

# Three-dimensional mass spectrometry imaging of biomedical tissues

Citation for published version (APA):

Vos, D. R. N. (2021). *Three-dimensional mass spectrometry imaging of biomedical tissues*. [Doctoral Thesis, Maastricht University]. Maastricht University. <https://doi.org/10.26481/dis.20210409nv>

## Document status and date:

Published: 01/01/2021

## DOI:

[10.26481/dis.20210409nv](https://doi.org/10.26481/dis.20210409nv)

## Document Version:

Publisher's PDF, also known as Version of record

## Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

## General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.umlib.nl/taverne-license](http://www.umlib.nl/taverne-license)

## Take down policy

If you believe that this document breaches copyright please contact us at:

[repository@maastrichtuniversity.nl](mailto:repository@maastrichtuniversity.nl)

providing details and we will investigate your claim.

# Three-Dimensional Mass Spectrometry Imaging of Biomedical Tissues

Dina Rebecca Naomi Vos

ISBN: 978-94-6423-183-0

PhD thesis Maastricht University

Three-Dimensional Mass Spectrometry Imaging of Biomedical Tissues

Cover design & layout: Naomi Vos

Printed by: ProefschriftMaken

© Naomi Vos, Maastricht, The Netherlands 2020

The research reported in this dissertation was carried out at Maastricht Multimodal Molecular Imaging Institute (M4I), Universiteitssingel 50, 6229 ER, Maastricht, The Netherlands

# Three-Dimensional Mass Spectrometry Imaging of Biomedical Tissues

## DISSERTATION

To obtain the degree of doctor at Maastricht University, on the authority of  
the Rector Magnificus  
Prof. dr. Rianne M. Letschert,  
in accordance with the decision of the Board of Deans,  
to be defended in public  
on Friday the 9<sup>th</sup> of April 2021 at 10 hours

By

**Dina Rebecca Naomi Vos**

Born in Hengelo (OV), The Netherlands, on June 20<sup>th</sup> 1991

## **Promotor**

Prof. Dr. Ron M.A. Heeren

## **Co-promotors**

Dr. Shane R. Ellis

Dr. Benjamin D. Balluff

## **Assessment Committee**

Prof. Dr. M. Honing (Chairman)

Prof. Dr. T.M. Hackeng

Dr. L. Dubois

Prof. Dr. M. Altelaar, Utrecht University

Dr. B. Heijs, Leiden University Medical Center



ITEA3



National Institutes  
of Health

The research presented in this thesis has been made possible with the funding from ITEA and RVO by means of project numbers ITEA151003 and ITEA14001 as well as funding from the National Institutes of Health (NIH) of the United States of America, Grant No. NIH R01 CA213492.

*Faith, it does not make things easy,  
it makes them possible*

Luke 1:37

*To my parents, siblings, Tommie †, and Boomer*



# Table of content

<b>Prologue</b>	General introduction	<b>9</b>
<b>Chapter One</b>	Experimental and data analysis considerations for three-dimensional mass spectrometry imaging in biomedical research	<b>21</b>
<b>Chapter Two</b>	Class-specific depletion of lipid ion signals in tissues upon formalin fixation	<b>43</b>
<b>Chapter Three</b>	Strategies for managing multi-patient 3D mass spectrometry imaging data	<b>79</b>
<b>Chapter Four</b>	2.5D N-glycan mass spectrometry imaging of esophageal adeno-carcinogenesis	<b>101</b>
<b>Chapter Five</b>	Valorization	<b>125</b>
<b>Epilogue</b>	General discussion and Summary (Samenvatting)	<b>131</b>
<b>References</b>		<b>143</b>
<b>List of Publications</b>		<b>163</b>
<b>Acknowledgements</b>		<b>167</b>
<b>Curriculum Vitae</b>		<b>173</b>



# ***Prologue***

## **General Introduction**



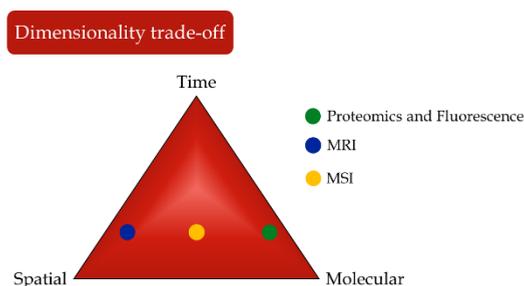
Living systems consist of one or more cells each of which plays a vital part in keeping the system alive. Cells in multicellular organisms thereby share their tasks by building different organs and tissues each with its role and individual processes. As these tissues are complex, dimensionality is a crucial consideration in biological research which encompasses:

- Space
- Time
- Molecular state

In particular, the study of the molecular complexity of the tissues requires researchers to find new ways to address the dimensionality of their research, as many different molecules are present within a tissue. Understanding this molecular complexity requires analytical chemistry to separate, identify, map the distribution of the individual chemical components, and investigate their changes over time. Therefore, dimensionality plays a key role in analytical chemistry, as the different aspects require either a spectral, spatial, or time dimension. In the spectral dimension, different molecules of a certain molecular class present in living systems are detected while in the spatial dimension the distribution of a single molecule in a living system is mapped. Investigation of changes in these dimensions in time makes up the time dimension in analytical chemistry. Although these dimensions provide insights into molecular complex living systems, they are limited. Multiple molecular classes are for instance involved in biological processes that require an increase in the spectral dimension to detect them all. These processes, in turn, take place at different histological regions inside the living systems necessitating an increase in the spatial dimension to map the distribution of the various molecules. Expanding the dimensions, both separate and combined is therefore needed to gain more insights into the molecular complexity of living systems. As biological tissues are living and volumetric, it is important to study the molecular complexity *in vivo* in the original three-dimensional (3D) conformation over time to be able to study the structural context of the sample in all dimensions. This thesis is essentially about studying molecular information of biological tissues in space (3D) over time.

In biomedical research, it is difficult to have all 3 previously mentioned points of molecular information, space, and time in one single research. The

analytical tools employed in this field differ per application or research question and are focused on either space, time, or molecular complexity. Studies are, for instance, done by employing magnetic resonance imaging (MRI) to study changes in the brain in patients with Parkinson's<sup>[1]</sup> or whether scoring of MRI features in the brain can be representative of small-vessel disease.<sup>[2]</sup> Other labs might use proteomics to determine differentially expressed proteins during disease progression or between healthy and disease tissue.<sup>[3, 4]</sup> Fluorescence measurements (wavelength and quenching measurements, *in situ* hybridization) can also be applied to investigate the relation between two genes or the interaction of proteins with membranes.<sup>[5, 6]</sup> Although the aforementioned techniques can always be used to study changes over time, there is a trade-off when it comes to insights into the complexity in the molecular and spatial dimensions they can provide (Figure 1). Common for proteomics and fluorescence, and most biomedical research, is that only a certain group of molecules is investigated or only those that are already known beforehand. The molecular complexity of the biological systems is investigated in one spectral dimension thereby not allowing a researcher to acquire a complete view of this complexity. Spatial information is needed as well as the localization of molecules inside biological systems can give more information about their role in biological processes. This spatial complexity is studied with MRI, however, this technique has a limitation in spatial resolution and is only applicable for a range of observable molecules. To gain a deeper understanding of the complexity of biological tissues, and have the best trade-off between all three dimensions, imaging techniques are needed that can map the molecular as well as the spatial complexity.

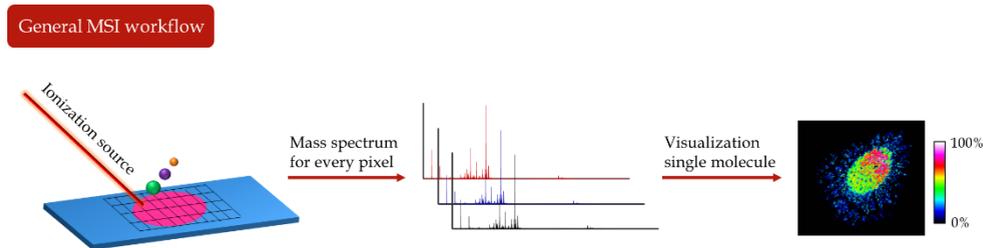


**Fig 1.** Diagram showing the trade-off between the different dimensions. Indicated is where the standard biomedical analytical tools (proteomics (green), fluorescence (green), and MRI (blue)) are located in comparison to MSI (yellow).

Many molecular imaging techniques, like positron emission tomography (PET) and near-infrared (NIR) imaging, allow the visualization of only one particular molecule that is known beforehand. Application of these techniques is targeted and they have a one-dimensional viewpoint focused on the molecule of interest. This is very useful for diagnostic purposes albeit less for research purposes where the goal often is the discovery of new molecules. A discovery approach in these cases needed using techniques that have a multidimensional viewpoint in, for instance, the molecular domain allowing detection of multiple molecules. Besides finding new molecules, spatial distributions of these molecules are important as well for biomarker discovery, investigating new classification models, or determining drug distributions and their metabolism. This requires a technique that not only has a multidimensional molecular viewpoint but also a spatial dimension. Mass spectrometry imaging further referred to as MSI, is one of those techniques that captures both the spatial complexity as well as the molecular complexity of biological systems (Figure 1). This duality makes it an advantageous additional research tool in biomedical research.

## Mass spectrometry imaging

MSI is an imaging technique for which no prior knowledge about the specific molecules present in the sample are needed as it detects and identifies a plethora of molecules from a specific class that is present in the tissue.<sup>[7]</sup> The only knowledge needed is which class to investigate due to specific sample preparation protocols for every target molecular class. At the same time, MSI maps the spatial distribution of every detected compound. It achieves this, as illustrated in Figure 2, by scanning tissue sections in a pixel-by-pixel manner where a mass spectrum is acquired for each pixel.<sup>[8]</sup> Molecular distributions are reconstructed based on the intensity in each pixel after the selection of a molecule; the higher the intensity the brighter the color depicted for this pixel. In this way, the chemical make-up of the tissue surface is mapped and elucidated. This chemical make-up is oftentimes complex as different histological tissue features have their unique molecular profile. MSI is, therefore, an advantageous analytical tool to investigate these local chemical complexities further.



**Fig 2.** Overview of a general MSI workflow. The tissue section is placed on a slide and scanned in a pixel-by-pixel manner by an ionization source. For each pixel, a mass spectrum is generated that are combined into an average spectrum for the whole tissue. After acquisition, a single molecule can be visualized based on the intensity of said molecule in every pixel.

MSI is a mass spectrometry-based technique. A mass spectrometer consists of an ionization source, a mass analyzer, and a detector.<sup>[9]</sup> MSI uses a local desorption and ionization process to enable spatially resolved analysis. Such an ionization source encompasses a desorption/ionization beam directed at the sample surface to locally generate ions. These ions are subsequently directed into the mass analyzer. Here the ions are separated based on their mass to charge ratio and transported to the detector to be measured. If enough ions with a certain  $m/z$  value hit the detector, a peak is recorded in the mass spectrum. This results in a mass spectrum where every detected mass signal can either belong to one specific molecule or multiple with almost similar mass. To determine whether one molecule or more is detected at a certain mass, a mass analyzer with a high resolving power is needed that can separate these small mass differences between the molecules. The height of the peak is a function of the concentration, ionization affinity, and ion suppression of this molecule in the experiment.

There are many different mass spectrometry imaging techniques with secondary ion mass spectrometry (SIMS), matrix-assisted laser/desorption ionization (MALDI), and desorption ionization (DESI) MSI being the most commonly applied in biomedical research. These techniques rely on different ionization mechanisms and therefore have their own unique capabilities in capturing the molecular and spatial complexity of biological systems.

SIMS employs an ion beam to desorb and ionize the surface molecules.<sup>[10]</sup> These so-called secondary ions are subsequently transferred into the mass analyzer, commonly a time-of-flight (ToF) analyzer. As these beams can be focused down to a very small spot, down to 50 nm<sup>[11]</sup>, SIMS is a coveted technique for single-cell imaging and other high spatial resolution imaging experiments ranging from nm to 10  $\mu\text{m}$  resolution. The downside of this high spatial resolution is that a lot of energy hits the sample in a single-pixel resulting in heavy fragmentation of the ionized molecules. Over the years, innovative molecular primary ion beams have been developed that cause less fragmentation, allowing more intact biomolecules to be analyzed.<sup>[12, 13]</sup>

Where SIMS uses energetic molecular ions to analyze surfaces, DESI uses charged solvent droplets at supersonic velocities to desorb and ionize molecules.<sup>[10]</sup> The composition of the solvent spray determines the detectable molecular classes, which are commonly metabolites and lipids. The solvent spray itself also determines the spatial resolution that can be achieved and is usually around 200  $\mu\text{m}$ .<sup>[14]</sup> A drawback of using charged solvent droplets is that the surface gets wet and it could lead to delocalization. However, this can largely be prevented by optimizing the instrument settings. Another benefit of DESI is that it operates at atmospheric pressure, which makes it useful for direct analyses of tissues that cannot withstand vacuum conditions.

The previously mentioned techniques are excellent for specific tissues or when high spatial resolutions are required. However, there is still a trade-off between spatial and molecular information. It is therefore that MALDI is the most employed technique in biomedical research as it can offer both high spatial resolutions down to the single-cell and cover the molecular complexity of biological tissues. Ionization in MALDI is based on laser desorption and ionization and requires sample surfaces to be covered with a fine organic crystal, termed matrix layer. This matrix crystallizes on the tissue surface and extracts the molecules of interest, based on the chosen solvent/matrix mixture. In addition, the matrix aids in the absorption of the laser light and ionization of the molecules incorporated into the matrix crystals.<sup>[10]</sup> Depending on the type of matrix applied, different molecular classes can be detected and visualized. The size of these matrix crystals as well as the laser focus spot size determine the spatial resolution that can be

achieved. MALDI-MSI experiments are, therefore, commonly performed at 20-50  $\mu\text{m}$  spatial resolution. In recent years developments in both instrumentation and matrix application methods have been made allowing for spatial resolutions of 5  $\mu\text{m}$ <sup>[15]</sup> and even 1  $\mu\text{m}$  has been reported by Spengler and Hubert in 2002.<sup>[16]</sup> These advantages make that MALDI-MSI can acquire images at spatial resolutions of 1-100  $\mu\text{m}$ . As MALDI-MSI allows researchers not only the detection of a large plethora of molecular classes but also the ability to choose which molecules to investigate by choosing a suitable matrix at high spatial resolutions, it is the most utilized MSI technique in biomedical research.

### **Biomedical tissue for mass spectrometry imaging**

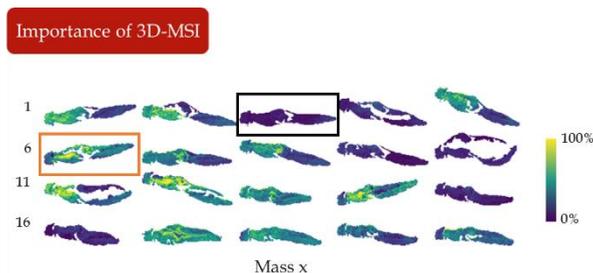
To be able to apply MSI in biomedical research the techniques have to be compatible with biomedical tissues. In most biomedical research, there are primarily two types of tissues used: animal or human. Before they become available for research purposes, they are stored and archived. When animal models are used, the tissues can generally be available as fresh frozen samples. This means that organs are directly put into liquid nitrogen to snap-freeze them after resection from the animal, thereby stopping all biological processes and preventing the degradation of the sample and its molecules.<sup>[17]</sup> This freezing does not alter the molecules and the molecular complexity of the tissue is preserved. The benefit of this is that the different molecular classes are readily accessible for all the before-mentioned MSI techniques. The tissues do not require extensive sample preparation to enable the detection and visualization of the molecular complexity with MSI. This is why these types of tissues are preferred.

However, clinical studies require a focus on human samples, obtained for instance during cancer resections. These resected cancers are typically scrutinized by a pathologist who will take a thin slice or more, stains the slice(s), and investigates if the margin of the resected material is free of cancerous cells. The remaining tissue of the resection cannot be discarded and has to be stored for a minimum amount of years according to the institution's respective guidelines. Pathology labs fixate the samples with formalin to enable long-term storage. This method cross-links proteins

together to prevent degradation, maintains the histology of the tissue, and improves the handleability by paraffin embedding, resulting in formalin-fixed paraffin-embedded (FFPE) samples.<sup>[18, 19]</sup> As a result, samples can be stored at room temperature which is much cheaper and easier compared to long-term sample storage in a deep freezer.

The molecular changes that result from these treatments make it harder for these samples to be measured with MSI since these chemical preservation steps alter the molecular information of the tissues compared to fresh frozen tissue. When these samples are embedded in paraffin it is even harder to analyze these samples as the process of embedding and its reversal results in the loss of certain molecular information. Some molecules, such as peptides and N-glycans, are demonstrated to be preserved and detectable by MALDI-MSI.<sup>[20, 21]</sup> These molecular classes, however, require extensive sample preparation steps as compared to lipid and metabolite MSI. Many innovative developments in sample preparation over the last decade now allow for routine measurements from FFPE tissues by MALDI-MSI. Even metabolites have been shown to be detectable from FFPE tissue.<sup>[22]</sup>

The compatibility of MSI with FFPE increased biomedical investigations using these types of samples as they are readily available in biobanks and allow larger patient cohort studies. These studies are commonly done in a two-dimension (2D) manner where only one section per sample is imaged that only provides a partial representation of the complete sample. Biomedical tissues are volumetric with different molecular profiles throughout the sample as different biological processes take place at different parts of the tissues making it spatially complex. Also, tumors are known to be heterogeneous with different molecular profiles at different tumor regions.<sup>[23, 24]</sup> To gain more insights into this tumor heterogeneity or a deeper understanding of the biological processes at play in tissue and how they influence each other, research needs to be expanded to the complete tissue. Hence MSI needs to expand its spatial dimension from 2D to 3D to fully capture the spatial and coinciding molecular complexity of biological systems (Figure 3).



**Fig 3.** Visualization of a mass in a 3D volume highlighting the need for 3D-MSI. If only section 3 (black box) was chosen the mass would have been classified as being absent in the tissue. If section 6 (orange box) would have been chosen it would have been classified as high abundant. In both cases, the statements would be incorrect as the intensity for this mass changes within the 3D volume.

## From two-dimensional to three-dimensional mass spectrometry imaging

Up until recently, 3D-MSI datasets were rare as one of the bottlenecks was the long measuring time per section. Over the last years, instrument developments have resulted in faster acquisition times, higher spatial resolution, and broader mass ranges. The introduction of faster MSI techniques started a new trend in mass spectrometry imaging: three-dimensional MSI. In MALDI and DESI, instead of one or a couple of sections being measured from a tissue sample, consecutive sections are measured up to a certain z-depth creating a 3D-MSI image. In SIMS, 3D imaging is achieved without sectioning. Instead, an image is generated from the surface and the imaged layer is sputtered away. This way a new layer is created that can be imaged and is unaffected by the previous image run. By alternating these imaging and sputtering cycles a 3D image is generated. Most of the 3D studies done so far have been proof-of-principle investigations. Slowly the applications are developing towards more biologically relevant investigations.

