for IF and H&E. This method was applied to an investigation of the differences between lipids associated with GCs and other regions in spleens using non-destructive F_{em} as a mono-modal registration medium[19] (Fig. 1e). Spleens of non-immunized controls were compared to those of immunized mice bearing or lacking the AID-GFP transgene (Fig. 1a). IMS was then used to identify m/z features after collection of F_{em} images of the AID-GFP fluorophore.

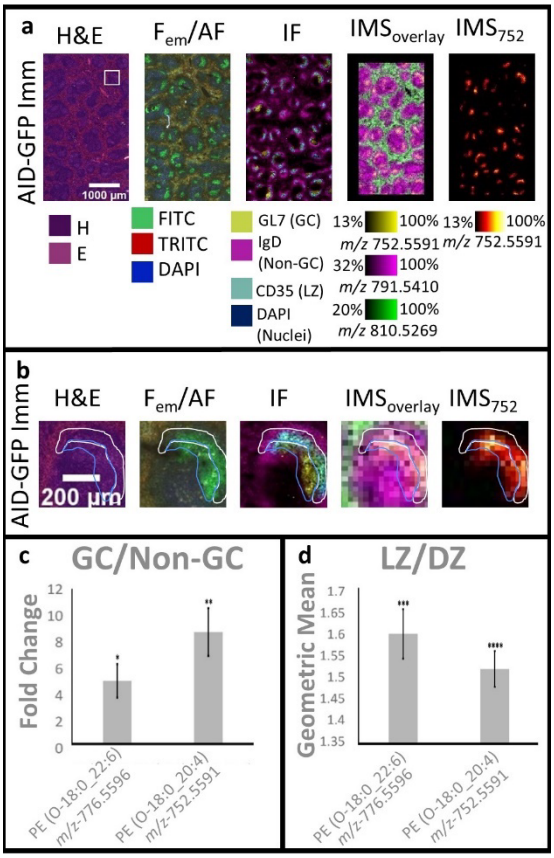


Figure 2. High accuracy registration of multimodal data. a) Representative registered images highlighting the types of detection. Rectangular areas of immunized AID-GFP transgenic (AID-GFP Imm) mouse spleen are shown with each section, from left to right: Hematoxylin and eosin (H&E); fluorescence emission/autofluroescence (F_{em}/AF); immunofluorescence (IF) after staining with mAb; IMS with three ions [*m/z* 752.5591, m*/z* 791.5410, and *m/z* 810.5269] overlaid for context of white pulp and red pulp; and a single ion image showing *m/z* 752.5591 (IMS_{752}). Intensity scales from least to greatest total ion intensity and color legends are displayed below each set of images. A 1000 µm scale bar is depicted in the H&E image. F_{em} was taken on the same section imaged by IMS. IF and H&E were then taken from serial sections to the IMS section. IF was used to identify micro-anatomic portions of lymphoid follicles, and included both indirect and direct staining of GL7, IgD, and CD35. b) Higher magnification images of a single representative GC (designated by a white box in 1a) are shown with the same sample order and modalities. GC LZ and DZ are demarcated by a yellow and blue outline respectively. c) The bar graph shows the ratio of ion intensities in GC to non-GC regions for the m/z features of 776.5596 and 752.5591 [identified by IMS MS/MS in Fig. 4 as PE (O-18:0 22:6) and PE (O-18:0 20:4), respectively] (p= 0.0409, p=0.0099, n=3). d) The geometric mean of the ratio of LZ/DZ ion intensity of two lipids is

1.6 and 1.5 for PE (O-18:0 20:4) and PE (O-18:0 22:6) (p=0.007, n=65, p=<0.0001, n=65). Replicates and magnified regions for WT Imm and WT Non-Imm samples can be found in the Fig. S2.

In addition to identification of GC within the section destined for IMS through F_{em}, we investigated lipid difference in subregions of the GC. Accordingly, the workflow incorporated IF staining of adjacent sections with antibodies specific for markers that not only would identify GC by independent criteria (IgDneg GL7+) but also would allow subdivision of the GC into functionally distinct domains termed the dark (DZ, CD35neg) and light (LZ, CD35+) zones. To compare the conventional use of serial sections to intra-section registration, we quantitated the error in overlap between adjacent sections. GC masks annotated for all AID-GFP mouse spleen serial section pairs (n =5) were used to calculate a Dice-Sorenson coefficient (DSC), a statistical means of determining the similarity of two samples that were registered as described by Patterson et al[19]. The average DSC was 0.81 (±0.3) for the five pairs, indicating that serial sections as registered can be expected to have 81% GC overlap (Fig. S1, Table S1).