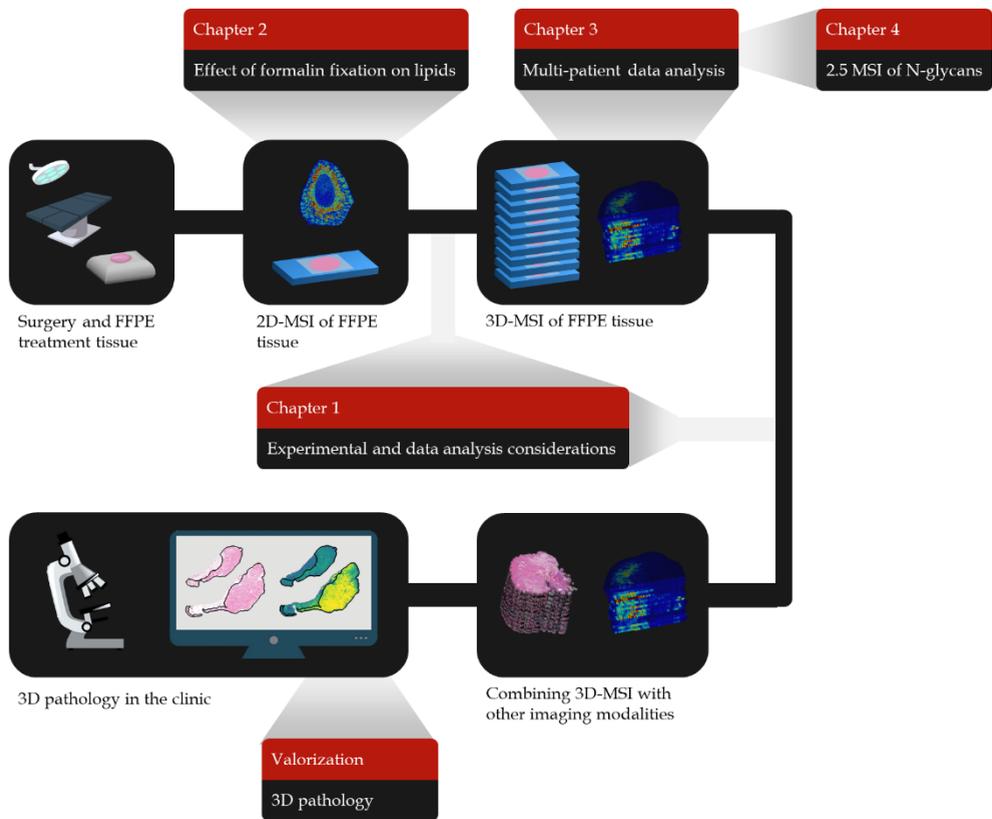
These advances in instrumentation and consequential migration to 3D-MSI also come with their difficulties. Samples have to be carefully prepared to ensure that the obtained molecular information is representative of the tissues. In addition, sample preparation has to be reproducible throughout

the entire 3D volume to prevent technical outliers and therefore gaps in the 3D stack. This requires the adaptation of established protocols and the development and inclusion of quality controls that can aid in tracking the reproducibility during all steps in a 3D workflow. Randomization of the samples to prevent batch effects is another aspect that has to be taken into account. Another big challenge is the amount of data that is generated and how to analyze them. There are no standard data analysis pipelines yet to determine possible technical outliers. Routine data analytics might fail to detect features due to a large amount of data and the large variety of features potentially incorporated in one 3D model. All these considerations, both in sample preparation and data analysis, means that experimental design is an important factor in 3D-MSI. The challenges resulting from increased dimensionality, both molecular and spatial, some of their solutions and applications are the central theme of this thesis.

## Scope of this thesis

In this thesis, 3D-MSI in biomedical research and its challenges are evaluated, innovated, and applied for cohort studies. In [chapter 1](#) an overview is given of recent 3D-MSI applications in biomedical research and experimental and data analysis considerations are being discussed. In [chapter 2](#) the effect of formalin-fixation (FF) on the lipidome is investigated to determine whether the lipidome is still representative of the tissue. This is achieved by comparing a fixed mouse and rat kidney with matched unfixed kidneys. Alteration of certain lipid classes upon formalin-fixation was explored under controlled MSI conditions. The impact of these alterations on the ability to draw biological conclusions based on lipid data on FF tissues was comprehensively studied. This is relevant as FFPE tissues are often used in biomedical research and care has to be taken with statements based on the lipid data as an indication of the state of the tissue. Based on these results high-throughput enzymatic digestions are employed as a standard for FFPE MALDI-MSI in the next chapters. A big challenge in any 3D-MSI experiment is the data analysis due to the large amount of data gathered during these experiments. In [chapter 3](#) a data analysis pipeline for 3D-MSI is therefore developed consisting of novel outlier detection methods to remove outliers from the data and a method to determine representativity. This allows the

determination of how many sections are required to be analyzed to be representative of the whole volume and their spacing. The developed data analysis pipeline is applied in **chapter 4** where esophageal adenocarcinogenesis is investigated by imaging N-glycans in resections from patients in different stages. The obtained knowledge in the previous chapters leads to the valorization chapter (**chapter 5**) where the potential usage of 3D-MSI in 3D digital pathology is discussed.



# ***Chapter One***

**Experimental and data  
analysis considerations for  
three-dimensional mass  
spectrometry imaging in  
biomedical research**

**D. R. Naomi Vos, Shane R. Ellis, Benjamin Balluff, Ron M. A. Heeren**  
**Molecular Imaging and Biology 2020**



Mass spectrometry imaging (MSI) enables the visualization of molecular distributions on complex surfaces. It has been extensively used in the field of biomedical research to investigate healthy and diseased tissues. Most of these MSI studies are conducted in a 2D-fashion where only a single slice of the full sample volume is investigated.

These 2D-MSI studies most often do not give a complete overview of the biological processes at play in a single tissue as tissues are volumetric and these processes occur within the whole volume. Ideally, these biological processes are investigated throughout the whole volume to gain a more comprehensive understanding of the full spatial complexity of biological processes.

Mass spectrometry imaging has therefore been expanded to the 3D-realm whereby molecular distributions within a 3D sample can be visualized. The benefit of investigating volumetric data has led to a quick rise in the application of single sample 3D-MSI investigations. As a result, experimental workflows and data analysis had to adapt beyond conventional 2D-MSI.

In this chapter, we discuss the current state of these aspects for 3D-MSI studies as well as ongoing developments in light of applying 3D-MSI to larger patient cohorts. Although throughput is not a limitation anymore, data handling, processing, and analysis remain bottlenecks. These barriers need to be addressed to enable 3D-MSI to be routinely applied to multi-sample studies. The experimental and data analysis considerations mentioned in this chapter will go a long way in guiding researchers into setting up their own (multi-patient) 3D-MSI studies as well as highlighting the remaining challenges and their potential solutions.

## Introduction

Mass spectrometry imaging (MSI) is a label-free molecular imaging technique for which no prior knowledge about the sample is needed. It enables the visualization of molecular distributions on solid surfaces using mass spectrometry by scanning the samples in a pixel-by-pixel manner where for each pixel a mass spectrum is generated. Visualization of molecular spatial distributions provides unique insights in many fields such as material science and biomedical research.<sup>[10, 25]</sup> In the latter, such surfaces are typically thin planar sections prepared from biological tissue. MSI has been used extensively to image the two-dimensional (2D) distributions of endogenous or exogenous (drugs & pharmaceuticals) compounds in such planar tissue sections for the study of tissue-based diseases, tissue pharmacokinetics, or the study of biomaterials in medical applications.<sup>[26, 27]</sup> However, the natural state of the original sample is volumetric, hence keeping the original three-dimensional (3D) information is important to be able to study the structural context of the sample in all dimensions.

In material science, where MSI has been employed since the 1960s<sup>[28]</sup>, it is proven that 3D-MSI can provide essential information. Secondary ion mass spectrometry (SIMS), for instance, has been used extensively for 3D elemental analysis of semiconductors, superconductors, glass, stainless steel, solid oxide fuel cell components, aerospace alloys, coatings, and biomaterials.<sup>[29, 30]</sup> The unique capability of obtaining molecular depth profiles has provided evidence on how 3D-SIMS can aid in understanding complex volumetric structures.<sup>[31-36]</sup>

The benefit of acquiring more relevant information through 3D-MSI is also of added value for biomedical research in which the added dimension can yield important contextual information about the biological tissue (see 'Benefits of 3D-MSI'). In conventional 2D imaging, a chemical snapshot of a single tissue section is generated. This snapshot corresponds to a specific location in a larger sample. Tissue stereology<sup>[37]</sup> has shown that a single section may not be representative of the whole sample. Tissues are seldom homogeneous, and the mix of cells across several 2D planes varies strongly, especially in a diseased state.<sup>[38]</sup> This introduces a potential sampling bias. In

clinical diagnostics, the discrepancy between the sampled sections and the variation across a whole tissue can lead to inaccurate predictions. This occurs especially in oncology where cancer can be spatially very heterogeneous.<sup>[39]</sup> This is corroborated by the use of radiomics in oncology, which has shown that 3D structural features contained in clinical *in vivo* scans harbor relevant clinical information.<sup>[40]</sup> An imaging feature was found for ovarian carcinoma that was predictive of the outcome after the first chemotherapy. In glioblastoma, a specific imaging pattern was able to predict epidermal growth factor receptor (EGFR) overexpression. Another prime example for the necessity of volumetric investigation is the healthy brain which has therefore been the blueprint for many 3D omics studies<sup>[41]</sup>

Consequentially, 3D-MSI has already analyzed a variety of solid tumor tissues, rodent brains, and other organs, and applications are rapidly increasing.<sup>[42, 43]</sup> While these studies provide successful use-cases, they also reveal technical aspects that require special attention and further development beyond conventional 2D-MSI. These aspects affect the whole study workflow and range from study design through sample preparation to data analysis. We will discuss them in this review, especially in the light of extending 3D-MSI to the analysis of larger sample (patient) cohorts to capture inter-individual effects. First, the most commonly used 3D-MSI techniques and their applications are briefly reviewed to illustrate the benefit of 3D-MSI in biomedical research. Then we highlight practical considerations in performing 3D experiments and 3D data analysis, as well as discuss the remaining challenges.

## 3D approaches in MSI

There are generally two different ways by which 3D-MSI can be achieved, depending on the desorption/ionization technique used: via surface sputtering or via the analysis of multiple serial sections (Figure 1).

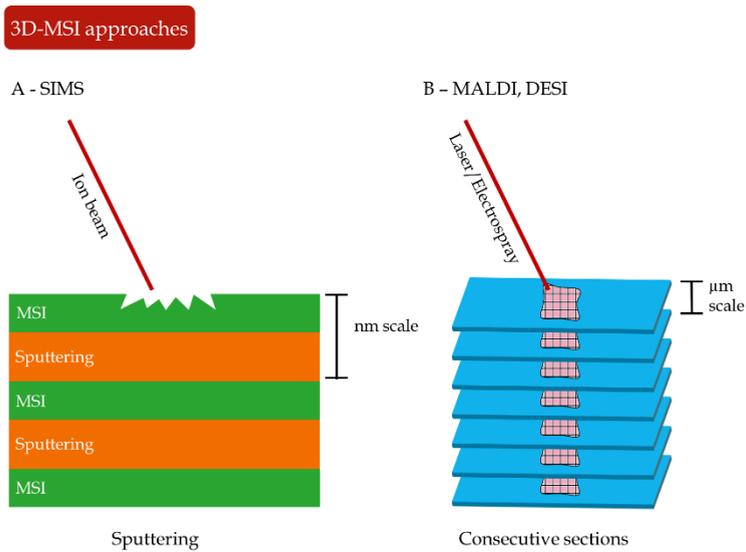
In 3D-SIMS, surface imaging and surface sputtering, where the impacting primary ion beam removes a nanometers-thin layer during the sputtering cycle, are continuously alternated until the desired depth is reached (Figure

1A). The depth-resolution depends on the type of ion beam and the associated sputter rate of the region being analyzed<sup>[44-46]</sup>, but usually ranges between 10 nm<sup>[47]</sup> and 1 nm.<sup>[48]</sup> The use of polyatomic or cluster-based ion beams enables small-sized biomolecules, such as lipids, to be released intact from the surface. Additionally, the chemical subsurface damage done by these beams is reduced or completely removed.<sup>[10, 49]</sup> This also allows an increased primary ion energy to be applied. Higher primary ion energy typically leads to higher secondary ion signals within a single pixel and results in improved image contrast.<sup>[49]</sup> The introduction of polyatomic ion guns, like the  $C_{60}^+$  and  $Bi_3^+$  ion guns with a smaller beam focus, even allows 3D analysis of single cells as it can be focused down to subcellular resolution and generates higher yields of intact biomolecules.<sup>[49, 50]</sup>

Soft-ionization techniques, in contrast, do not rely on energetic sputtering but rather on the soft desorption/extraction of biomolecules. In matrix-assisted laser desorption/ionization (MALDI) and desorption electrospray ionization (DESI), the two most common soft-ionization techniques used for MSI, a 3D-image is obtained by serial measurements of consecutive sections from a sample (Figure 1B). The final 3D dataset is then obtained via a reconstruction procedure. The overall sample volume that can be analyzed with these techniques in a given amount of time is much higher than with SIMS. The analysis volume depends on the sectioning thickness (usually 10-20  $\mu\text{m}$ )<sup>[17, 51, 52]</sup>, the number of sections, and the spacing between the sections. MALDI and DESI are therefore often employed to measure whole tissues in 3D as a result of the large volume of analytical capabilities. However, compared to SIMS, the depth-resolution between sections is limited by the thickness of the tissue section, which typically ranges from 4  $\mu\text{m}$  (paraffin-embedded tissue) to 12  $\mu\text{m}$  (frozen tissue). Molecules within a section are samples from an unknown extraction depth determined by the MALDI matrix or DESI solvent often assumed to be constant across a sample.

In MALDI-MSI, 3D applications have seen a recent rise in popularity due to advances in sample preparation and instrumentation, such as lasers with high repetition rates and fast-moving stages, which made high-throughput analysis possible.<sup>[15, 53, 54]</sup> As a consequence, whole tissue sections can now be measured within minutes, which allows a full 3D dataset to be acquired within a day.

Compared to MALDI-MSI, DESI-MSI requires less sample preparation as no external matrix is required to extract the molecules of interest. This does not necessarily mean a higher throughput for DESI-MSI over MALDI-MSI. Due to a limited spray focus, DESI has a lower spatial resolution than MALDI or SIMS. Various efforts are ongoing to increase the lateral spatial resolution that would allow for the acquisition of better quality images from different histological regions within a tissue.<sup>[55]</sup>



**Figure 1.** Schematic representation of the different approaches in 3D mass spectrometry imaging (MSI). (A) Secondary ion mass spectrometry (SIMS) continuously alternates imaging and sputtering cycles. (B) Matrix-assisted laser desorption/ionization (MALDI) or desorption electrospray ionization (DESI) rely on the sectioning of the sample into a stack of consecutive sections which are then analyzed individually and their data reassembled afterward.

## Benefits of 3D-MSI

Early 3D-MSI experiments have demonstrated feasibility and have paved the way for subsequent applications with different ionization modalities.<sup>[56-58]</sup> Here, we highlight selected applications of 3D-MSI that demonstrate its benefit or the need to contextualize the molecular information into an additional spatial dimension in certain situations.

### 3D SIMS-MSI

SIMS has been extensively used for 3D-MSI studies of single cells and tissues as reviewed by Fletcher.<sup>[34, 59]</sup> Here we highlight two applications focused on interfaces between drugs and surfaces. Biomaterials, like coronary stents, are frequently studied with three-dimensional SIMS. Coronary stents are coated with a polymer that contains an anti-inflammatory drug that is released over time to prevent blocking of the stent. To visualize the distribution of the drug sirolimus in a stent, drug/matrix-coated metal coupons were imaged with 3D-SIMS.<sup>[60]</sup> A gold ion beam was used for imaging in conjunction with a carbon-cluster ( $C_{60}$ ) ion beam for sputtering, which results in low residual molecular damage after sputtering. This is needed since sirolimus is a mid-sized (914 g/mol) pharmaceutical compound. Sputter rates have been determined under the same measurement conditions on a similar polymer material to allow accurate determination of the depth scale. It was found that large areas of the surface and subsurface channels were composed primarily of sirolimus, followed by a drug-depleted region, and lastly, a relative homogeneous drug distribution in the polymer matrix. Comparing these 3D distributions with the elution rates showed that elution occurs relatively quickly from the drug-enriched surface area while it proceeds more gradually for the subsurface regions.

Supramolecular hydrogels and their capability to enhance skin permeation of vitamin c (ascorbic acid) and its precursor (ascorbyl glucoside) have also been investigated with 3D-SIMS of *ex vivo* porcine skin.<sup>[61]</sup> Researchers demonstrated that the hydrogel enhances the skin permeation of both compounds and preserves the conversion of the precursor into vitamin c until it reaches the epidermal layer, the intended target (Figure 2A). As the depth to the epidermis is *a priori* unknown and sputtering rates might change for each layer within a sample, the authors used two methods to calculate the *z*-resolution during sputtering. Optical profilometry was applied to determine the depth of the sputter craters. This provides an estimation of the average depth, and a phospholipid ion marker was monitored during sputtering to determine when the epidermis layer is reached.

### 3D MALDI-MSI

The application of 3D MALDI-MSI has seen a steady increase over the past years. Some of these have been discussed in earlier reviews, though this is only a small fraction.<sup>[10, 62]</sup> Thereby missing interesting papers like the investigation of pathogenesis in *Francisella* infection by Scott et al.<sup>[63]</sup> and the feasibility assessment of 3D MALDI-MSI with FTICR by Jones et al.<sup>[64]</sup> One of the most investigated organs in the field of MALDI-MSI is the murine brain due to its molecular but morphologically well-charted complexity.<sup>[65, 66]</sup> The high-lipid content in the brain, which is a molecular class well compatible with MALDI- or DESI-MSI, also contributes to the frequent usage of brain sections in MALDI-MSI.

Lipid changes upon traumatic brain injury were investigated in rat brain with 3D MALDI-MSI to gain more understanding of the biochemical alterations caused by this neurological damage that is a major cause of death and disability in children and young adults.<sup>[67]</sup> The researchers obtained rat brain tissue sections every 200–250  $\mu\text{m}$  covering the entire area of the cortical-impact injury. Tissue sections were analyzed individually using MALDI-MSI of lipids at 70  $\mu\text{m}$  spatial resolution, resulting in a voxel size of 70x70x220  $\mu\text{m}$  in the final 3D dataset. The authors observed that certain lipids are specific for either the lesion site (e.g. PC (42:9) ( $m/z$  856.598)) or the ventricles (e.g. PC ( $m/z$  797.580)) that change from rostral to caudal while others co-localize with both regions (e.g.  $m/z$  741.545). Acquisition of the 3D model helped the authors gain more insight into the changes happening throughout the brain since the injury-related molecules are transferred via these ventricles throughout the brain and other organs. The 3D results provided a more global view of the impact of the traumatic brain injury and provided insight into pathological phenomena remote from the injury site.