Table 1. GC Lipids revealed through all data mining strategies. From left to right the m/z value, identification of the lipid found the MS/MS imaging, matches to the LIPIDMAPS database, p value for a t-test between GC and non-GC regions, p value for a ratio paired t-test between LZ and DZ, ppm error in identification, manual interpretation discovery, segmentation discovery, or data-driven image fusion discovery.

m/z	Lipid ID	DB Matches	P value GC vs. non-GC	P value LZ vs. DZ	ppm error*	Man. Int.	Seg.	Image Fusion
671.4647	PA(18:1_16:1)	6	0.09	0.0007	0.070			222.1
699.4957	PA(18:1_18:1)	6	0.03	0.0002	0.36			277.2
699.4957	PA(18:0_18:2)	7	0.03	0.0002	0.36			277.2
699.4957	PA(20:2_16:0)	7	0.03	0.0002	0.36			277.2
714.5069	PE(18:2 16:0)	4	0.04	0.2	0.053			102.5
716.5224	PE(18:0_16:1)	3	0.1	0.9	0.059			243.8
725.5120	PA(20:3_18:0)	8	0.007	0.02	0.53			63.2
740.5246	PE(18:1_18:2)	4	0.01	0.01	2.9			112.8
742.5389	PE(18:0_18:2)	5	0.04	0.0006	1.0			290.5
746.5130	PE(P-16:0_22:6)	6	0.005	<0.0001	1.5			280.0
748.5273	PE(O-16:0_22:6)	6	0.007	0.2	0.37			236.6
752.5591	PE(O-18:0_20:4)	5	0.01	<0.0001	0.32	X	X	219.0
762.5088	PE(16:0_22:6)	4	0.03	0.2	2.6			167.0
772.5314	PE(P-18:1_22:6)	5	0.03	0.01	5.0			163.3
776.5596	PE(O-18:0_22:6)	5	0.05	<0.0001	0.88	X	X	244.7
786.5303	PS(18:0_18:2)	8	0.02	0.0004	2.9			279.4
812.5460	PS(18:0_20:3)	2	0.03	0.3	3.0			37.1
857.5182	PI(16:0_20:4)	16	0.009	0.002	0.82			400.6
883.5360	PI(18:1_20:4)	6	0.003	0.1	3.3		X	565.8
887.5609	PI(18:0_20:3)	14	0.0006	0.07	3.9		X	252.9

* Note that ppm error was determined from a tune mix doped IMS experiment.

Data Mining. Overall, 1,375 m/z features were detected at a S/N > 3 by IMS, including a variety of lipids with diverse patterns of localization to substructures of spleen that included red and white pulp. In addition to these constitutive features of splenic micro-anatomy, GC form in the white pulp after lymphocyte activation by immunization generates T cell help. Mice were immunized to increase size, differentiation, and numbers of GC which were observed in all imaging modalities when comparing immunized to non-immunized controls (Fig. 2a, Fig. S2). In F_{em} images, a difference in GC localized GFP expression can be seen between samples with and without AID-GFP (Fig. 2a, Fig. S2). AF detected in the DAPI and TRITC channels enhanced the identification GC in the FITC channel by distinguishing GC from other portions of the white pulp highlighted by the AF.

The IMS data were first analyzed using manual interpretation (Fig. 1f). Two ions of interest were selected by virtue of their association with in-section AID-GFP, m/z 752.5591 and m/z 776.5596 (Fig. 2a). A ratio paired T-test applied to the ion intensity was performed to determine significance of correlation and anti-correlation throughout this work. Specifically, GCs

were compared to non-GC regions. Because, AID-GFP does not distinguish the GC sub-regions[23], IF of adjacent sections was employed to identify the LZ and DZ.

Data were further analyzed for significant differences in GC LZs and DZs [24] as identified in F_{em} and IF microscopy images. To obtain ion intensity for statistical analysis, we used QuPath and a home-built R program to extract ion intensity values for all GC and non-GC regions identified through F_{em}. Sub-regions of GC, LZ and DZ identified through IF were annotated in QuPath[25], and compared. Pairs of GC LZs and DZs were identified based on shortest Euclidian distance (Fig. S3).

Ions discovered through manual interpretation, m/z 752.5591 and m/z 776.5596, were mapped to the GC (~8-fold and ~5-fold enrichment; Fig. 2 b, c), and each of these lipid species was further enriched in the LZ compared to the DZ (~1.6-fold and ~1.5-fold enrichment within the GC; Fig. 2b, d).

Spatially shrunken centroids segmentation which circumvents the potential for cognitive bias introduced through manual interpretation by computationally determining ROIs [20] (Fig. 1f). This approach generated a list of four ions that localize to GC, m/z 752.5591, 776.5596, 883.5360, and 887.5609 (m/z 883.5360 p=0.0025, n=3; m/z 887.5609 p=0.0087, n=3) [20],[26](Table S2). Of these, the first two (m/z 752.5591 and 776.5596) matched the ions discovered by manual interpretation, and all localized to GCs but not all localized to LZ or DZ (m/z 883.5360, p=0.12, n =108 and m/z 887.5609, p=0.070, n=106) (Fig. 2; Fig. S6 q-r; Table 1).