With recent advances in instrumentation, especially in laser repetition rates, synchronized stage movement, and electronics, 3D-MSI datasets can now be generated much faster, up to 20-50 times, allowing the acquisition of biological or technical replicates. This was recently demonstrated by Paine et al., who studied medulloblastoma, the most common malignant pediatric brain tumor, in six mouse brains by 3D MALDI-MSI of lipids. The voxel size was 50x50x150  $\mu\text{m}$ , and a total of 223 sections were measured in this study.<sup>[68]</sup>

A semi-supervised segmentation of all tissue sections was performed first to find the boundaries of the primary tumors to facilitate the analysis of this large amount of data. A comparison of metastasizing vs. non-metastasizing primary tumors revealed ten lipids associated with medulloblastoma metastasis. As these lipids were observed to be very heterogeneous in their distribution throughout the tumor, single section 2D-MSI could have missed the significance of these lipids due to sampling bias. This experiment demonstrated the added value of 3D-based analyses for tumor marker discovery and exploration. Even if multiple sections are investigated from one sample, it is important to prepare and analyze the sections in a reproducible manner to reduce inter-section technical variation. Paine et al. have ensured this by batch preparation (20–25 sections) and a quality assurance approach that involved tuning the signal intensity of each imaging run on the matrix to ensure similar ion yields on all sections.<sup>[68]</sup>

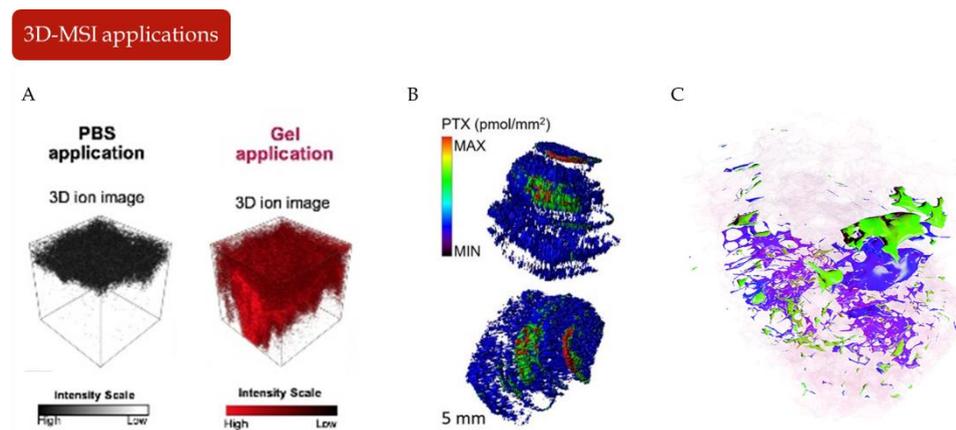
This is especially true if MSI signals are to be compared quantitatively. For instance, Giordano et al. have employed 3D MALDI-MSI to quantify the distribution of the anti-cancer drug paclitaxel in mesothelioma tumor-bearing mice.<sup>[69]</sup> Tumor sections were taken every 500  $\mu\text{m}$  and MALDI-MSI experiments were performed at 100  $\mu\text{m}$  lateral pixel size. The signal of paclitaxel was normalized to an isotopically labeled version of paclitaxel as an internal standard and quantified using a concentration series applied to a separate section. It was discovered that the distribution of the drug is influenced by the cellular heterogeneity of the tumor microenvironment showing its highest concentration at the edge of the tumor and a lower abundance in the center of the tumor with necrotic and fibrotic regions (Figure 2B). The more accurate description of the distribution of the compounds obtained demonstrates the added value of 3D-MSI for the quantitation of pharmaceutical compounds.

### **3D DESI-MSI**

Few studies have so far used DESI in a 3D fashion. An early 3D DESI-MSI study has mapped the distribution of anabolic steroid esters through bovine muscle tissue previously injected with these compounds.<sup>[70]</sup> The spatial resolution in x and y was 500  $\mu\text{m}$  and the samples were spatially separated

in the z-direction by one centimeter. The 3D profile was made at the injection site and demonstrates that the esters are indeed inside the tissue. This excludes external contamination which is an important element in the investigation of illegal treatment of animals.

In a more recent study, the unique capability of MSI to describe intratumoral molecular heterogeneity has been for the first time extended to 3D by investigating a human colorectal adenocarcinoma biopsy by 3D DESI-MSI at  $100 \times 100 \times 100 \mu\text{m}$  voxel size.<sup>[42]</sup> The application of advanced data analysis methods including deep learning and parametric t-SNE mapping helped to identify tumor subgroups and characteristic metabolites that were not detected by classical methods such as principal component analysis (Figure 2C). This also exemplifies that, as the field of 3D-MSI is evolving, novel and advanced data analysis procedures are needed to extract all biological information from 3D-MSI data.



**Figure 2.** Applications of 3D mass spectrometry imaging (MSI). (A) 3D SIMS-MSI was used to study the permeation of ascorbic acid through *ex vivo* skin samples after the application of PBS or a hydrogel. The results showed a deeper permeation when a hydrogel is used. This article was published in [61], Copyright Elsevier (2019). (B) Quantitative 3D MALDI-MSI of the anti-cancer drug paclitaxel in a malignant pleural mesothelioma tumor. Images indicate the drug is mostly located at the edge of the tumor. Figure adapted from [69], Licensed under CC BY 4.0: <http://creativecommons.org/licenses/by/4.0/>. (C) The application of advanced data analysis methods to a 3D DESI-MSI dataset of a human colorectal cancer sample revealed several metabolically different tumor subgroups highlighting the heterogeneity of tumors in three dimensions. Figure adapted from [42] - Published by The Royal Society of Chemistry.

## 3D-MSI experimental considerations

The benefit of performing 3D mass spectrometry imaging also comes with some additional requirements and considerations in terms of sample preparation, experimental design, and data acquisition. The evaluation of these workflow elements is crucial to reliably correlate and compare findings throughout a 3D-MSI volume. Below, we will discuss several considerations that have a direct impact on the outcome of a 3D study.

### Single-cell preparation for SIMS

In the case of SIMS for single-cell analysis, cells need to be prepared in such a way that the 3D shape and integrity of the cells are maintained. One way of achieving this is by chemical fixation of the cells to preserve the cellular architecture of the cell.<sup>[71-73]</sup> In this process of fixation, certain molecules, like salts, can diffuse while proteins are cross-linked together.<sup>[71, 73]</sup> Another approach that retains the integrity of the cells and prevents diffusion of molecules is to prepare frozen-hydrated cells through cryo-fixation of cells in their native hydrated state.<sup>[74]</sup> This method achieves higher ion yields for certain chemical species but requires the instrumentation to be equipped with a cold-stage to maintain the frozen-hydrated state of the sample throughout the entire workflow.<sup>[74, 75]</sup> UHV systems can result in the sublimation of water from the 3D samples which can subsequently deteriorate during cryo-imaging experiments. Alternatively, cells can be prepared for 3D-SIMS analyses by freeze fracturing where cells are trapped in a frozen sandwich, which is broken before analysis. This method maintains both cellular and molecular integrity but can lead to the loss of the top part of the cells and the fracture plane is not always reproducible. This makes it difficult to characterize both complementary fractured surfaces.<sup>[74, 75]</sup> All these sample preparation methods have their advantages (e.g. maintaining cellular integrity, preventing diffusion) and drawbacks (e.g. diffusion of molecules, unreproducible fracture planes), and choosing the best-suited method depends on the research question at hand.

## Serial section preparation

In the case of 3D-MALDI and 3D DESI-MSI, experiments might require the collection of dozens to hundreds of (semi-)consecutive sections from one sample. As a result, caution is required during sectioning to exclude the loss of individual sections and retain the appropriate order and spacing in the 3D stack. Keeping the samples in the same orientation in a consistent shape will, later on, simplify the alignment of the sections into a 3D volume. Minor misalignment of sections can be digitally corrected afterward.<sup>[76]</sup> This registration is often based on prominent morphological features, and heavy distortion can render molecular images unsuitable for inclusion in the 3D volume. This can be overcome by embedding the tissue in embedding materials like optimal cutting temperature (OCT) or gelatin.<sup>[19, 52, 77, 78]</sup> Care has to be taken that the right embedding material is chosen, compatible with the measurement. Embedding with OCT is generally not recommended for MSI as it contains a polymer that easily ionizes in positive-ion mode and may cause ion suppression or mask tissue-specific signals. Synthetic polymer-based embedding media also run the risk of smearing across tissue samples during sectioning.<sup>[79]</sup> Gelatin and CMC are biopolymer-based and generate less spectral interference, which renders them more compatible with MSI experiments of biological surfaces.<sup>[80]</sup> Care has to be taken that biomolecules do not delocalize during the embedding process. Using an embedding medium also allows the use of fiducial markers that can aid in the registration process. Fiducial markers are placed in the embedding block and may help to determine the position, orientation, and distortion of each section. They are also very suitable for use in the co-registration of MSI data with images from different imaging modalities.<sup>[81-83]</sup>

## Experiment design

3D-MALDI and 3D DESI-MSI experiments consist of the sequential measurement of serial tissue sections. If the sections or experiments have not been properly randomized, the entire contiguous parts of the 3D volume might be affected by a technical bias, which cannot be distinguished from a biological effect (Figure 3A). It is, therefore, necessary to use randomization in both sample preparation and data acquisition. Serial sections should be

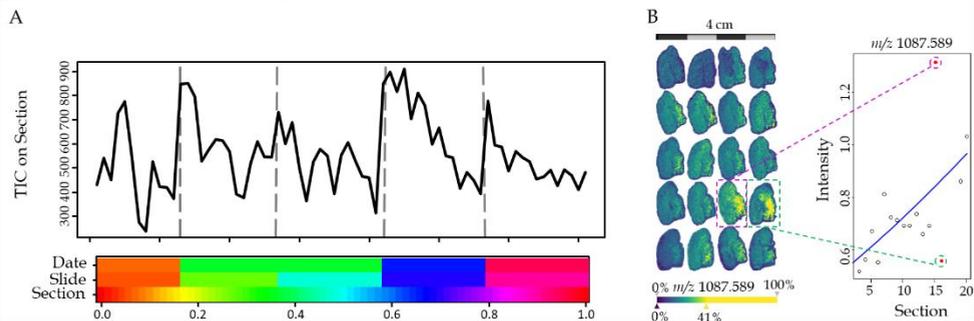
distributed in a randomized fashion within and between the slides already during the sectioning process. This overcomes potential batch effects resulting from technical variance in slide preparation (cutting, on tissue chemistry, matrix application) and instrument-related batch effects such as a decrease or drift of system sensitivity during long runs of acquisition.

While randomization itself will certainly reduce a systematic technical bias that can be confused with a biological effect, it will not recognize technical outliers. This becomes even more important when multiple 3D datasets from different tissues are to be measured and compared. The total project execution duration exceeds classical 2D-MSI projects by a factor corresponding to the average number of sections per 3D sample. The total sample preparation of a multi-sample 3D-MSI project can spread over weeks and months as only a single sample can usually be cut at a time. This increases the possibility of observing time-related batch effects between and within 3D-MSI datasets. A way to monitor experimental quality is the application and integration of quality control measurements. Quality control compounds can be applied before measurement or matrix application, thereby independently assessing either instrument performance or sample preparation, respectively. For instance, protein and peptide standards can be spotted onto the slides next to the tissue to determine matrix quality or digestion efficiency on a slide-by-slide basis.<sup>[84, 85]</sup> Unfortunately, this is not yet routinely applied in the field of 3D-MSI. Only recently, we have demonstrated the usefulness of quality controls to track digestion efficiency in a multi-patient 3D-MSI study.<sup>[85]</sup> In this study, we incorporated cytochrome *c* as quality control for digestion efficiency on a slide basis. By performing principal component analysis (PCA) on the cytochrome *c* spots and using a 95% error ellipse, we could identify 11 slides on which the digestion was insufficient. This translated to 22 sections out of a total 280 measured giving a dropout rate of < 8%. In addition to tracking sample preparation quality, it is also a necessity to monitor day-to-day instrumentation performance as instrument sensitivity can vary from one experiment to another. Monitoring instrumentation performance with a quality control compound measured over time can assist in the identification of technical outliers.<sup>[86]</sup>

## Data acquisition

Large volume measurements require careful planning of the total workflow, from sample preparation to data acquisition. Time can be a critical factor and influence the resulting 3D molecular dataset. Sample preparation should happen “just-in-time” before the measurements otherwise degradation of the samples might occur which might lead to false discoveries along the volume.<sup>[87, 88]</sup> Total acquisition time should be optimized to minimize the experimental variance within one 3D-MSI dataset while keeping throughput as high as possible. Dataset size is another critical factor. It increases linearly with the number of sections and samples and is quadratically dependent on the spatial resolution. The resulting total dataset size often determines the choice of optimal spatial resolution, the mass range, and the number of sections per sample. This also ensures during data analysis the various software packages and algorithms can read a full dataset in memory if so desired. We have recently proposed a procedure to determine the minimum number of sections needed to retain a minimum amount of the full information extracted from a 3D MALDI-MSI experiment.<sup>[85]</sup> This representativity analysis is based on calculating the correlation coefficient of a subset of samples to the full 3D dataset. In our study, we found that every third of the measured 20 sections were needed to reach a minimum correlation of 0.99 to the full 3D dataset. This way, a preselected, representative sample is defined to prospectively save sections from alike samples and reduce the data size of the study.

### Experimental considerations



**Figure 3.** Experimental considerations in 3D-MSI. (A) When no randomization has been performed, the fluctuation of the overall signal intensity across every single section, represented by its total ion count (TIC), can be caused by other factors such as date of the experiment, the slide, or the position of the section on the slide. (B) Assuming proper randomization and that closely spaced sections should have similar molecular signals, we have recently proposed a z-direction based regression of molecular signals. Outliers are detected whose molecular signals are not in line with previous and successive sections. This article was published in [85], Copyright Elsevier (2019).

## 3D-MSI data analysis considerations

### Outlier detection

One of the challenges that MSI faces is the lack of a generic and reliable experimental outlier detection method to ensure a high degree of comparability for both inter- and intra-sample comparisons. Recently, we proposed a detection method to check for any possible outliers in 3D-MSI datasets. The method is based on z-directed regression analyses within a 3D volume and identifies sections as outliers whose molecular signals deviate statistically too much from the expected signal intensity (Figure 3B).<sup>[85]</sup> This method is generic and captures most experimental biases unless too many sections are affected at the same time from that bias thereby driving the regression analysis. This could be the case if many consecutive sections are on the same slide. These sections undergo all experimental steps at the same time and therefore all get affected by matrix application or a poor MS instrument performance at the same time.

### 3D image reconstruction

It is of interest in many studies to align and stack the individual MSI images of a completed 3D-MSI dataset to construct a 3D volume, which can be used for an advanced interpretation of the spatial context of the molecular images. Embedded fiducial markers can be used as reference points to automatically co-register and stack the single sections (Figure 4A).<sup>[81]</sup> However, the 3D volume reconstruction can be completed without fiducial markers. This is often performed manually by using either anatomical features in optical images that are already aligned to the MS images<sup>[89, 90]</sup> or well-structured MS images to spatially align the consecutive MSI slices (Figures 4B-C).<sup>[67, 68]</sup> These approaches are only suitable for tissues with well-defined and visible structures in the optical or MS images. For the task of aligning and reconstructing 3D visualizations, different software applications have been used<sup>[58, 91-93]</sup>, including ImageJ (<https://imagej.nih.gov/ij/>)<sup>[94]</sup>, R software (<https://www.r-project.org/>), SCiLS Lab (Bruker Daltonik GmbH, Bremen, Germany), or Autoaligner® combined with Imaris (both from Bitplane, Zurich, Switzerland).<sup>[69]</sup> With Autoaligner® image features are used to align the sections which are built into a volume with Imaris.<sup>[92]</sup> In the case of ImageJ, there are different ways to align and can be done by first aligning stained sections using an automated rigid-body registration and then align the MSI images to this. With SCiLS, the sections are manually co-registered together based on the shape of the sections.

### Multimodal 3D-imaging

3D-MSI also opens the possibility of correlating the volumetric molecular information with data from other (3D) imaging modalities.<sup>[95]</sup> A common example is the combination of 3D MALDI-MSI data with standard histology techniques. Therefore, Lotz and coworkers analyzed 162 consecutive sections of an oral squamous cell carcinoma with MALDI-MSI, conventional hematoxylin and eosin, and immunohistochemistry. The transfer of these two modalities (MSI and optical microscopy) into the same coordinate system allowed all 3D data to be viewed next to each other, which led to a better understanding of the functional heterogeneity within the tumor as spatial context and progression can readily be determined from the H&E

(Figure 5A).<sup>[96]</sup> The visual integration of MSI data with data from *in vivo* imaging techniques also harbors great potential by combining the molecular and anatomical information from different spatial scales. In this context, Abdelmoula et al. developed an automatic co-registration between 3D-MSI and MRI that tackles the differences in spatial scales and coordinate systems, the lack of direct spatial-correspondences, or nonlinear tissue deformations (Figure 5B).<sup>[97]</sup>

### **Spatial statistics in 3D-MSI**

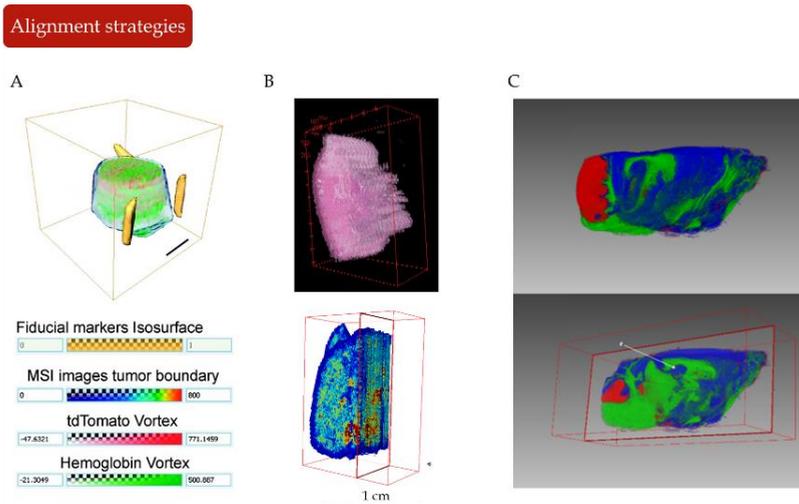
3D-MSI data puts new demands on data analysis strategies themselves. It is already known in 2D-MSI that the spatial context of the single mass spectra is a factor that can be exploited for multivariate analyses or has to be accounted for when performing inferential statistical analyses. Spatial-aware segmentation, as reported by Alexandrov and Kobarg<sup>[98]</sup>, would have to be extended to the clustering of voxels. Likewise, statistical approaches that account for the inherent spatial autocorrelation in MSI data would have to be extended to the third dimension.<sup>[99]</sup> Morphometric characteristics of MSI data have been shown by Picard de Muller et al. to carry biomedical relevant information, which would also have to be calculated on voxels instead of 2D pixels.<sup>[100]</sup>

Recently, Abdelmoula et al. have translated and extended the non-linear t-SNE dimensionality reduction to the interactive visual exploration of 3D-MSI datasets using a hierarchical version of t-SNE. They have demonstrated that this approach can rapidly identify regions of interest within large high-dimensionality 3D-MSI data sets.<sup>[101]</sup>

### **Other data analysis challenges in 3D-MSI**

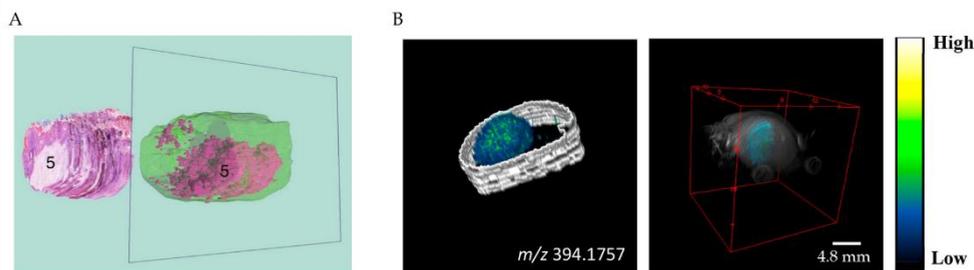
Other, less apparent data analysis challenges remain, such as the calculation of average spectra across a 3D volume when serial sections are not evenly spaced. This requires the development of a weighted average procedure that uses the information of the z-position of all sections that are taken into account. A practical challenge is the annotation of tissue regions, already a

huge bottleneck in 2D-MSI studies, and even more so in 3D-MSI studies, as multiple sections need annotation. Semi-supervised learning approaches where partial annotations by a pathologist are combined with deep learning on the MSI data are proposed for the classification of the remaining tissue by Inglese et al.<sup>[42]</sup> If the sections are consecutive and have already been aligned to form a 3D model, one could also consider interpolating the partial annotations throughout the volume. In that light, one could also foresee the translation of the already existing approaches from unbiased tissue stereology to the molecular 3D-MSI data, where stringent sampling methods and geometrical and statistical principles are used to obtain accurate and precise three-dimensional information.<sup>[37]</sup> Ultimately, novel data analysis approaches need to be developed that a) will enable the researcher to extract the added information provided by the third dimension from a 3D-MSI dataset and b) does not increase the manual workload for the researcher and maintains it similar to the 2D-scale.



**Figure 4.** Alignment strategies in tissue-section based 3D-MSI. (A) Alignment based on embedded fiducial markers requires a previous embedding of the sample into a medium but delivers good results. Adapted with permission from [81]. Copyright (2012) American Chemical Society (B) An alignment based on optical images requires the optical images to be linked to the MSI data before data analysis e.g. during the experiment. (C) Alignment based on only the MS images requires visible and well-defined structures in the images. The coarseness of these structures should match the spatial resolution of the MS images. Figure adapted from [68], Licensed under CC BY 4.0: <http://creativecommons.org/licenses/by/4.0/>.

### 3D-MSI integration



**Figure 5.** Integration of 3D-MSI with other modalities. (A) A 3D MALDI-MSI dataset consisting of 162 consecutive sections of an oral squamous cell carcinoma was combined with the corresponding 3D stack of histological images. Both volumes were placed in the same space allowing the co-visualization of both modalities at the same depth. This article was published in [96], Copyright Elsevier (2017). (B) A 3D MALDI-MSI dataset of glioblastoma in a mouse brain was also combined with MRI. The visualization on the left shows a molecule that highlights the tumor area and on the right, the same ion is overlaid with the MRI image. Adapted with permission from [97]. Copyright (2019) American Chemical Society.

## Conclusion

Mass spectrometry imaging is an added tool for the spatially resolved analysis of molecules in biological tissue specimens in many fields of biomedical research. 3D-MSI studies are gaining momentum as the benefit of the added information from the 3<sup>rd</sup> dimension is slowly beginning to outweigh the additional experimental and post-experimental efforts. While the throughput is not a limitation anymore nor by sample preparation<sup>[54]</sup> nor by instrumentation speed<sup>[53]</sup>, data handling, processing, and analysis remain bottlenecks. Dedicated software solutions exist but still cannot alleviate the additional data analysis workload that results from 3D studies. This holds especially true for histological annotations, spectral and image alignments, co-registration for integration with other imaging modalities, experimental outlier detection, and smart 3D visualizations. This has been recognized by the field shown by the many recent publications. Once these barriers are alleviated, 3D-MSI will be routinely extended to multi-sample studies. This results in a rapid three-dimensional, fully spatially aware molecular biology method that will cause a paradigm shift in patient diagnostics.

## Conflict of interest

The authors declare that they have no conflict of interest.

## Acknowledgments

The work was performed in the M4I research program and was financially supported by the Dutch Province of Limburg as part of the “LINK” program. D.R.N.V has been partially funded by ITEA and RVO (ITEA151003/ITEA 14001). B.B received financial support from the European Union (ERA-NET TRANSCAN 2; Grant No. 643638). S.R.E received financial support from the Dutch Research Council (NWO) through the VIDI program (project number TTW.VI.Vidi.198.011) and the Australian Research Council Future Fellowship funding scheme (FT190100082).

# Experimental and data analysis considerations for three-dimensional mass spectrometry imaging in biomedical research

# ***Chapter Two***

## **Class-specific depletion of lipid ion signals in tissues upon formalin fixation**

**D. R. Naomi Vos**, Andrew P. Bowman, Ron M. A. Heeren, Benjamin Balluff,  
Shane R. Ellis

*International Journal of Mass Spectrometry* **2019** 446, 116212



Fresh frozen tissues are preferred for mass spectrometry imaging, however, these can be difficult to obtain, especially when large sample sizes like patient cohorts are required. Tissues that are more readily available for these ends are formalin-fixed paraffin-embedded (FFPE) tissues as these are widely used in pathology and stored in large biobanks.

FFPE tissues are first fixed with formalin due to its ability to prevent tissue degradation and better preserve tissue morphology. By embedding the tissues with paraffin after fixation they can be stored long-term at room temperature. Although this type of tissue processing is favorable for pathology labs, for MSI it poses a challenge as certain classes of molecules, like lipids, cannot be detected anymore and tissues require longer sample preparation to be analyzed with MSI.

Most molecular classes that are therefore analyzed on FFPE tissues with MSI are proteins and N-glycans. However, recently the detection of lipids and metabolites from formalin-fixed paraffin-embedded tissues with matrix-assisted laser desorption/ionization mass spectrometry imaging (MALDI-MSI) was demonstrated. This poses exciting possibilities for future studies, however, the resulting lipid profiles differ markedly from those typically generated from fresh frozen tissue.

In this chapter, we want to shed light on the effects of formalin-induced modifications on lipids in tissues by comparing 24h formalin-fixed (FF) and untreated mouse and rat kidney tissues using time-of-flight MALDI-MSI and Fourier transform MALDI-MSI combined with tandem mass spectrometry.

These analyses revealed distinct and class-specific depletion of ion signals of certain lipid species upon FF, in particular for those containing amine functionalities. As MSI moves closer towards clinical application and necessitates an increased analysis of formalin-fixed tissues our results provide a route to rationalize the observed tissue-dependent lipid compositions observed from such tissues.

## Introduction

Formalin fixation (FF), especially when combined with paraffin embedding, is the gold-standard approach for tissue preservation in pathology. The mechanism of preservation arises from the formalin-induced chemical modification of biomolecules, primarily via the reaction of formaldehyde with primary amines and thiols.<sup>[18]</sup> The most common outcome of this reaction is the crosslinking of biomolecules through the formation of methylene bridges which deactivates enzymatic activity enabling long-term storage. Despite the induced chemical modifications a variety of molecular analysis methods have demonstrated that FF tissues can still yield rich biochemical information.<sup>[102-107]</sup>

The detection of lipids and metabolites has recently been demonstrated from formalin-fixed paraffin-embedded (FFPE) tissues using matrix-assisted laser desorption/ionization mass spectrometry imaging (MALDI-MSI).<sup>[22, 108]</sup> Importantly, the resulting spectral profiles were found to enable tissue-type differentiation between two renal tumor types and revealed tissue-specific distributions. However, the observed lipid profiles differed markedly from those typically generated from fresh frozen tissue, consisting of notably less detected lipid species than typically observed from fresh frozen tissue. These differences were largely attributed to lipid removal during the tissue processing steps that require organic solvents, as has also been reported using ultrahigh performance liquid chromatography.<sup>[107]</sup> However, it is reasonable to expect lipid alterations also originating from the fixation process itself.<sup>[109]</sup> Using extracts from cell cultures Cacciatore et al. reported that the majority of lipid species can still be detected after FF.<sup>[107]</sup> However, little information on the changing ion abundances and the possibility of class-specific depletion induced by FF was provided. Such depletion has previously been observed for FF brain extracts analyzed with LC-MS/MS.<sup>[109]</sup> Furthermore, previous MSI studies have observed spectral changes when analyzing FF tissues compared to fresh frozen tissues<sup>[110-112]</sup>, although broad identification based on high mass accuracy and MS/MS and detailed insight into the changing ion abundances of many lipid classes was not provided. Given the increasing role of MSI in clinical research<sup>[27, 113]</sup> and its potential for tissue classification based on local lipid composition, it is

important to understand the alterations in lipid profiles induced by FF and FFPE in tissues analyzed by MSI.

In this study, we explicitly investigate the effects of FF on the nature and abundance of lipid signals observed from mouse and rat kidney tissue using time-of-flight MALDI-MSI and Fourier-transform MALDI-MSI combined with tandem mass spectrometry. By exploiting the high mass accuracy and monoisotopic precursor ion selection these results unequivocally demonstrate the broad lipid class-specific depletion of ion signals observed during direct MSI analysis of untreated and FF tissues using high-confidence, on-tissue lipid identification. They also confirm previous studies demonstrating high-quality MSI can be obtained for many lipid species from FF tissues. These results can help researchers rationalize and understand the observed MSI data from FF and FFPE tissue and determine if and how reflective they are of the native tissue environment.

## Experimental

### Materials

ULC/MS-CC/SFC grade methanol absolute and water was purchased from Biosolve Chimie SARL (Dieuze, France). Anhydrous chloroform, crystalline norharmane ( $C_{11}H_8N_2$ ), and zinc sulfate heptahydrate (99% A.C.S. reagent) were purchased from Sigma-Aldrich (Zwijndrecht, The Netherlands). Formalin fixative (3.7-4.0%) containing zinc sulfate ( $\leq 1\%$  according to manufacturer) and acetate buffered at pH 5.6-5.8 (Unifix) was purchased from Klinipath BV (Duiven, The Netherlands). Indium tin oxide-coated conductive (ITO) glass slides ( $R_s$  4-8  $\Omega$ /sq) were purchased from Delta Technologies Ltd (Loveland, Colorado, USA). Hematoxylin solution modified according to Gill and entellan<sup>®</sup> new was purchased from Merck KGaA (Darmstadt, Germany). Eosine-Y, Alcoholic was purchased from Avantor<sup>®</sup> Performance Materials B.V. (Deventer, the Netherlands), and coverslips were purchased from Thermo Scientific (Waltham, Massachusetts, USA).

## Sample preparation

Two mouse kidneys were sourced from the same animal at John Hopkins University School of Medicine under the ethical approval by the Institutional Animal Care and Use Committee at the Johns Hopkins University School of Medicine (MO17M190). Two rat kidneys were sourced from the same rat at Maastricht University Medical Centre under the ethical approval by the institution's review board, the Animal Ethical Committee, the Inspection for Animal Welfare, and the Dutch Central Committee of Animal experiments (AVD107002016720). All organs were snap-frozen by immediately transferring them into liquid nitrogen after removal and stored at  $-80^{\circ}\text{C}$  until fixation and/or sectioning. One kidney of each animal was fixed for 24 hours in a zinc sulfate buffered formalin fixation solution following a standard pathology protocol. Prior to fixation, the tissues were first acclimated for an hour at  $-20^{\circ}\text{C}$  after removing it from  $-80^{\circ}\text{C}$ . This was followed by moving the sections to room temperature for an hour to ensure they were completely thawed before emerging them in the fixation solution. After fixation for 24 hours, the fixation solution was removed and the fixed kidney was stored back at  $-80^{\circ}\text{C}$  until sectioning. The other kidney of each animal was left untreated and kept at  $-80^{\circ}\text{C}$  until sectioning. Both fixed and untreated kidneys were transversally sectioned at the thickest part of the kidneys at  $-20^{\circ}\text{C}$  on a Microm HM525 cryotome (Thermo Fisher Scientific, Waltham, Massachusetts, USA). Sections were cut 10-12  $\mu\text{m}$  thick and thaw-mounted onto ITO glass slides.

For the MALDI-ToF measurements, slides were coated with 8 layers of norharmane matrix solution (7 mg/mL norharmane in 2:1 chloroform:methanol (v:v)). An automated TM-Sprayer (HTX Technologies, LLC, North Carolina, USA) was used for matrix application with a flowrate of 0.12 mL/min, 30 seconds drying time between each layer, and a temperature of  $30^{\circ}\text{C}$ . Slides for MALDI-Orbitrap analysis were prepared using identical parameters but coated with 15 layers of norharmane matrix.

After MSI slides were subsequently washed with 2 x 70% ethanol followed by Milli Q water for 3 minutes each before staining with hematoxylin for 3 minutes. Excess of hematoxylin was removed with a 3-minute water wash followed by eosin staining for 30 seconds followed by another 3-minute

water wash to remove the excess of eosin. This was followed by a 1-minute ethanol wash and 30 seconds xylene wash before coverslips were placed on top of the slides using Entellan as a mounting medium to adhere the coverslips to the sections. Optical images of the H&E stained tissues were acquired using a MIRAX scanner (Zeiss, Breda, The Netherlands).

Washing of the untreated rat kidneys with zinc sulfate was performed by submerging a slide with rat kidney sections in a 1% zinc sulfate solution in water for 30 seconds.

## **MALDI-MSI measurements**

All MALDI-ToF-MSI was acquired at 50  $\mu\text{m}$  spatial resolution on a rapifleX MALDI-ToF/ToF instrument<sup>[15]</sup> (Bruker Daltonik GmbH, Bremen, Germany). Tissues from each kidney (mouse and rat) were analyzed in both negative-ion and positive-ion mode using adjacent sections for each polarity and data acquired in the  $m/z$  range of 200-2000. The laser scanned across a 45x45  $\mu\text{m}^2$  area using a 50x50  $\mu\text{m}^2$  raster size with 200 laser shots summed for each position. Laser power and focus were optimized for each tissue section.

MSI experiments at high mass resolution and accuracy were performed using an Orbitrap Elite Mass Spectrometer (Thermo Fisher, Bremen, Germany) equipped with an ion-funnel-based MALDI interface (Spectrograph LLC, Kennewick, WA, USA) as recently reported.<sup>[114]</sup> Three untreated and FF mouse kidney sections were each measured in the negative-ion mode at a pixel size of 50 x 50  $\mu\text{m}^2$  between  $m/z$  200-2000. Data was acquired using a nominal mass resolution of 240,000 (FWHM at  $m/z$  400), an injection time of 250 ms, and a laser repetition rate of 1 kHz. In addition, one FF and unfixed section from both mouse and rat kidney tissues were analyzed in positive- and negative-ion mode using the data-dependent acquisition (DDA)-imaging method as recently described.<sup>[115]</sup> Data was acquired using a stage step size of 25 x 50  $\mu\text{m}^2$ . This approach acquires one full-scan Fourier transform mass spectrometry spectrum in the Orbitrap and one DDA-MS/MS scan in the ion trap at each 50 x 50  $\mu\text{m}^2$  area of tissue. Acquisition in this manner allows simultaneous acquisition of both a high

mass resolving power MSI dataset, accurate mass measurements of all detected lipids, and ion trap MS/MS spectra of lipids using monoisotopic precursor selection. DDA-imaging was performed in both negative-ion and positive-ion mode on consecutive tissue sections and data acquired within the  $m/z$  range 200-2000 using an ion injection time of 250 ms and a nominal mass resolution of 240,000 (FWHM at  $m/z$  400) for the Orbitrap (MSI) scans. Typical mass accuracy was  $\leq 2$  ppm. Both ion activation and mass analysis of the fragment ions were performed in the dual pressure linear ion trap using resonant collision-induced dissociation (CID). In negative-ion mode fragmentation was performed using a normalized collision energy of 38 (manufacturer units) and an activation  $q$  of 0.25. In positive-ion mode fragmentation was performed at a normalized collision energy of 30 (manufacturer units) and an activation  $q$  of 0.17 to allow for the detection of the low  $m/z$  fragment ion at  $m/z$  184 produced from PC and SM lipids.

## MSI data analysis

MALDI-ToF data was imported into SCiLS software version 2016b (SCiLS GmbH, Bremen, Germany) for further analysis. To define regions-of-interest (ROIs) and extract region-specific spectra, MSI data were co-registered with the optical images of the post-MSI H&E-stained tissue sections.

High mass resolution imaging data (excluding DDA-imaging data) was converted into imzML using ImageInsight software (Spectrograph LLC, Kennewick, WA, USA) and imported into SCiLS software and the publically accessible METASPACE annotation platform.<sup>[116]</sup> Regions-of-interest (ROI) were manually defined by visual correlation with the H&E stained tissues. ROI-spectra were then extracted from the cortex and medulla regions of each section and loaded into mMass software<sup>[117]</sup> for peak picking and peak intensity determination. For MSI data visualization this data was also converted to mzXML format using RawConverter software<sup>[118]</sup> and visualized using in-house scripts for MATLAB (version R2014a, The MathWorks, Natick, USA) as previously described.<sup>[114]</sup>

DDA-imaging data were converted to mzXML format using a RawConverter<sup>[118]</sup> and analyzed using in-house scripts for MATLAB (version

R2014a, The MathWorks, Natick, USA) as described above and previously.<sup>[115]</sup> DDA-MS/MS data was analyzed using Thermo Xcalibur 2.3.26 (Thermo Fisher, Bremen, Germany).