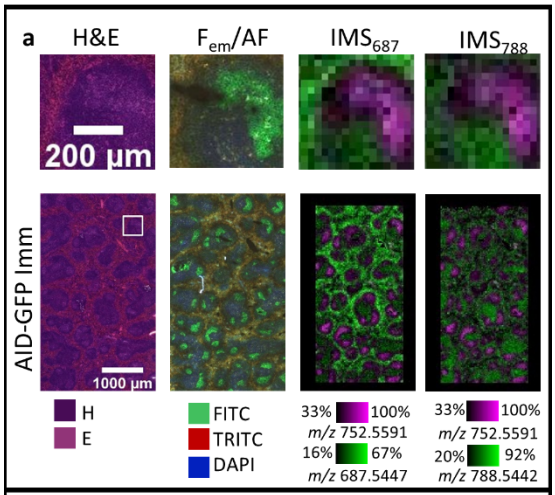
Although segmentation enabled the identification of four ions of interest localizing to GCs (Fig. 2a; Table 1), this approach is well suited only for determining ions that directly correlate to a specific tissue sub-region. Data-driven image fusion connects the spatial and informational content of two imaging modalities by constructing a cross-modality model using highly multivariate linear regression to enable predictive and data mining applications (Fig. S4, Fig. S5)[21]. In previous work, data-driven image fusion has been used for image enhancement such as spatial sharpening, out-of-sample prediction, and image denoising.[21]

We hypothesized that by fusing IMS images with those of F_{em}, the linear models produced through fusion processes could uncover new correlative relationships enabling fluorophore-directed data mining. Accordingly, we tested the use of data-driven image fusion to provide a deeper understanding of all correlative relationships between IMS and F_{em} data in GCs.

From the fusion of a high resolution (15 µm) IMS image and F_{em}, 16 GC-specific ions were revealed (Table 1), of which four were those highlighted by segmentation-based analyses (Fig. 2a, c). Integration of the image fusion algorithm into the workflow allowed identification of a far greater number of candidates for GC-associated ions along with species that were anti-correlated (e.g. m/z 687.5447 and m/z 788.5442, Fig.3; Fig. S5; Table S3).

GC areas annotated in F_{em} images served as a means for identifying GC (p=0.0099, n=3, slope= 219.0 for green channel) and non-GC regions for statistical analysis (p=0.04, n=3, slope= -57.5 for green channel and p=0.04, n=3, slope = -202.6 for green channel respectively) (Fig. 3; Table S3; Fig. S5). The ion m/z 752.5591 is shown for contrast with non-GC ions m/z 687.5447 and m/z 788.5442 (Fig. 3). In contrast to manual interpretation and segmentation, ten additional ions revealed through data-driven image fusion were higher in GCs and exhibited a pattern of LZ >DZ (Table 1; Fig. S4-5).

Molecular Identification. Due to the large number of potential isomers at these m/z values, mass accuracy alone is not enough to specifically identify lipids. For example, the phosphatidylethanolamine ether species PE(O-40:6) and PE(O-38:4) are isomers of the phosphatidylethanolamine plasmalogen species PE(P-40:5) and PE(P-38:3), respectively. Liquid chromatography-tandem mass spectrometry (LC-MS/MS) determined the presence of both ether and plasmalogen species for these ions of interest in total lipid extracts from whole spleen tissue (Fig. 4a). Thus, a spatial component was needed to confirm the identity of the ions that correlate to F_{em} signals. IMS-MS/MS was performed with sectioned spleens of immunized transgenic AID-GFP mice. The MS/MS-based imaging experiment found these ions to be ether lipids PE(O-18:0_20:4) (Fig. 4 b-d) and PE(O-18:0_22:6) and not the isomeric plasmalogens (Fig. 4e-f). In addition, the co-localization of the specific fragment ions from these ether lipids with F_{em} signals reveals that these species are enriched in splenic GC.



b

m/z	Lipid ID	DB Matches	P value non-GC vs GC	ppm error
687.5447	PE-Cer(d36:1)*	-	0.04	1.7
788.5442	PS(18:1_18:0)	7	0.04	0.72

Figure 3. Identification of anti-correlating germinal center (not-GC) ions by image fusion. a) Shown are representative registered images highlighting the localization of anti-correlating GC ions. From left to right, the following image types are pictured: H&E with scale bar, Fem , IF, IMS showing an overlay of non-GC ion m/z 687.5447 and GC ion m/z 752.5591, and IMS showing an overlay of non-GC ion m/z 788.5442. LZ and DZ as identified by IF are outlined in yellow and blue respectively. b) From left to right, the m/z value, identification, matches to the LIPIDMAPS database in the IMS MS/MS spectrum, statistical significance, and ppm error in mass identification are listed. These two ions were identified as PE-Cer (d36:1) by accurate mass and PS(18:1 18:0) through IMS MS/MS, respectively. A complete listing of results from image fusion is in Table 1; further information is in Extended Data (Fig. S4-5).

In addition to plasmalogen and ether species, image fusion enabled the identification of a variety of phosphatidylethanolamine (PE), phosphatidic acid (PA), glycerophosphoserine (PS), and glycerophosphoinositol (PI) lipids that were enriched in GC, with some observed at higher intensity in GC LZ. Fatty acid tails of 16:0 and 18:0 were most common. We observed many repeats of fatty acid tails 20:3, 20:4, and 22:6. In GC, five

out of eight lipids had unsaturated fatty acid tails, whereas in GC LZ, all eight had at least one unsaturated fatty acid tail. Two ions, m/z 687.5447, PE-Cer(d36:1) (phosphatidylethanolamine ceramide), and m/z 788.5442, PS(18:1_18:0), were identified as anti-correlating with GC (Fig. 3). This unique combination of IMS with biologically driven microscopy modalities, advanced image registration, multimodal data mining, and spatially driven identification provides a pipeline for elucidating molecular drivers of biological processes. As a test of the technology, this process revealed an enrichment of ether and plasmalogen lipid species in GC, a metabolically stressed environment central to the qualities of antibody responses and humoral memory.

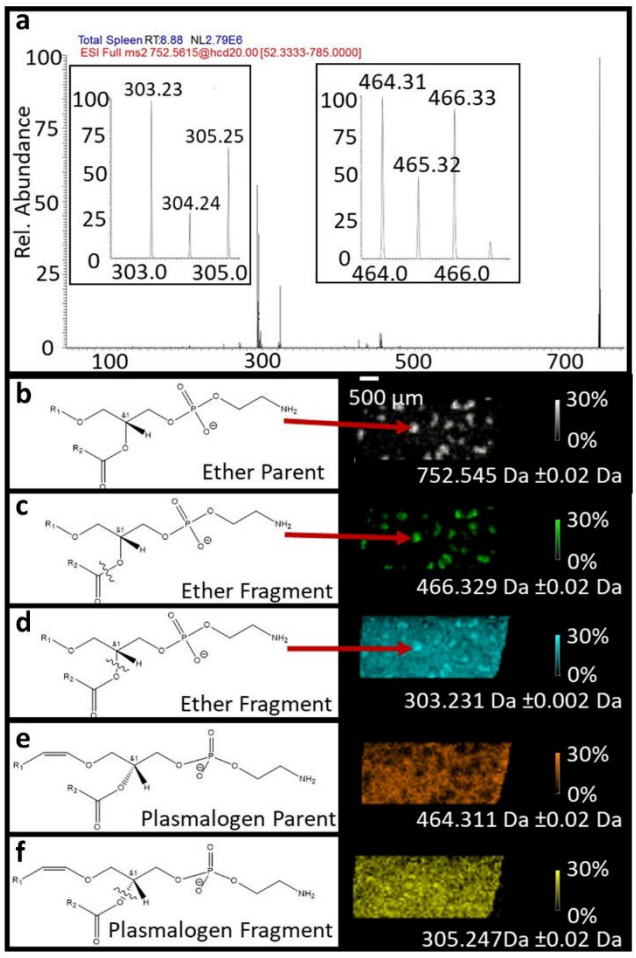


Figure 4. Identification of species localizing to germinal centers as ether linked lipids. a) LC-MS/MS fragmentation spectra of total splenocytes show common fragments for both plasmalogen and ether lipids (enlarged) from a parent mass of m/z 752.545. b) Shown to the left is the chemical structure of the parent ether ion and to the right the corresponding ion image. c,d) The correlating ether fragments are depicted with the chemical structure on the left and ion image on the right. e,f,g,) Similarly, plasmalogen parent ion structure and fragments are shown with chemical structure on the left and corresponding ion images on the right.

We have developed a multimodal imaging process that combines high spatial resolution IMS with microscopy utilizing a transgenic fluorophore to identify micro-anatomical regions of biological interest. Our approach incorporates high accuracy registration and various data mining tools, including data-driven image fusion, to fully integrate multiple imaging modalities collected from a single tissue section and across adjacent sections, enabling discovery of molecular drivers of immune response. Unambiguous identification of GC and the assessment of lipid abundances in light and dark zones was made possible by combining F_{em} of the transgenic tracking allele with traditional microscopy approaches (i.e. stained and IF microscopy). While data-driven image fusion has previously been used for spatial applications in image sharpening and out-of-sample prediction, the evidence presented here indicates that it can also be applied to mine highly dimensional data to find correlations between modalities by interpreting the linear models constructed during the fusion process. When compared to conventional approaches, the yield of structure-associated molecules was enhanced four- to five-fold, as 16 GC-associated lipid species were determined.

We identified three key processes in multimodal imaging as (1) registration, (2) data mining, and (3) molecular identification. Histological depth differences between serial sections are becoming larger challenges as the spatial resolution of IMS increases[27–29] due to the small size of single cells within a tissue. In addition to histological depth differences, accurate data alignment correlating H&E or IF to IMS becomes central as spatial resolution increases and regions of interest approach single cells. Importantly, the technologies presented here should be applicable to fusion of IMS, fluorescence, and spatial transcriptomic or protein data[30,31].

The unexpected finding that the prevalence of a series of ether lipid species is higher in GC frames new hypotheses, i.e., that molecular programing of GC lymphocytes is tied to increased ether lipid synthesis and that these species are functionally important in humoral immunity. A higher abundance of ether lipids in the spleen and white blood cells has been reported, but the exact role of these ether lipids remains uninvestigated[32]. Ether lipid synthesis begins in the peroxisome and is completed in the ER[33]. Disruption of this pathway in peroxisome biogenesis disorders such as Zellweger spectrum (PBD-ZSD) or by gene-targeting generates decreased ether lipid levels[33–35]. In this light, it was striking that image analysis of IMS uncovered GC PE lipids with the same tail lengths as their ether and plasmalogen counterparts. Most notably, PE(16:0_22:6) localized to GCs as did its ether lipid counterpart PE(O-16:0_22:6) while its plasmalogen derivative, PE(P-16:0_22:6), localized not only to GC but within them to their LZ (Table 1; Fig. S5o-q; Table S3). This enrichment along a pathway suggests that GC have enhanced peroxisomal activity, resulting in increased abundance of PE-ether lipids.