## Lipid identification

Accurate  $m/z$  values for precursor ions obtained from the DDA-imaging data were used to first determine the sum-composition formula of the detected lipid species. Supporting sum-composition fragments or those allowing assignment of individual acyl chains were acquired from the DDA-MS/MS data of the untreated tissues. Assignment of the fragments in the MS/MS spectra was performed using the online ALEX<sup>123</sup> database (<http://alex123.info/ALEX123/MS.php>) and the proposed nomenclature by Pauling et al. for all identified fragments reported in the ALEX<sup>123</sup> database.<sup>[119]</sup> In some cases, several supporting fragments not reported in the ALEX<sup>123</sup> databases were also observed and annotated based on prior literature reports.<sup>[120]</sup>

Selected MS/MS measurements were performed using higher-energy collisional dissociation (HCD) combined with Orbitrap detection of fragment ions to aid in the identification of several unknown ion species. HCD spectra were obtained using a 1.0 Da isolation window, normalized collision energy of 40 (manufacturer units), and detection of fragment ions at a mass resolution setting of 240,000 (FWHM at  $m/z$  400) in the Orbitrap. HCD-MS/MS spectra were acquired from a selected list of precursor ions using a 5000 ms injection time while continuously moving the stage over the tissue at 1 mm/sec. Data for each precursor was acquired for two minutes prior to spectral averaging for analysis.

For both broader identification of sum-composition lipid species and determining the percentage contribution of each lipid class to the total identified lipid signal in the mouse kidney dataset the ALEX<sup>123</sup> software package was used<sup>[115, 121, 122]</sup>. For each replicate, an average spectrum of each sample was first generated. Average spectra were then used to identify lipids using an  $m/z$  tolerance of  $\pm 2.5$  mDa, a relative intensity threshold of 0.1% base peak intensity (defined as the most intense identified lipid in the

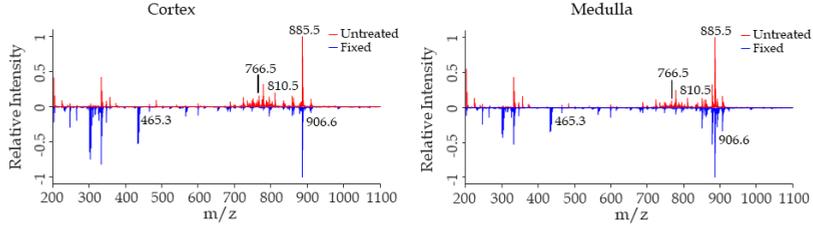
averaged spectra), and searching for the following lipid classes as [M-H]<sup>-</sup> ions: PA, PE, PS, PI, PG, their corresponding ether and lyso variants, SHexCer, Cer, GM1, GM2, GM3, and CL. The percentage contribution of each lipid class to the total identified lipid signal was then calculated using Tableau Desktop version 2018.2 (Tableau Software, CA, USA). Finally, the percentage contribution of each class within a fixed and unfixed replicate was averaged to generate the final dataset.

## Results and Discussion

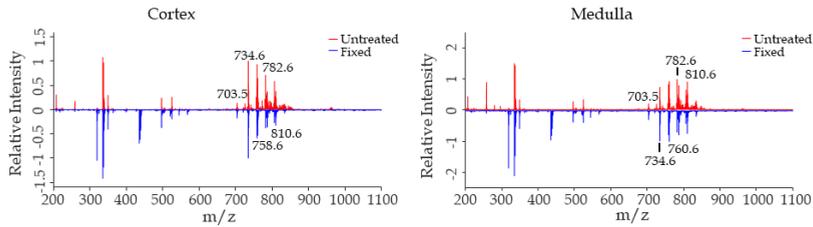
Three transversal sections taken from both untreated and formalin-fixed (FF) mouse kidney tissue were analyzed using MALDI-ToF-MSI in both negative-ion and positive-ion mode. Analysis of this data revealed drastic differences in the mean spectra generated from the cortex and medulla regions from untreated and FF tissues (Figure 1). Spectra from individual sections are provided in Figures 2 and 3. In particular, the depletion of a variety of lipid-related peaks is observed in the negative-ion mode upon FF, especially in the  $m/z$  700-850 range. We note the additional freeze/thaw cycle undergone by the fixed tissue is expected to have a negligible influence on the observation made during this study. Previous work has demonstrated that the primary artifact induced by repetitive freeze/thaw cycles is an increase in free fatty acid signals, and this effect becomes more pronounced with more cycles. Given only one additional freeze/thaw cycle was performed, and that no increase in free fatty acid signals is observed in the fixed tissue, this effect on the data analysis can be excluded.

## Region-specific spectral comparison

## A - Negative mode

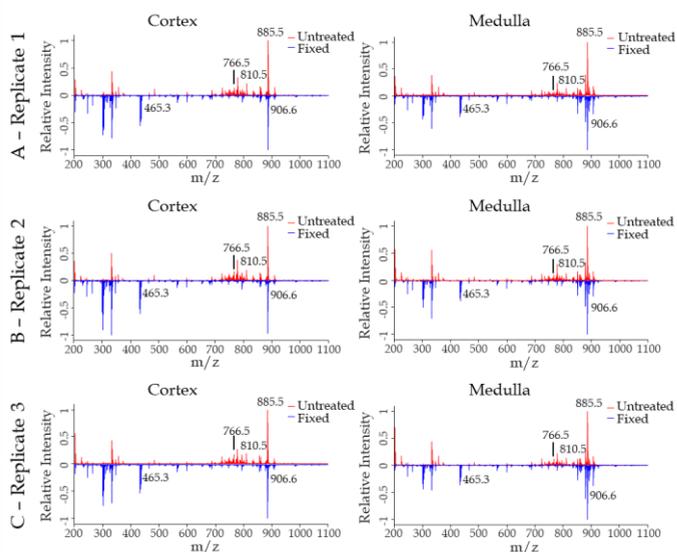


## B - Positive mode



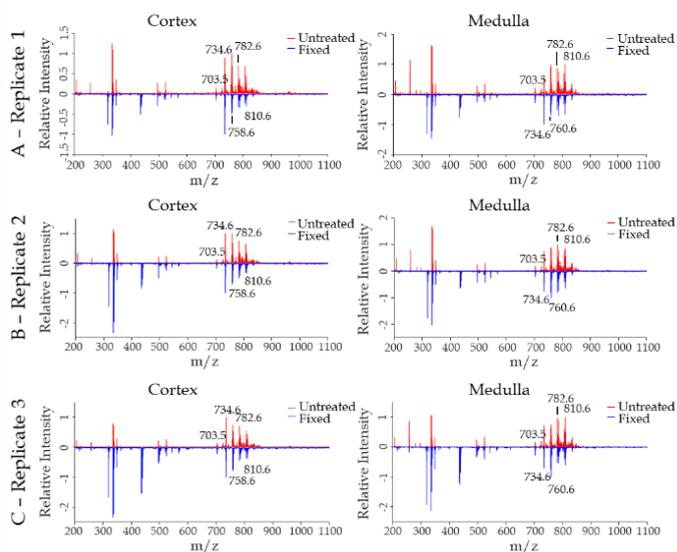
**Figure 1.** Mean cortex and medulla spectral comparison of the generated MALDI-ToF data acquired in negative-ion (A) and positive-ion (B) mode. Data was acquired from untreated ( $n=3$ , top, red trace) and formalin-fixed ( $n=3$ , bottom, blue trace) mouse tissues. Intensities have been normalized to the most abundant lipid-related peak in each spectrum.

Negative-ion mode replicate spectra



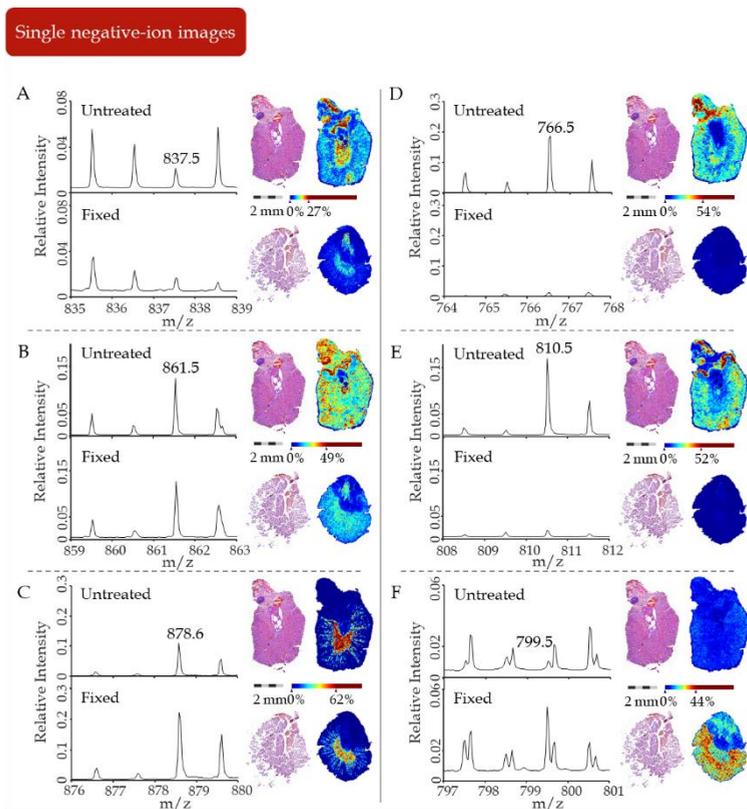
**Figure 2.** Mean region-of-interest spectra acquired in negative-ion mode from the cortex (left) and medulla (right) of each untreated (top, red trace) and formalin-fixed (bottom, blue trace) mouse kidney tissue sections. Intensities have been normalized to the most abundant lipid-related peak in each spectrum.

Positive-ion mode replicate spectra



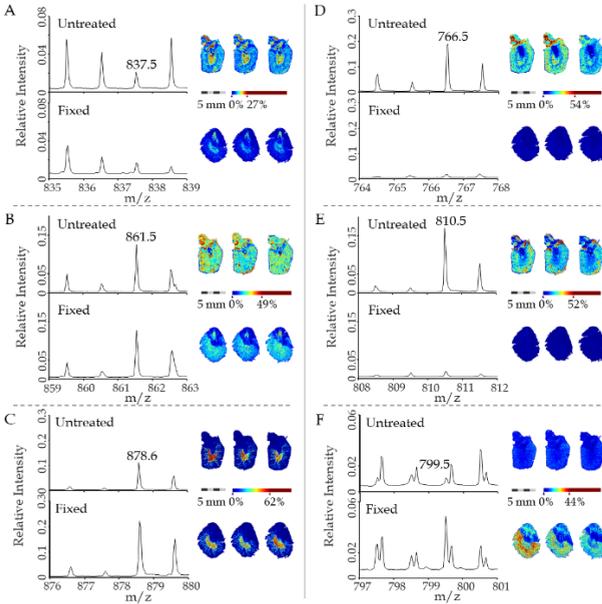
**Figure 3.** Mean region-of-interest spectra acquired in positive-ion mode from the cortex (left) and medulla (right) of each untreated (top, red trace) and formalin-fixed (bottom, blue trace) mouse kidney tissue sections. Intensities have been normalized to the most abundant lipid-related peak in each spectrum.

Figure 4 shows selected ion images and narrow  $m/z$  ranges acquired from a single FF and untreated mouse kidney tissue sections in the negative-ion mode. These data reveal in greater detail the specific depletion of certain lipid signals upon FF. The corresponding mean spectra and MSI data from all replicates are provided in Figure 5 and are consistent with the data shown in Figure 4. Figures 4A-C show example ion distribution images and spectral ranges that show little alteration upon FF. Importantly, and in line with earlier reports<sup>[22, 110, 111]</sup>, consistent ion images are observed from both fixed and unfixed tissues indicating that the solvent-based fixation process does not significantly alter the spatial distributions of lipids detectable in FF tissues. In contrast, Figures 4D and 4E show example ion images and spectral ranges acquired in the negative-ion mode that are significantly depleted upon FF (e.g.,  $m/z$  766.5 and 810.5). Several ions are also observed at higher abundance upon FF, for example,  $m/z$  799.5 shown in Figure 4F, which is observed primarily from the cortex regions of the kidney. Analogous changes are also observed in identically-treated rat kidney tissues (Figures 6 and 7). The potential influence on the obtained spectra by the zinc sulfate present in the fixation solution was also investigated. Spectra acquired from fresh frozen tissue with and without washing in 1% aqueous zinc sulfate solution are provided in Figure 8. Although a strong desalting effect was observed in the washed tissue leading to a reduction of  $[M+Na]^+$  and  $[M+K]^+$  ions, crucially, the detected lipid species and classes were very similar, providing evidence the observed spectral changes are not the result of the zinc sulfate within the fixation solution.



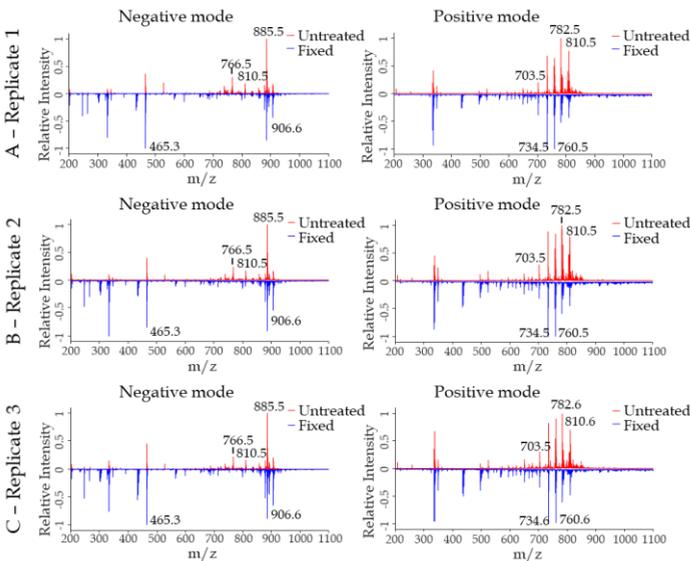
**Figure 4.** Selected  $m/z$  ranges and ion images in the negative-ion mode that were (A-C) not significantly altered upon fixation, (D, E) depleted upon fixation, and (F) enhanced upon fixation of the mouse kidney tissue. Spectra represent the average signal acquired from all three technical replicates while one representative MSI dataset from each fixed and unfixed dataset is shown. All spectra have been normalized to their base peak and all images normalized to the total-ion-current and generated using hotspot removal (quantiles of the image have been calculated between 0-99.0%). MSI data correspond to the labeled peak in each case and the optical image of the post-MSI H&E stained tissue is provided adjacent to the MSI data (note some distortion of the fixed tissues during H&E staining was observed. This distortion was not observed in the optical image of the sample prior to MSI, neither is it observed in the MSI data itself. This post-MSI distortion thus does not influence the acquired spectra data. ROI's are consistent with the optical images acquired prior to MSI where cortex and medulla regions are also visible).

## Negative-ion images replicates



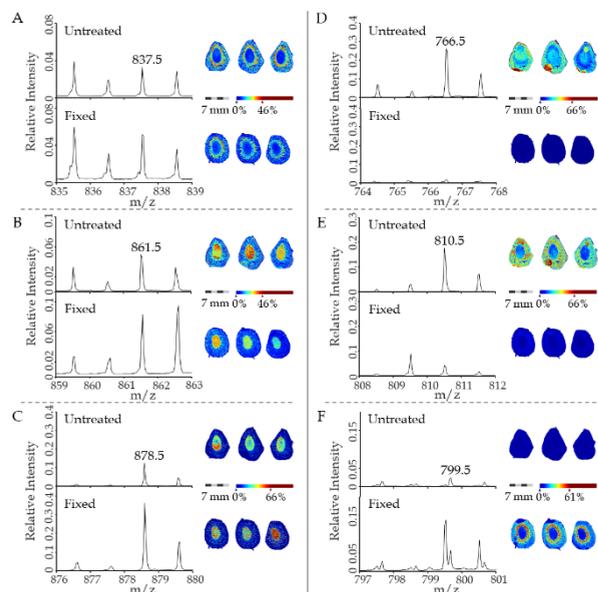
**Figure 5.** Selected  $m/z$  ranges and ion images of all three mouse kidney replicates acquired in the negative-ion mode that were (A-C) not significantly altered upon fixation, (D, E) depleted upon fixation, and (F) enhanced upon fixation. Spectra represent the average signal acquired from all three technical replicates. All spectra have been normalized to the base peak of each corresponding data and all images normalized to the total-ion-current and generated using hotspot removal (quantiles of the image have been calculated between 0-99.0%). MSI data correspond to the labeled peak in each case.

## Replicate spectra rat kidney



**Figure 6.** Mean spectral comparison of the generated MALDI-ToF data acquired in (A) negative-ion and (B) positive-ion mode. Data was acquired from untreated ( $n=3$ , top, red trace) and formalin-fixed ( $n=3$ , bottom, blue trace) rat tissue sections. Intensities have been normalized to the most abundant lipid-related peak in each spectrum and intensities from each tissue averaged across the entire measurement region.

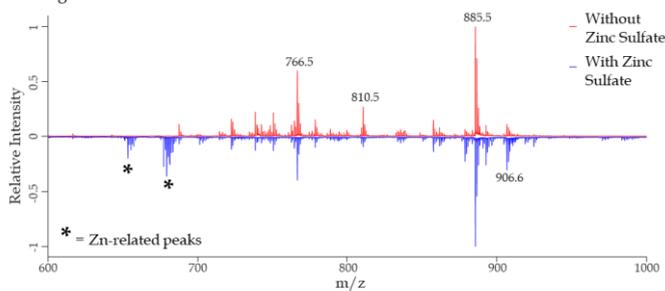
Negative-ion images rat kidney replicates



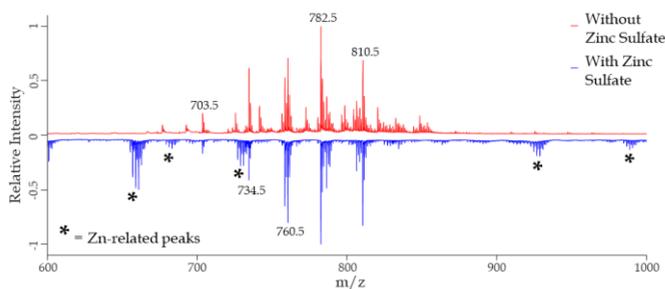
**Figure 7.** Selected  $m/z$  ranges and ion images of all three rat kidney replicates acquired in the negative-ion mode that were (A-C) not significantly altered upon fixation, (D, E) depleted upon fixation, and (F) enhanced upon fixation. Spectra represent the average signal acquired from all three technical replicates. All spectra have been normalized to the base peak of each corresponding data and all images normalized to the total-ion-current and generated using hotspot removal (quantiles of the image have been calculated between 0-99.0%). MSI data correspond to the labeled peak in each case.