The peroxisome also generates reactive oxygen species (ROS)[36]. Plasmalogen ether lipids scavenge reactive oxygen species[37]. This capability has not been documented for non-plasmalogen ether lipids, but the structural similarity suggests a connection in synthesis pathways and roles[38]. Starting 3.5 d after immunization, GC form in the follicles of secondary lymphoid organs and are sites of B-cell proliferation, differentiation, and selection that are central to promoting antibody affinity increases as well as vaccine efficacy and humoral immunity[23]. Substantial AID- mediated mutational[39] and nutrient[23,40] stresses appear to be present in GCs. This micro-anatomic structure consists of LZ and DZs in which the native oxygen levels vary, such that hypoxia is present in an LZ>DZ pattern[23]. While there is strong evidence of connections between hypoxia and inflammation[41,42], much remains unknown as to the effect of this hypoxic microenvironment on lipid synthesis within these

regions[43]. The role of ether lipids in the adaptive immune microenvironment has not yet been explored, and thus, management of ROS and their levels are crucial for lymphocyte physiology[44]. This point, in conjunction with known metabolic stresses in GC[23,39,45] and influences of hypoxia on ROS generation[23], suggests that a model in which higher plasmalogen and ether lipid abundance in GC reflects a physiological role in which ether lipid production indicates the need to maintain optimal ROS levels[36].

CONCLUSION

Two key methodological advances documented here – use of engineered alleles that track gene expression by linking a fluorophore to the normal gene product, and application of data-driven image fusion for data mining – should be widely applicable to a variety of experiments in a broad range of biological systems. Gene-editing technologies such as CRISPR-Cas9 will further expand an already abundant supply of transgenes that mark specific biological pathways and cell-types. Moreover, this new application of image fusion as a means of elucidating ions of interest co-localizing with a specific fluorophore will enable unique applications of data mining, including applications in settings where unambiguous marking of a region of interest by other modalities exists.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website.

Expanded methods and supplementary figures (PDF)

AUTHOR INFORMATION

Corresponding Author

* E-mail: jeff.spraggins@vanderbilt.edu, mark.boothby@vumc.org

ORCID

Marissa A. Jones: 0000-0002-9346-727X
Sung Hoon Cho: 0000-0002-6463-4032
Nathan Heath Patterson: 0000-0002-0064-1583
Raf Van de Plas: 0000-0002-2232-7130
Jeffrey M. Spraggins: 0000-0001-9198-5498
Mark R. Boothby: 0000-0003-2593-8276
Richard M. Caprioli: 0000-0001-5859-3310

Author Contributions

M.A.J. designed and performed the research, analyzed data, and wrote the paper. S.H.C. developed and performed experiments with mice, and designed aspects of the research. N.H.P. designed and developed interactive multimodal data analysis programs and pipelines, guided analyses using the programs, and wrote portions of the paper pertaining to them. R.V.P. performed and analyzed data-driven image fusion. J.M.S. assisted in experimental design, data acquisition, data interpretation and visualization, and manuscript preparation. M.R.B. and R.M.C. designed experiments, interpreted data, and edited the manuscript.

Notes

The authors declare not competing financial interest.

ACKNOWLEDGMENT

This work was supported by NSF DGE-1445197 (R.M.C.), NIH grants P41 GM103391 (Richard M. Caprioli.), U45-DK120058 (Richard M. Caprioli) R01 AI113292 and R01 HL106812 (Mark R. Boothby). The Vanderbilt Mass Spectrometry Research Center and Core are gratefully acknowledged, particularly M. Wade Calcutt and Emilio Rivera for expert assistance with LCMS/MS method development. Whole slide imaging was performed in the Digital Histology Shared Resource at Vanderbilt University Medical Center (www.mc.vanderbilt.edu/dhsr).

REFERENCES

(1) Caprioli, R. M.; Farmer, T. B.; Gile, J. Molecular Imaging of Biological Samples: Localization of Peptides and Proteins Using MALDI-TOF MS. *Anal. Chem.* **1997**, *69* (23), 4751–4760.

(2) Cornett, D. S.; Mobley, J. A.; Dias, E. C.; Andersson, M.; Arteaga, C. L.; Sanders, M. E.; Caprioli, R. M. A Novel Histology-Directed Strategy for MALDI-MS Tissue Profiling That Improves Throughput and Cellular Specificity in Human Breast Cancer. *Mol. Cell. Proteomics* **2006**, *5* (10), 1975–1983. https://doi.org/10.1074/mcp.M600119-MCP200.

(3) Lotz, J. M.; Hoffmann, F.; Lotz, J.; Heldmann, S.; Trede, D.; Oetjen, J.; Becker, M.; Ernst, G.; Maas, P.; Alexandrov, T.; et al. Integration of 3D Multimodal Imaging Data of a Head and Neck Cancer and Advanced Feature Recognition. *Biochim. Biophys. Acta - Proteins Proteomics* **2017**, *1865* (7), 946–956. https://doi.org/10.1016/J.BBAPAP.2016.08.018.

(4) Feider, C. L.; Woody, S.; Ledet, S.; Zhang, J.; Sebastian, K.; Breen, M. T.; Eberlin, L. S. Molecular Imaging of Endometriosis Tissues Using Desorption Electrospray Ionization Mass Spectrometry. *Sci. Rep.* **2019**, *9* (1). https://doi.org/10.1038/s41598-019-51853-y.

(5) Zavalin, A.; Yang, J.; Haase, A.; Holle, A.; Caprioli, R. Implementation of a Gaussian Beam Laser and Aspheric Optics for High Spatial Resolution MALDI Imaging MS. *J. Am. Soc. Mass Spectrom.* **2014**, *25* (6), 1079–1082. https://doi.org/10.1007/s13361-014-0872-5.

(6) Norris, J. L.; Caprioli, R. M. Analysis of Tissue Specimens by Matrix-Assisted Laser Desorption/Ionization Imaging Mass Spectrometry in Biological and Clinical Research. **2013**. https://doi.org/10.1021/cr3004295.

(7) Heeren, R. M. A.; Bruinen, A. L.; Mascini, N. E.; Fisher, G. L.; Porta, T.; Ellis, S. R. Multimodal Molecular Imaging: Insight into the Complexity of Biological Surfaces through Speed, Resolution and Identification. *Microsc. Microanal* **2015**, *21*, 1927. https://doi.org/10.1017/S1431927615011952.

(8) Oliveira, F. P. M.; Manuel, J.; Tavares, R. S. *Medical Image Registration: A Review*.

(9) Alexandrov, T. MALDI Imaging Mass Spectrometry: Statistical Data Analysis and Current Computational Challenges. *BMC Bioinformatics* **2012**. https://doi.org/10.1186/1471-2105-13-S16-S11.

(10) Norris, J. L.; Caprioli, R. M. Analysis of Tissue Specimens by Matrix-Assisted Laser Desorption/ Ionization Imaging Mass Spectrometry in Biological and Clinical Research. https://doi.org/10.1021/cr3004295.

(11) Whitehead, S. N.; Chan, K. H. N.; Gangaraju, S.; Slinn, J.; Li, J.; Hou, S. T. Imaging Mass Spectrometry Detection of Gangliosides Species in the Mouse Brain Following Transient Focal Cerebral Ischemia and Long-Term Recovery. *PLoS One* **2011**, *6* (6), e20808. https://doi.org/10.1371/journal.pone.0020808.

(12) Amstalden Van Hove, E. R.; Blackwell, T. R.; Klinkert, I.; Eijkel, G. B.; Heeren, R. M. A.; Glunde, K. Integrated Systems and Technologies Multimodal Mass Spectrometric Imaging of Small Molecules Reveals Distinct Spatio-Molecular Signatures in Differentially Metastatic Breast Tumor Models. **2010**. https://doi.org/10.1158/0008-5472.CAN-10-0360.

(13) Abdelmoula, W. M.; Regan, M. S.; Lopez, B. G. C.; Randall, E. C.; Lawler, S.; Mladek, A. C.; Nowicki, M. O.; Marin, B. M.; Agar, J. N.; Swanson, K. R.; et al. Automatic 3D Nonlinear Registration of Mass Spectrometry Imaging and Magnetic Resonance Imaging Data. *Anal. Chem.* **2019**, *91* (9), 6206–6216. https://doi.org/10.1021/acs.analchem.9b00854.

(14) Hanselmann, M.; Kirchner, M.; Renard, B. Y.; Amstalden, E. R.; Glunde, K.; Heeren, R. M. A.; Hamprecht, F. A. Concise Representation of Mass Spectrometry Images by Probabilistic Latent Semantic Analysis. *Anal. Chem.* **2008**, *80* (24), 9649–