Effect of zinc sulfate

A - Negative mode



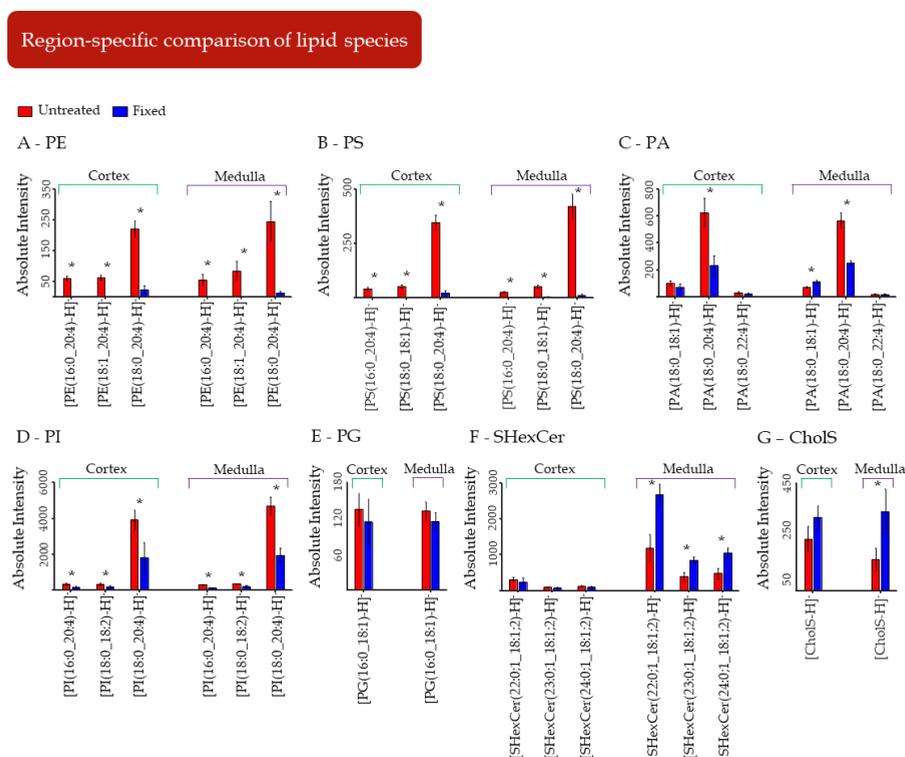
B - Positive mode



**Figure 8.** Mean spectral comparison of the generated MALDI-ToF data acquired in (A) negative-ion and (B) positive-ion mode from fresh frozen tissue unwashed (top, red trace) and washed with aqueous 1% zinc sulfate solution (bottom, blue trace). Peaks originating from the 1% zinc sulfate are marked with an asterisk.

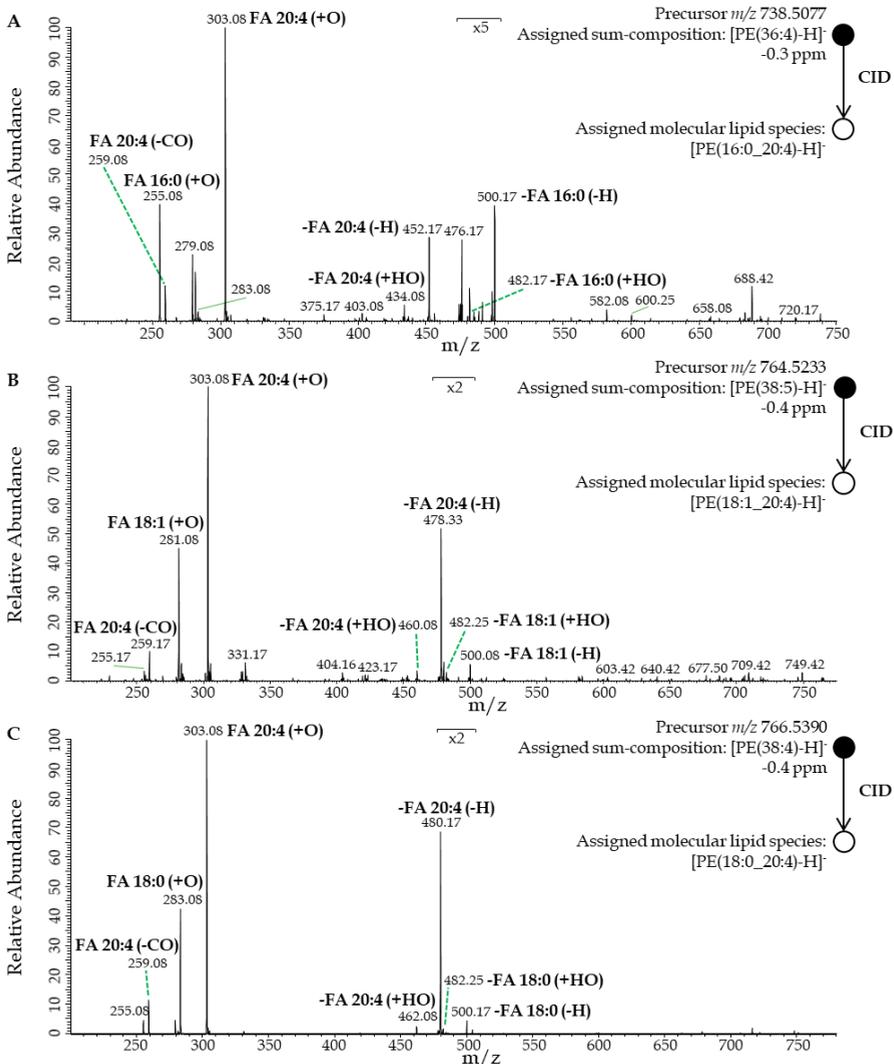
The above results provide evidence of class-specific lipid depletion of ion signals upon FF. To elucidate the specific lipid classes affected by FF we further analyzed both FF and fresh frozen mouse and rat kidney tissues using high mass resolution MALDI-MSI and MS/MS. Figure 9 shows the average ion intensities for a variety of lipid species detected in the cortex and medulla regions of three untreated and FF mouse kidney sections. MS/MS data supporting the molecular-lipid species lipid identifications are provided in Figures 10-15. As these comparisons are made from the same tissue regions comprising the same original composition, the observed differences must arise via FF-induced processes. Figures 9A and 9B show the intensities of several phosphatidylethanolamines (PE) and phosphatidylserines (PS) lipid species. Both PE and PS lipids show a dramatic and statistically significant reduction in signal intensity after FF that is consistent with prior observations using both MSI and LC-MS/MS<sup>[109, 110]</sup> and the known mechanism of formalin fixation. More specifically, formalin reacts with the primary amine groups of the PE and PS headgroups, effectively cross-linking the lipids with other amine-containing lipids and proteins. This crosslinking is hypothesized to inhibit the extraction of free lipids into the matrix and their subsequent ionization by MALDI. Phosphatidic acid (PA) species show a varied response upon fixation (Figure 9C). The ions corresponding to [PA(18:0\_18:1)-H]<sup>-</sup> and [PA(18:0\_22:4)-H]<sup>-</sup> remaining relatively unaffected by fixation. This can be rationalized by the lack of an amine group in PA lipids. In contrast [PA(18:0\_20:4)-H]<sup>-</sup> shows an apparent ~2 fold reduction upon FF. This result is initially surprising due to the lack of an amine group that can undergo crosslinking. However, PA lipids can also form via in-source fragmentation of other phospholipids, most significantly from PS lipids following the loss of the serine headgroup (-87 Da). Thus part of the [PA(18:0\_20:4)-H]<sup>-</sup> ion population likely arises from the abundant [PS(18:0\_20:4)-H]<sup>-</sup> species that is reduced upon FF (Figure 9B). Thus results for PA lipids reflect a convolution of both endogenous PA lipids and those arising from in-source fragmentation of other lipid species. Phosphatidylinositol (PI) lipids showed a ~2-fold reduction after FF (Figure 9D). We speculate this is due to either minor degradation of endogenous lipids during the fixation process and/or partial reaction of the hydroxyl groups of PI with formalin. Phosphatidylglycerol (PG) remains largely unaffected by fixation (Figure 9E). This is in contrast to previous work that has reported both a decrease<sup>[109]</sup> and an increase<sup>[110]</sup> in PG ion intensities after

FF, although we note we only observed one PG species at sufficient intensity for analysis. Interestingly both sulfatides (SHexCer, Figure 9F) and cholesterol sulfate (CholS, Figure 9G) exhibited higher intensities after FF within the the medulla. This observation can possibly be explained by an increase in ionization efficiency, resulting from the decreased amounts of other phospholipids classes that may lead to charge competition and ionization suppression. A similar effect resulting in increased signals for glycosphingolipids following removal of phospholipids using a phospholipase C digestion protocol has been reported previously by Vens-Cappell et al.[123]



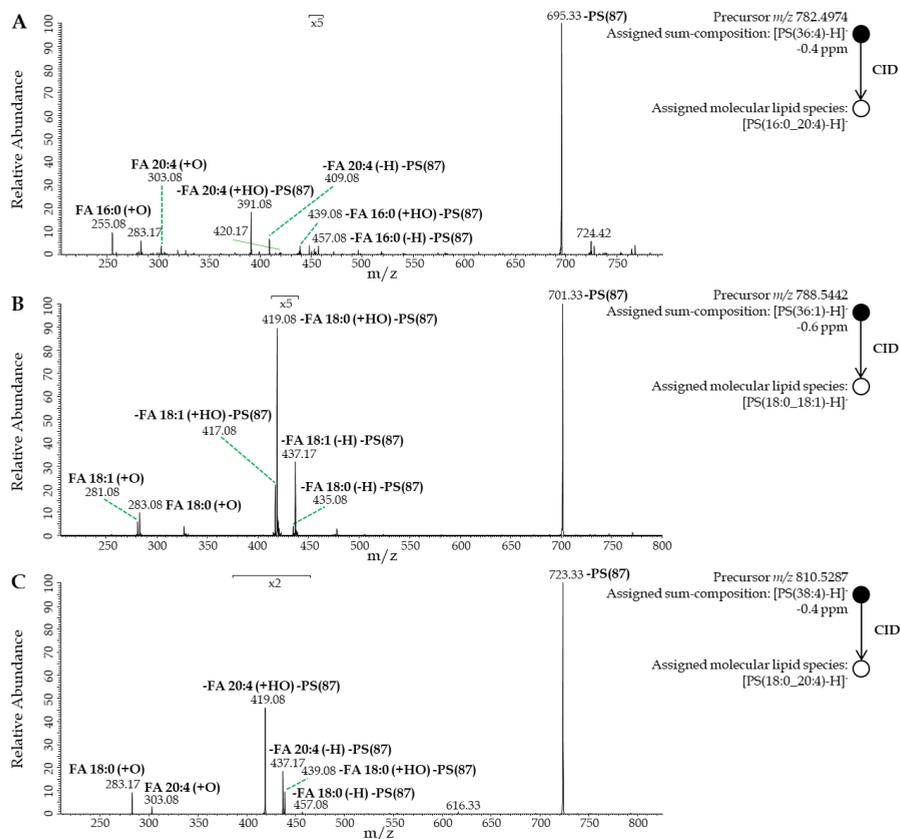
**Figure 9.** Barplot comparison of the absolute ion intensities acquired from the cortex and medulla of fixed and unfixed mouse kidney tissue sections ( $n=3$  each) for selected lipid species representative of seven different lipid classes. (A) phosphatidylethanolamine (PE), (B) phosphatidylserine (PS), (C) phosphatidic acid (PA), (D) phosphatidylinositol (PI), (E) phosphatidylglycerol (PG), (F) sulfatide (SHexCer), and (G) cholesterol sulfate (CholS). Significant changes ( $P < 0.05$ ) between the untreated and the fixed tissues are indicated using an asterisk. Error bars represent  $\pm 1$  standard deviation.

## Identification PE lipids



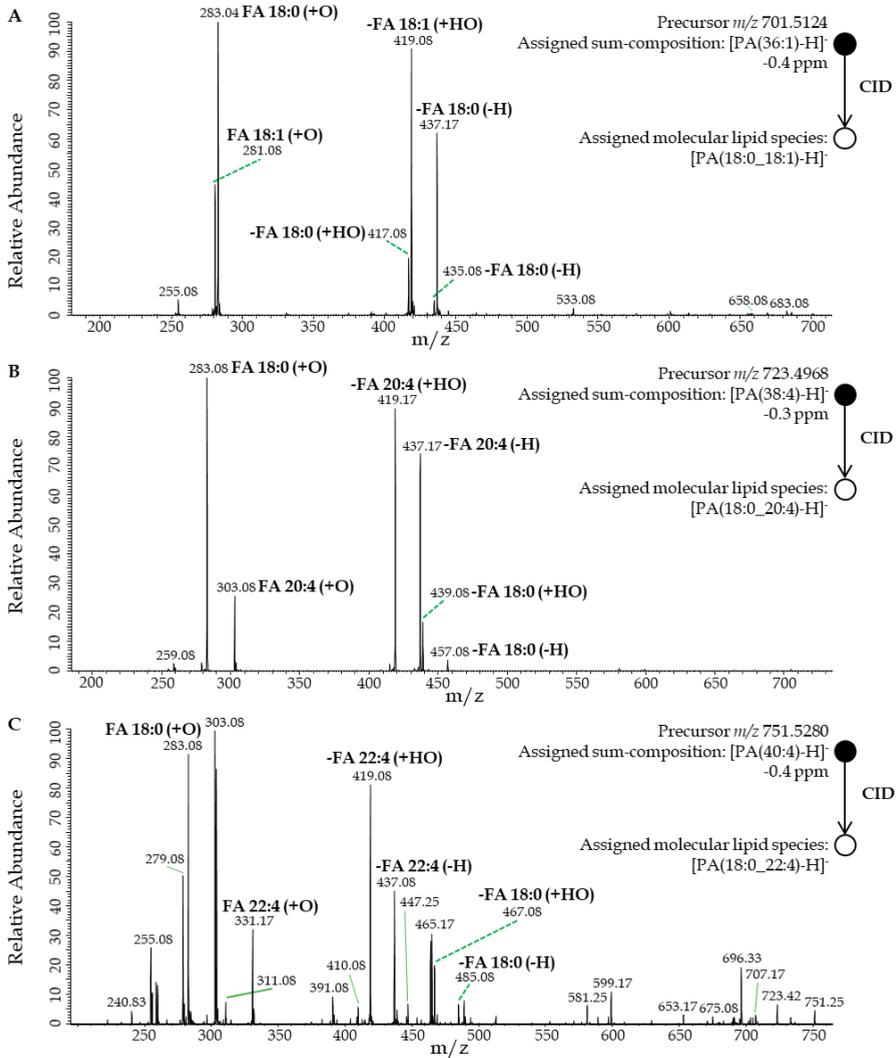
**Figure 10.** Ion trap MS/MS spectra of deprotonated PE lipid species acquired from an untreated mouse kidney tissue section. Lipids are identified as (A) [PE(16:0\_20:4)-H]<sup>-</sup>, (B) [PE(18:1\_20:4)-H]<sup>-</sup>, and (C) [PE(18:0\_20:4)-H]<sup>-</sup>. Full-scan MS<sup>1</sup> data acquired in the Orbitrap was used to determine the precursor ion  $m/z$  and assign sum-composition formula based on the high mass accuracy measurement. The MS/MS spectra acquired in parallel in the ion trap using resonance CID was then used to identify the dominant molecular lipid species. It is likely additional isobaric and isomeric lipids are also co-isolated and fragmented giving rise to additional peaks in the MS/MS spectra.

## Identification PS lipids



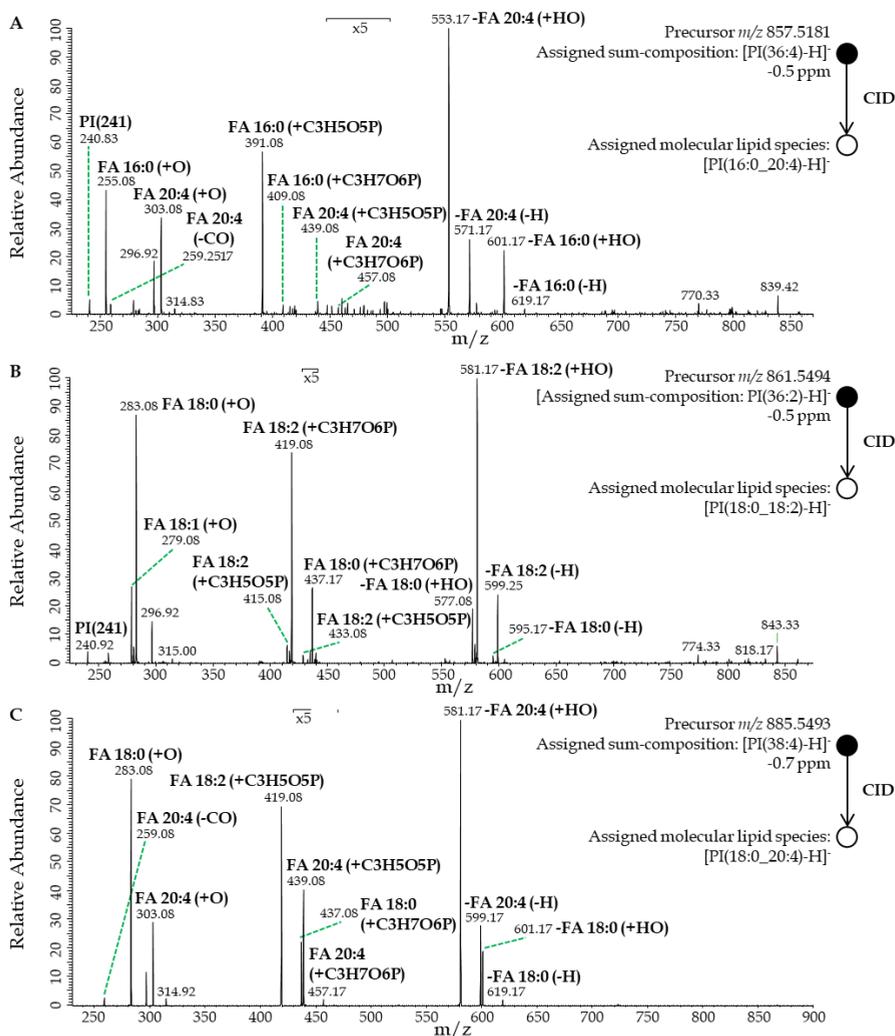
**Figure 11.** Ion trap MS/MS spectra of deprotonated PS lipid species acquired from an untreated mouse kidney tissue section. Lipids are identified as (A) [PS(16:0\_20:4)-H], (B) [PS(18:0\_18:1)-H], and (C) [PS(18:0\_20:4)-H]. Full-scan MS<sup>1</sup> data acquired in the Orbitrap was used to determine the precursor ion  $m/z$  and assign sum-composition formula based on the high mass accuracy measurement. The MS/MS spectra acquired in parallel in the ion trap using resonance CID was then used to identify the dominant molecular lipid species. It is likely additional isobaric and isomeric lipids are also co-isolated and fragmented giving rise to additional peaks in the MS/MS spectra.