9658. https://doi.org/10.1021/ac801303x.
(15) Chughtai, K.; Jiang, L.; Greenwood, T. R.; Glunde, K.; Heeren, R. M. A. Mass Spectrometry Images Acylcarnitines, Phosphatidylcholines, and Sphingomyelin in MDA-MB-231 Breast Tumor Models. *J. Lipid Res.* **2013**, *54* (2), 333–344. https://doi.org/10.1194/jlr.M027961.
(16) Oetjen, J.; Aichler, M.; Trede, D.; Strehlow, J.; Berger, J.; Heldmann, S.; Becker, M.; Gottschalk, M.; Kobarg, J. H.; Wirtz, S.; et al. MRI-Compatible Pipeline for Three-Dimensional MALDI Imaging Mass Spectrometry Using PAXgene Fixation. *J. Proteomics* **2013**, *90*, 52–60. https://doi.org/10.1016/J.JPROT.2013.03.013.
(17) Verbeeck, N.; Yang, J.; De Moor, B.; Caprioli, R. M.; Waelkens, E.; Van de Plas, R. Automated Anatomical Interpretation of Ion Distributions in Tissue: Linking Imaging Mass Spectrometry to Curated Atlases. *Anal. Chem.* **2014**, *86* (18), 8974–8982. https://doi.org/10.1021/ac502838t.
(18) Holzlechner, M.; Bonta, M.; Lohninger, H.; Limbeck, A.; Marchetti-Deschmann, M. Multisensor Imaging□From Sample Preparation to Integrated Multimodal Interpretation of LA-ICPMS and MALDI MS Imaging Data. **2018**. https://doi.org/10.1021/acs.analchem.8b00816.
(19) Patterson, N. H.; Tuck, M.; Van de Plas, R.; Caprioli, R. M. Advanced Registration and Analysis of MALDI Imaging Mass Spectrometry Measurements through Autofluorescence Microscopy. *Anal. Chem.* **2018**, acs.analchem.8b02884. https://doi.org/10.1021/acs.analchem.8b02884.
(20) Bemis, K. D.; Harry, A.; Eberlin, L. S.; Ferreira, C. R.; van de Ven, S. M.; Mallick, P.; Stolowitz, M.; Vitek, O. Probabilistic Segmentation of Mass Spectrometry (MS) Images Helps Select Important Ions and Characterize Confidence in the Resulting Segments. *Mol. Cell. Proteomics* **2016**, *15* (5), 1761–1772. https://doi.org/10.1074/mcp.O115.053918.
(21) Van de Plas, R.; Yang, J.; Spraggins, J.; Caprioli, R. M. Image Fusion of Mass Spectrometry and Microscopy: A Multimodality Paradigm for Molecular Tissue Mapping. *Nat Methods* **2015**, *12* (4), 366–372. https://doi.org/10.1038/nmeth.3296.
(22) Crouch, E. E.; Li, Z.; Takizawa, M.; Fichtner-Feigl, S.; Gourzi, P.; Montaño, C.; Feigenbaum, L.; Wilson, P.; Janz, S.; Papavasiliou, F. N.; et al. Regulation of AID Expression in the Immune Response. *J. Exp. Med.* **2007**, *204* (5), 1145–1156. https://doi.org/10.1084/jem.20061952.
(23) Cho, S. H.; Raybuck, A. L.; Stengel, K.; Wei, M.; Beck, T. C.; Volanakis, E.; Thomas, J. W.; Hiebert, S.; Haase, V. H.; Boothby, M. R. Germinal Centre Hypoxia and Regulation of Antibody Qualities by a Hypoxia Response System. *Nature* **2016**, *537*, 234–238. https://doi.org/doi:10.1038/nature19334.
(24) Prism Ratio Paired T-test www.graphpad.com.
(25) Bankhead, P.; Loughrey, M. B.; Fernández, J. A.; Dombrowski, Y.; McArt, D. G.; Dunne, P. D.; McQuaid, S.; Gray, R. T.; Murray, L. J.; Coleman, H. G.; et al. QuPath: Open Source Software for Digital Pathology Image Analysis. *Sci. Rep.* **2017**, *7* (1), 16878. https://doi.org/10.1038/s41598-017-17204-5.
(26) Bemis, K. D.; Harry, A.; Eberlin, L. S.; Ferreira, C.; Van De Ven, S. M.; Mallick, P.; Stolowitz, M.; Vitek, O. Cardinal: An R Package for Statistical Analysis of Mass Spectrometry-Based Imaging Experiments. *Bioinformatics* **2015**. https://doi.org/10.1093/bioinformatics/btv146.
(27) Patterson, N. H.; Tuck, M.; Lewis, A.; Kaushansky, A.; Norris, J. L.; Van de Plas, R.; Caprioli, R. M. Next Generation Histology-Directed Imaging Mass Spectrometry Driven by Autofluorescence Microscopy. *Anal. Chem.* **2018**, *90* (21), 12404–12413. https://doi.org/10.1021/acs.analchem.8b02885.
(28) Schober, Y.; Guenther, S.; Spengler, B.; Römpp, A. Single Cell Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry Imaging. *Anal. Chem.* **2012**, *84* (15), 6293–6297. https://doi.org/10.1021/ac301337h.
(29) Kompauer, M.; Heiles, S.; Spengler, B. Atmospheric Pressure MALDI Mass Spectrometry Imaging of Tissues and Cells at 1.4-Mm Lateral Resolution. *Nat. Methods* **2017**, *14* (1), 90–96. https://doi.org/10.1038/nmeth.4071.