## Identification PA lipids



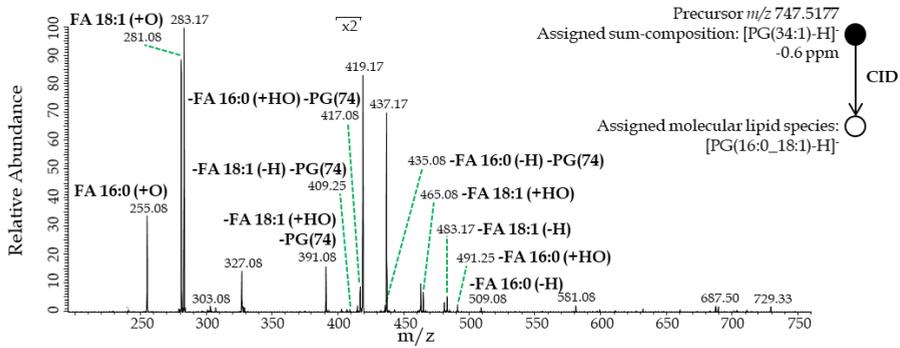
**Figure 12.** Ion trap MS/MS spectra of deprotonated PA lipid species acquired from an untreated mouse kidney tissue section. Lipids are identified as (A) [PA(18:0\_18:1)-H]<sup>-</sup>; (B) [PA(18:0\_20:4)-H]<sup>-</sup>; and (C) [PA(18:0\_22:4)-H]<sup>-</sup>. Full-scan MS<sup>1</sup> data acquired in the Orbitrap was used to determine the precursor ion  $m/z$  and assign sum-composition formula based on the high mass accuracy measurement. The MS/MS spectra acquired in parallel in the ion trap using resonance CID was then used to identify the dominant molecular lipid species. It is likely additional isobaric and isomeric lipids are also co-isolated and fragmented giving rise to additional peaks in the MS/MS spectra.

## Identification PI lipids



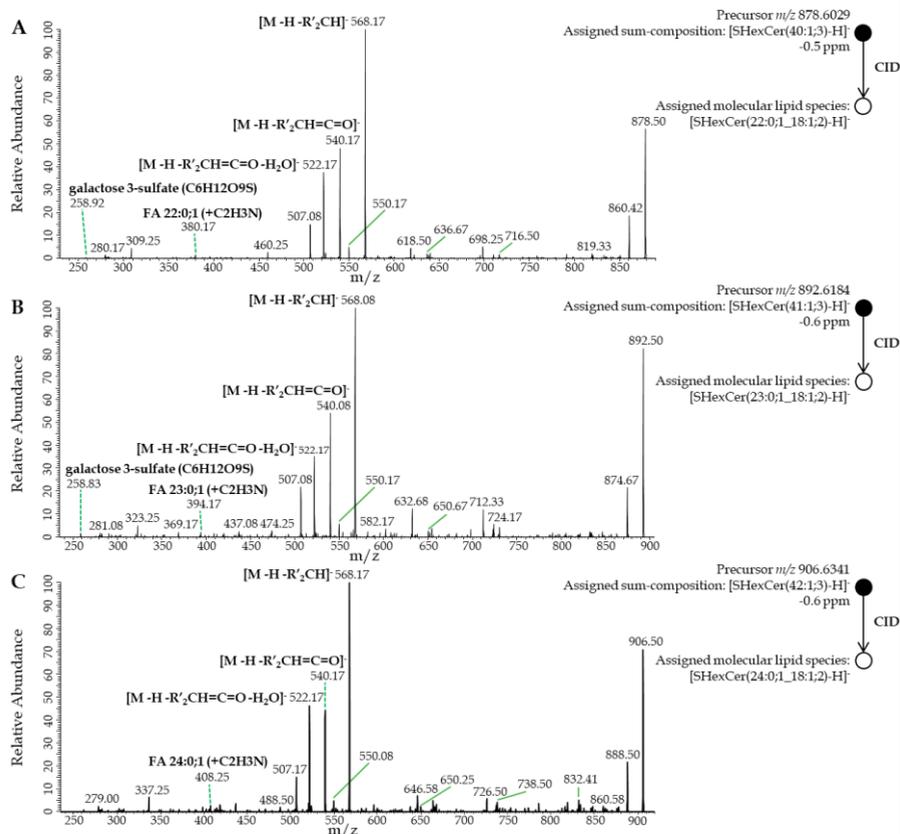
**Figure 13.** Ion trap MS/MS spectra of deprotonated PI lipid species acquired from an untreated mouse kidney tissue section. Lipids are identified as (A) [PI(16:0\_20:4)-H], (B) [PI(18:0\_18:2)-H], and (C) [PI(18:0\_20:4)-H]. Full-scan MS<sup>1</sup> data acquired in the Orbitrap was used to determine the precursor ion  $m/z$  and assign sum-composition formula based on the high mass accuracy measurement. The MS/MS spectra acquired in parallel in the ion trap using resonance CID was then used to identify the dominant molecular lipid species. It is likely additional isobaric and isomeric lipids are also co-isolated and fragmented giving rise to additional peaks in the MS/MS spectra.

## Identification PG lipid



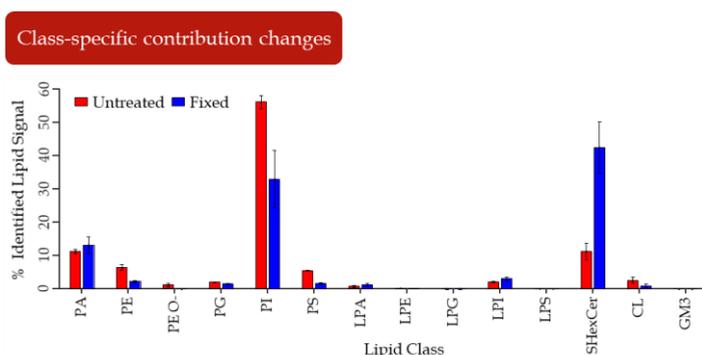
**Figure 14.** Ion trap MS/MS spectra of a deprotonated PG lipid species acquired from an untreated mouse kidney tissue section. Lipid is identified as [PG(16:0\_18:1)-H]<sup>-</sup>. Full-scan MS<sup>1</sup> data acquired in the Orbitrap was used to determine the precursor ion  $m/z$  and assign sum-composition formula based on the high mass accuracy measurement. The MS/MS spectra acquired in parallel in the ion trap using resonance CID was then used to identify the dominant molecular lipid species. It is likely additional isobaric and isomeric lipids are also co-isolated and fragmented giving rise to additional peaks in the MS/MS spectra.

## Identification SHexCer lipids



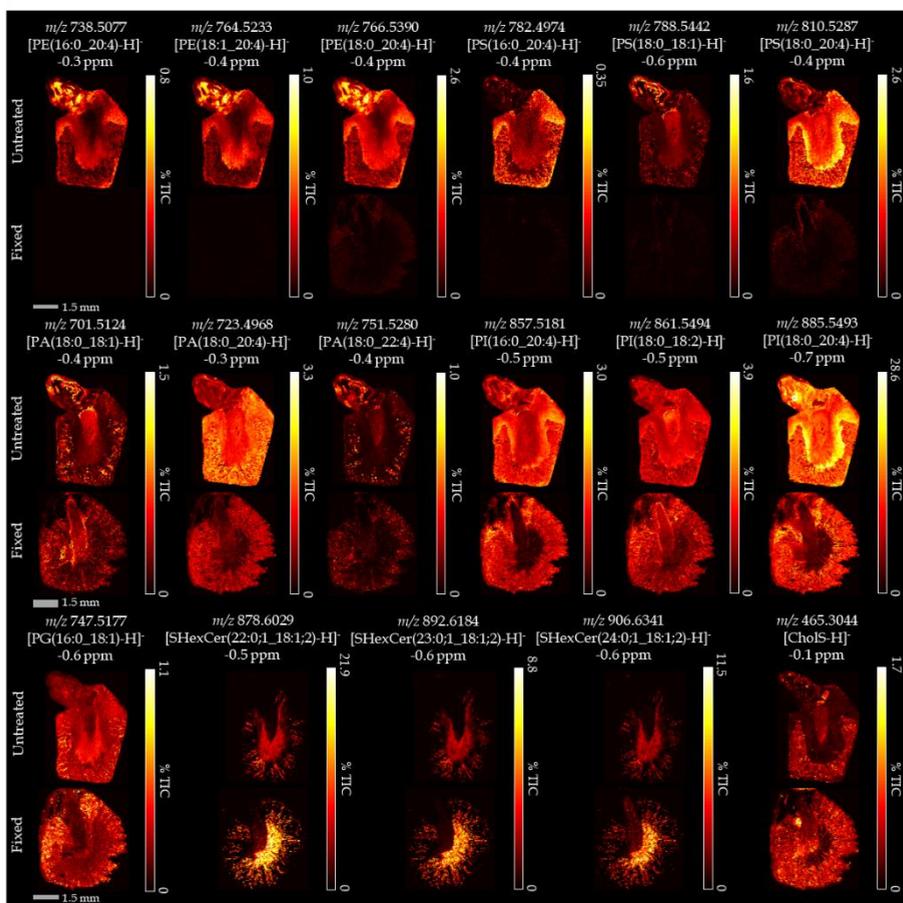
**Figure 15.** Ion trap MS/MS spectra of deprotonated sulfatide lipid species acquired from an untreated mouse kidney tissue section. Lipids are identified as (A) [SHexCer(22:0;1<sub>18:1;2</sub>)-H], (B) [SHexCer(23:0;1<sub>18:1;2</sub>)-H], and (C) [SHexCer(24:0;1<sub>18:1;2</sub>)-H]. Full-scan MS<sup>1</sup> data acquired in the Orbitrap was used to determine the precursor ion  $m/z$  and assign sum-composition formula based on the high mass accuracy measurement. The MS/MS spectra acquired in parallel in the ion trap using resonance CID was then used to identify the dominant molecular lipid species. It is likely additional isobaric and isomeric lipids are also co-isolated and fragmented giving rise to additional peaks in the MS/MS spectra.

From a broader lipidomics perspective, Figure 16 shows the changes in relative contributions to the total identified sum-composition lipid signal (cf., “Lipid Identification” section in the Methods) obtained from both fixed and unfixed mouse kidney tissues in the negative-ion mode MS1 data (i.e., lipids are assigned based on accurate  $m/z$  values). The full list of identified sum-composition lipid species is provided in Supporting Information Table S1 (<https://doi.org/10.1016/j.ijms.2019.116212>). Consistent with the above, Figure 16 shows an increased contribution from sulfatides and a reduced contribution from a variety of lipid species, in particular PI, PE, and PS species. We note that the relative decrease in PE species appears to be less than that shown in Figure 9. This is due to the MS<sup>1</sup>-based identification routine incorrectly assigning several ions originating from the in-source fragmentation of PC species to their corresponding [M-CH<sub>3</sub>]<sup>-</sup> ions as isomeric deprotonated PE lipids. For example, based only on accurate mass  $m/z$  718.5386 is assigned as [PE(34:0)-H]<sup>-</sup>. However, MS/MS data reveals an almost exclusive presence of 16:0-related fragments, suggesting this ion is instead [PC(32:0)-CH<sub>3</sub>]<sup>-</sup> (data not shown). Figure 17 shows the corresponding ion distribution images for all species discussed above in Figure 9 and reveals consistent ion distributions for all lipid species observed in both untreated and FF tissues. For interested readers the complete negative-ion mode datasets from each mouse kidney replicate can be accessed via the METASPACE online annotation platform (<https://metaspace2020.eu/>).<sup>[124]</sup> Similar class-dependent depletion of lipid signals was also obtained in rat kidney tissue sections (Figures 18 and 19).



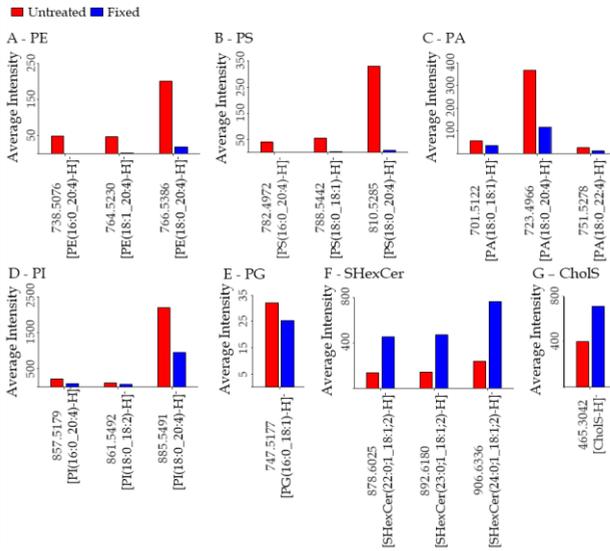
**Figure 16.** Barplot showing the changes in the contribution of lipid classes to the total identified lipid signal (cf., “Lipid Identification” section in the Methods) from fixed and unfixed mouse kidney tissue ( $n=3$  each). Error bars represent  $\pm 1$  standard deviation.

## High mass resolution ion images



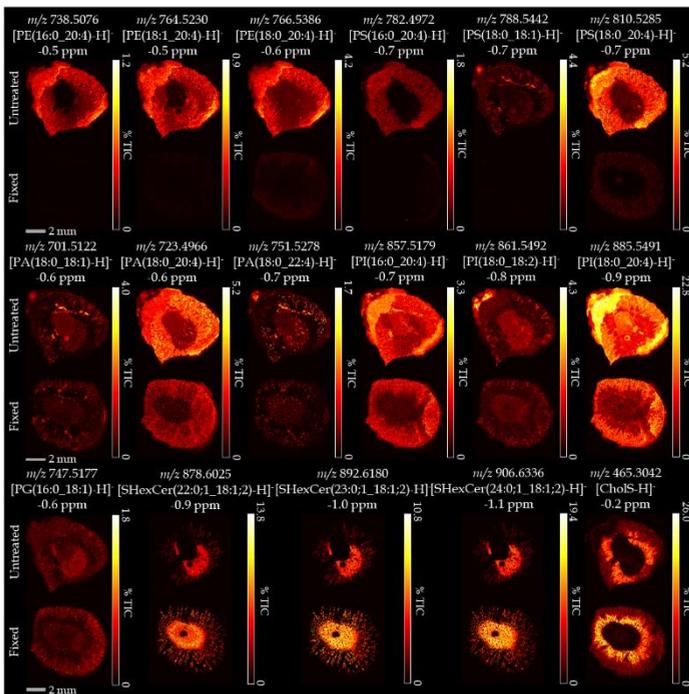
**Figure 17.** MALDI-Orbitrap ion distribution images for the lipid species shown in Figure 9. Each image has been normalized per-pixel to the total ion count and plotted against the maximum normalized intensity observed from the untreated (fresh frozen) tissues, with the exception of SHexCer and CholS images that are plotted against the maximum normalized intensity observed from the FF tissues. Confirmatory MS/MS spectra are provided in Figures 10-15.

## Region-specific comparison of rat kidney lipid species



**Figure 18.** Barplot comparison of the absolute ion intensities acquired from fixed and unfixed rat kidney tissue sections ( $n=1$ ) for selected lipid species representative of seven different lipid classes. Intensities are the average across the entire tissue region. (A) phosphatidylethanolamine (PE), (B) phosphatidylserine (PS), (C) phosphatidic acid (PA), (D) phosphatidylinositol (PI), (E) phosphatidylglycerol (PG), (F) sulfatide (SHexCer), and (G) cholesterol sulfate (CholS).

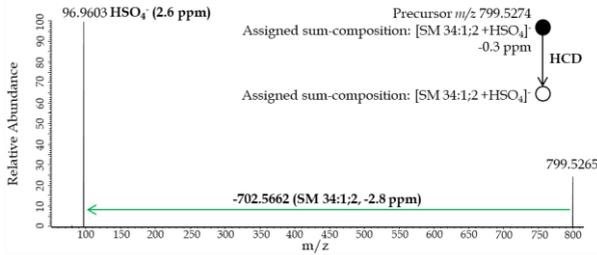
## High mass resolution rat kidney ion images



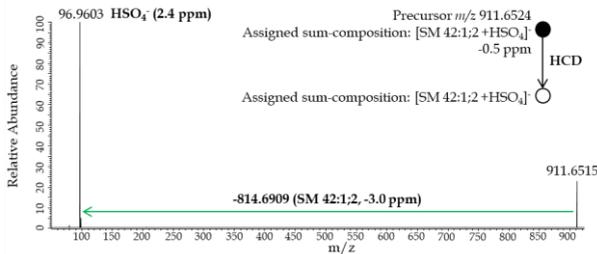
**Figure 19.** MALDI-Orbitrap ion distribution images for the lipid species shown in Figure 18 from both fixed and untreated rat kidney tissue. Each image has been normalized per-pixel to the total ion count and plotted against the maximum normalized intensity observed for the untreated (fresh frozen) tissues, with the exception of SHexCer and CholS images that are plotted against the maximum normalized intensity observed for the FF tissues.

An interesting observation in negative-ion mode was the presence of the  $[M+HSO_4]^-$  adducts for sphingomyelin (SM) in the fixed mouse and rat kidney tissue. From mouse kidney tissue, these ions were observed in the Orbitrap data at  $m/z$  799.5274 and  $m/z$  911.6524 ( $m/z$  799.5 is also shown in Figures 1 and 4). The measured  $m/z$  values provided in the high-resolution Orbitrap analysis yielded no logical matches in either the LIPIDMAPS database or the ALEX<sup>123</sup> lipid calculator. However, HCD-MS/MS spectra for these ions (Figures 20-23) revealed a sole fragment ion at  $m/z$  96.9603 for both precursors which was assigned to the  $HSO_4^-$  anion (mass errors  $\leq 2.5$  ppm). This anion formed following neutral losses of 702.5662 and 814.6909 from the precursor ions at  $m/z$  799.5274 and 911.6524, respectively. These neutral losses correspond to the neutral masses of the sphingomyelin (SM) species SM 34:1;2 and SM 42:1;2 (-2.8 ppm and -3.0 ppm), respectively, providing strong evidence for the assignment of these ions as  $[M+HSO_4]^-$  ions of these SM lipids. The source of the hydrogen sulfate is likely from the zinc sulfate-containing formalin solution. The ability to detect SM lipids in the negative-ion mode, in addition to the more acidic sphingolipids such as SHexCers, can allow the direct correlation of their distributions within a single dataset that could serendipitously prove beneficial in the study of sphingolipid metabolism.