(30) Perkel, J. M. Starfish Enterprise: Finding RNA Patterns in Single Cells. *Nature* **2019**, *572* (7770), 549–551. https://doi.org/10.1038/d41586-019-02477-9.
(31) Vickovic, S.; Eraslan, G.; Salmén, F.; Klughammer, J.; Stenbeck, L.; Schapiro, D.; Äijö, T.; Bonneau, R.; Bergenstråhle, L.; Navarro, J. F.; et al. High-Definition Spatial Transcriptomics for in Situ Tissue Profiling. *Nat. Methods* **2019**, *16* (10), 987–990. https://doi.org/10.1038/s41592-019-0548-y.
(32) Braverman, N. E.; Moser, A. B. Functions of Plasmalogen Lipids in Health and Disease. *Biochim. Biophys. Acta - Mol. Basis Dis.* **2012**, *1822* (9), 1442–1452. https://doi.org/10.1016/j.bbadis.2012.05.008.
(33) Dean, J. M.; Lodhi, I. J. Structural and Functional Roles of Ether Lipids. *Protein Cell* **2018**, *9* (2), 196–206. https://doi.org/10.1007/s13238-017-0423-5.
(34) Braverman, N. E.; Raymond, G. V; Rizzo, W. B.; Moser, A. B.; Wilkinson, M. E.; Stone, E. M.; Steinberg, S. J.; Wangler, M. F.; Rush, E. T.; Hacia, J. G.; et al. Peroxisome Biogenesis Disorders in the Zellweger Spectrum: An Overview of Current Diagnosis, Clinical Manifestations, and Treatment Guidelines. *Mol. Genet. Metab.* **2016**, *117* (3), 313–321. https://doi.org/10.1016/j.ymgme.2015.12.009.
(35) Lodhi, I. J.; Wei, X.; Yin, L.; Feng, C.; Adak, S.; Abou-Ezzi, G.; Hsu, F.-F.; Link, D. C.; Semenkovich, C. F. Peroxisomal Lipid Synthesis Regulates Inflammation by Sustaining Neutrophil Membrane Phospholipid Composition and Viability. *Cell Metab.* **2015**, *21* (1), 51–64. https://doi.org/10.1016/j.cmet.2014.12.002.
(36) Sandalio, L. M.; Rodríguez-Serrano, M.; Romero-Puertas, M. C.; del Río, L. A. Role of Peroxisomes as a Source of Reactive Oxygen Species (ROS) Signaling Molecules; Springer, Dordrecht, 2013; pp 231–255. https://doi.org/10.1007/978-94-007-6889-5_13.
(37) Farooqui, A. A.; Horrocks, L. A. Book Review: Plasmalogens: Workhorse Lipids of Membranes in Normal and Injured Neurons and Glia. *Neurosci.* **2001**, *7* (3), 232–245. https://doi.org/10.1177/107385840100700308.
(38) Zoeller, R. A.; Morand, O. H.; Raetz, C. R. A Possible Role for Plasmalogens in Protecting Animal Cells against Photosensitized Killing. *J. Biol. Chem.* **1988**, *263* (23), 11590–11596.
(39) Crouch, E. E.; Li, Z.; Takizawa, M.; Fichtner-Feigl, S.; Gourzi, P.; Montaño, C.; Feigenbaum, L.; Wilson, P.; Janz, S.; Papavasiliou, F. N.; et al. Regulation of AID Expression in the Immune Response. *J. Exp. Med.* **2007**, *204* (5), 1145–1156. https://doi.org/10.1084/jem.20061952.
(40) Raybuck, A. L.; Cho, S. H.; Li, J.; Rogers, M. C.; Lee, K.; Williams, C. L.; Shlomchik, M.; Thomas, J. W.; Chen, J.; Williams, J. V; et al. B Cell-Intrinsic MTORC1 Promotes Germinal Center-Defining Transcription Factor Gene Expression, Somatic Hypermutation, and Memory B Cell Generation in Humoral Immunity. *J. Immunol.* **2018**, *200* (8), 2627–2639. https://doi.org/10.4049/jimmunol.1701321.
(41) Colgan, S. P.; Taylor, C. T. Hypoxia: An Alarm Signal during Intestinal Inflammation. *Nat Rev Gastroenterol Hepatol* **2010**, *7* (5), 281–287. https://doi.org/10.1038/nrgastro.2010.39.
(42) Konisti, S.; Kiriakidis, S.; Paleolog, E. M. Hypoxia—a Key Regulator of Angiogenesis and Inflammation in Rheumatoid Arthritis. *Nat. Rev. Rheumatol.* **2012**, *8* (3), 153–162. https://doi.org/10.1038/nrrheum.2011.205.
(43) Quiñonez-Flores, C. M.; González-Chávez, S. A.; Pacheco-Tena, C. Hypoxia and Its Implications in Rheumatoid Arthritis. In *J Biomed Sci*; 2016; Vol. 23. https://doi.org/10.1186/s12929-016-0281-0.
(44) Chen, X.; Song, M.; Zhang, B.; Zhang, Y. Reactive Oxygen Species Regulate T Cell Immune Response in the Tumor Microenvironment. *Oxid. Med. Cell. Longev.* **2016**, *2016*, 1–10. https://doi.org/10.1155/2016/1580967.
(45) Jellusova, J.; Cato, M. H.; Apgar, J. R.; Ramezani-Rad, P.; Leung, C. R.; Chen, C.; Richardson, A. D.; Conner, E. M.; Benschop, R. J.; Woodgett, J. R.; et al. Gsk3 Is a Metabolic Checkpoint Regulator in B Cells. *Nat. Immunol.* **2017**, *18* (3), 303–312. https://doi.org/10.1038/ni.3664.