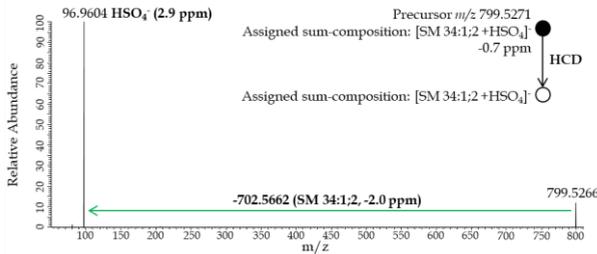
Taken together, the above results rationalize the observations in the initial MALDI-ToF data shown in Figures 1 and 4. In Figure 4 the presence of ions at  $m/z$  837.5, 861.5, and  $m/z$  878.5 that are not massively depleted upon FF is observed. Using the DDA-imaging data these ions are identified as  $[PI(16:0_{18:1})-H]^-$ ,  $[PI(18:0_{18:2})-H]^-$ , and  $[SHexCer(22:0;1_{18:1;2})]^-$ , respectively and are lipid classes that do not contain an amine group and are thus expected to be tolerant to FF. On the other hand, the ions observed at  $m/z$  766.5 and  $m/z$  810.5 (Figures 4D and 4E) are identified as  $[PE(18:0_{20:4})-H]^-$  and  $[PS(18:0_{20:4})-H]^-$  ions, respectively. These both correspond to amine-contacting lipids and their signal depletion is consistent with the known mechanisms of fixation. Finally, as described above, the ion detected at  $m/z$  799.5 in Figure 4 is assigned as a  $[M+HSO_4]^-$  ion of SM 34:1;2 and is an artifact of the sulfate-containing formalin solution.

Identification SM lipid  $m/z$  799,5

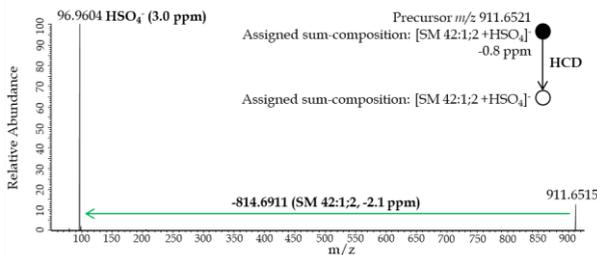
**Figure 20.** Identification of  $[\text{SM}(34:1;2)+\text{HSO}_4]^-$  in the negative-ion mode analysis of formalin-fixed mouse kidney tissue. HCD-MS/MS of the  $m/z$  799.5265 precursor ion demonstrating the formation of an  $\text{HSO}_4^-$  fragment corresponding to the neutral loss of  $\text{SM}(34:1;2)$ .

Identification SM lipid  $m/z$  911,7

**Figure 21.** Identification of  $[\text{SM}(42:1;2)+\text{HSO}_4]^-$  in the negative-ion mode analysis of formalin-fixed mouse kidney tissue. HCD-MS/MS of the  $m/z$  911.6515 precursor ion demonstrating the formation of an  $\text{HSO}_4^-$  fragment corresponding to the neutral loss of  $\text{SM}(42:1;2)$ .

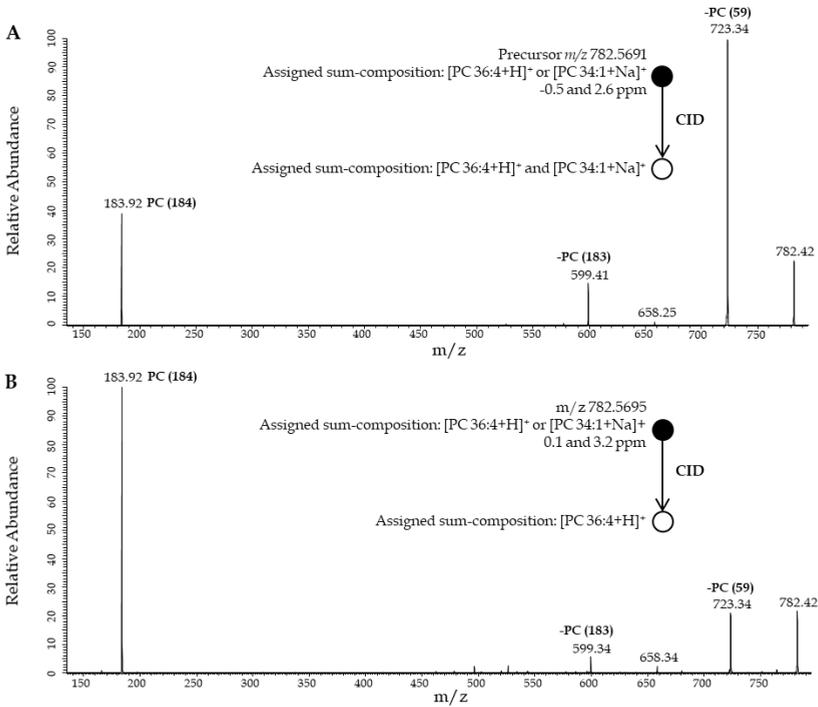
Identification rat kidney SM lipid  $m/z$  799,5

**Figure 22.** Identification of  $[\text{SM}(34:1;2)+\text{HSO}_4]^-$  in the negative-ion mode analysis of formalin-fixed rat kidney tissue. HCD-MS/MS of the  $m/z$  799.5266 precursor ion demonstrating the formation of an  $\text{HSO}_4^-$  fragment corresponding to the neutral loss of  $\text{SM}(34:1;2)$ .

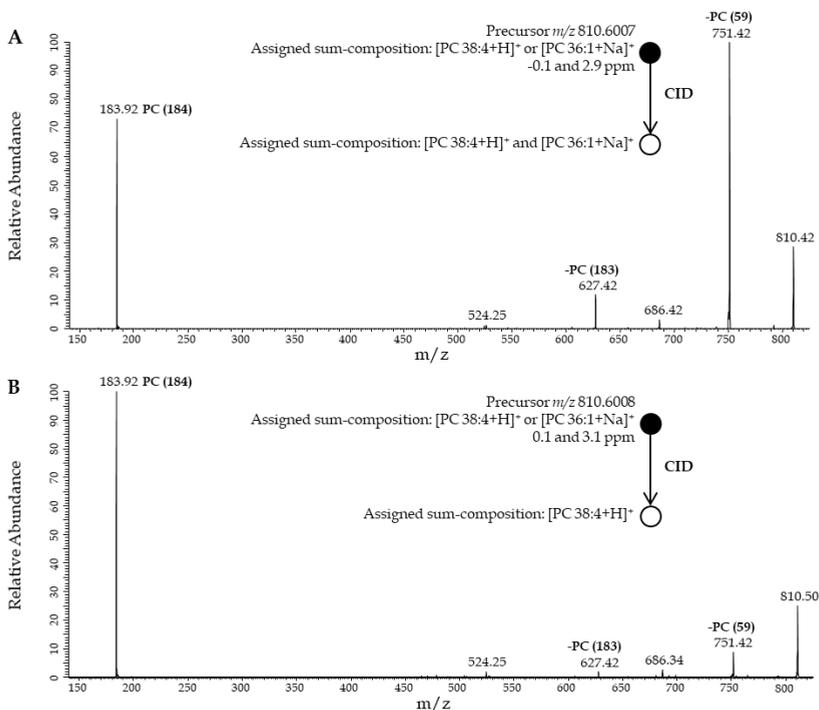
Identification rat kidney SM lipid  $m/z$  911,7

**Figure 23.** Identification of  $[\text{SM}(42:1;2)+\text{HSO}_4]^-$  in the negative-ion mode analysis of formalin-fixed rat kidney tissue. HCD-MS/MS of the  $m/z$  911.6515 precursor ion demonstrating the formation of an  $\text{HSO}_4^-$  fragment corresponding to the neutral loss of  $\text{SM}(42:1;2)$ .

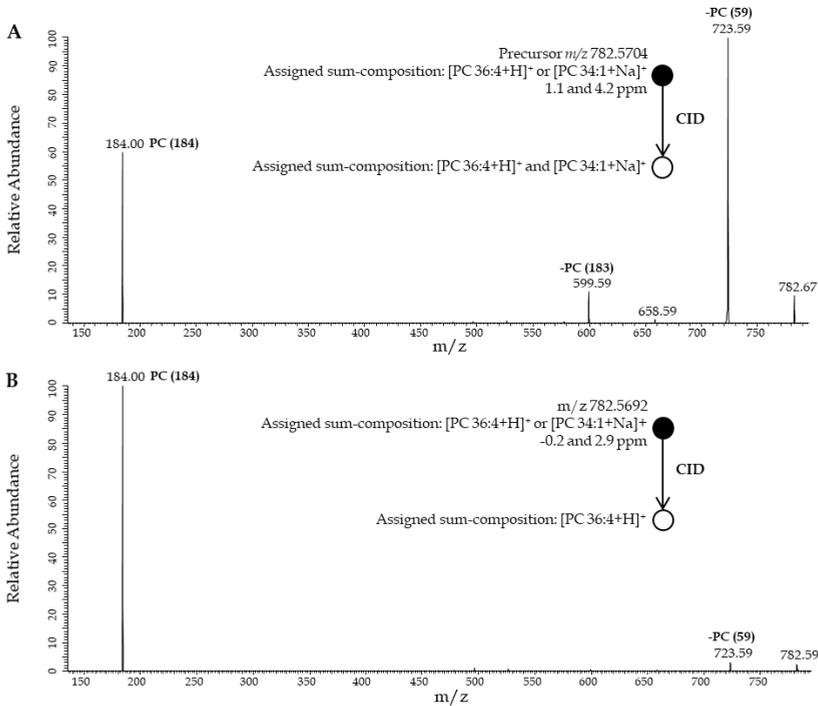
In the positive-ion mode, the combined high mass accuracy and MS/MS measurement enabled the differences shown in Figure 1B to be identified. As expected positive-ion mode spectra were dominated by PC and SM lipids in both FF and untreated tissues. Although at first look the spectra appear similar, being dominated by abundant  $[M+H]^+$  ion in both cases, an increase in alkali adducts, most notably  $[M+Na]^+$  ions, is also observed in the untreated tissues. The shift towards protonated ions in the fixed mouse and rat tissue is attributed to the aqueous fixation solution which effectively removes endogenous sodium and potassium from the tissue. This effect is analogous to studies utilizing aqueous washing to desalt tissues prior to MSI<sup>[125, 126]</sup> and is consistent with work demonstrating a shift in adduct formation of PC lipids before and after fixation, with the specific shift depending on the composition of the formalin fixation solution.<sup>[110, 111]</sup> However,  $[M+Na]^+$  and  $[M+H]^+$  adducts of different PC lipid species can remain unresolved even with the high mass resolving power realized here. To demonstrate the shift towards protonated species in the FF tissue we have analyzed the CID-MS/MS spectra acquired from precursor ions of  $m/z$  782.57 and  $m/z$  810.60 in both positive-ion mode datasets. These two ion populations represent isobaric combinations of  $[PC\ 36:4\ +H]^+ / [PC\ 34:1\ +Na]^+$  and  $[PC\ 38:4\ +H]^+ / [PC\ 36:1\ +Na]^+$  respectively. In the untreated tissue, both the 59 and 183 neutral loss fragments are observed at high abundance and are attributed to the losses of  $N(CH_3)_3$  (-59 Da) and phosphocholine (-183 Da) respectively (Figures 24A, 25A (from mouse tissue) and 26A, 27A (from rat tissue)). These fragment ions are known to form to a far greater extent for alkali adducted PC lipids compared to their protonated variants<sup>[127]</sup>, thereby providing evidence for a significant population of  $[M+Na]^+$  ions. By contrast, in the fixed tissue the  $m/z$  184 fragment is formed almost exclusively and is the main expected fragment from protonated PC lipids (Figures 24B, 25B (from mouse tissue) and 26B, 27B (from rat tissue)). These data demonstrate that FF does not significantly influence the detection of PC and SM lipids in positive-ion mode but simply results in a shift in adduct profile. In this case, this results in the almost exclusive formation of protonated PC lipids in the fixed tissue.

Identification PC lipid  $m/z$  782.6

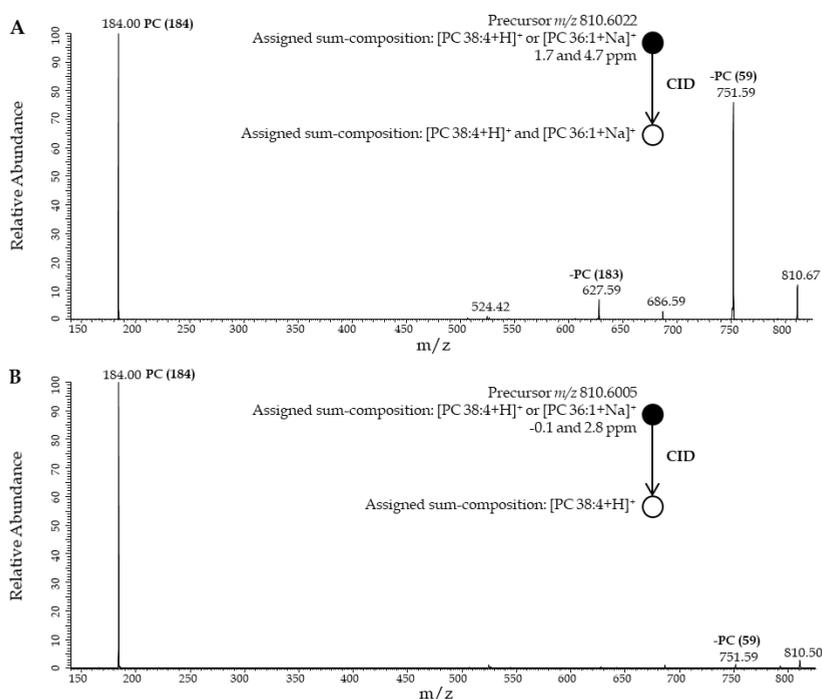
**Figure 24.** Ion trap MS/MS spectra of a selected PC-related ion observed at  $m/z$  782.57 in the positive-ion mode acquired from (A) untreated mouse kidney tissue section and (B) fixed mouse kidney tissue section. Full-scan MS<sup>1</sup> data acquired in the Orbitrap was used to determine the precursor ion  $m/z$  and assign sum-composition formula based on the high mass accuracy measurement. The MS/MS spectra acquired in parallel in the ion trap using resonance CID was then used to identify the dominant molecular lipid species. In the untreated tissue fragments suggesting the presence of both [PC(34:1)+Na]<sup>+</sup> ( $m/z$  723.3 and  $m/z$  599.4) and [PC(36:4)+H]<sup>+</sup> ( $m/z$  184) are observed. In the fixed tissue a far greater contribution from [PC(36:4)+H]<sup>+</sup>-related fragments is observed which is attributed to the removal of salts from the tissue during the fixation process.

Identification PC lipid  $m/z$  810.6

**Figure 25.** Ion trap MS/MS spectra of a selected PC-related ion observed at  $m/z$  810.60 in the positive-ion mode acquired from (A) untreated mouse kidney tissue section and (B) fixed mouse kidney tissue section. Full-scan MS<sup>1</sup> data acquired in the Orbitrap was used to determine the precursor ion  $m/z$  and assign sum-composition formula based on the high mass accuracy measurement. The MS/MS spectra acquired in parallel in the ion trap using resonance CID was then used to identify the dominant molecular lipid species. In the untreated tissue fragments suggesting the presence of both [PC(36:1)+Na]<sup>+</sup> ( $m/z$  751.4 and  $m/z$  627.4) and [PC(38:4)+H]<sup>+</sup> ( $m/z$  184) are observed. In the fixed tissue a far greater contribution from [PC(38:4)+H]<sup>+</sup>-related fragments is observed which is attributed to the removal of salts from the tissue during the fixation process.

Identification rat kidney PC lipid  $m/z$  782.6

**Figure 26.** Ion trap MS/MS spectra of a selected PC-related ion observed at  $m/z$  782.57 in the positive-ion mode acquired from (A) untreated rat kidney tissue section and (B) fixed rat kidney tissue section. Full-scan MS<sup>1</sup> data acquired in the Orbitrap was used to determine the precursor ion  $m/z$  and assign sum-composition formula based on the high mass accuracy measurement. The MS/MS spectra acquired in parallel in the ion trap using resonance CID was then used to identify the dominant molecular lipid species. In the untreated tissue fragments suggesting the presence of both [PC(34:1)+Na]<sup>+</sup> ( $m/z$  723.6 and  $m/z$  599.6) and [PC(36:4)+H]<sup>+</sup> ( $m/z$  184) are observed. In the fixed tissue a far greater contribution from [PC(36:4)+H]<sup>+</sup>-related fragments is observed which is attributed to the removal of salts from the tissue during the fixation process.

Identification rat kidney PC lipid  $m/z$  810.6

**Figure 27.** Ion trap MS/MS spectra of a selected PC-related ion observed at  $m/z$  810.60 in the positive-ion mode acquired from (A) untreated rat kidney tissue section and (B) fixed rat kidney tissue section. Full-scan MS<sup>1</sup> data acquired in the Orbitrap was used to determine the precursor ion  $m/z$  and assign sum-composition formula based on the high mass accuracy measurement. The MS/MS spectra acquired in parallel in the ion trap using resonance CID was then used to identify the dominant molecular lipid species. In the untreated tissue fragments suggesting the presence of both [PC(36:1)+Na]<sup>+</sup> ( $m/z$  751.6 and  $m/z$  627.6) and [PC(38:4)+H]<sup>+</sup> ( $m/z$  184) are observed. In the fixed tissue a far greater contribution from [PC(38:4)+H]<sup>+</sup>-related fragments is observed which is attributed to the removal of salts from the tissue during the fixation process.

## Conclusion

Using MALDI-MSI combined with accurate mass measurements and MS/MS we have systematically investigated the effect of formalin fixation on lipids in kidney tissue, a widely used model tissue for MSI studies. In line with the known crosslinking mechanism of formalin, our results show a significant depletion of ion signals from primary amine-containing phospholipids, namely PE and PS that is consistent with earlier reports.<sup>[109, 110]</sup> We did not observe potential intermediate reaction products of the fixation process such as *N*-formylation *N*-methylation as previously reported for LC-MS<sup>[109]</sup> or cross-linked lipid species. The latter observation can be rationalized by the generation of a cross-linked network of lipids and proteins. Although the depleted lipids only contain a single amine group they may still form an interconnected covalent network with other cross-linked species that is not efficiently extracted into the matrix, ionized, and that exceeds the accessible mass range. It is noteworthy that despite the solvent-based fixation procedure high-quality MSI data can still be acquired from fixed tissues for many different classes with minimal delocalization compared to unfixed tissues (e.g., Figures 4 and 17). The potential recovery of cross-linked lipids by optimized antigen retrieval steps may provide an exciting future avenue of research to increase the lipid coverage achievable when analyzing FF tissues.

FF is not used in isolation, but routinely in combination with paraffin embedding (FFPE tissues). Given the recent demonstration of lipid detection from clinically-relevant FFPE tissues using MALDI-MSI and the potential diagnostic power of such analyses<sup>[22]</sup> our results can help rationalize the detected lipid species in these applications (independent of potential losses arising from the paraffin embedding process itself) and provide insight into the biochemical conclusions that can, and cannot, be drawn in these studies.

## Acknowledgments

This work has been made possible with the financial support of the LINK program of the Dutch province of Limburg and has been funded by ITEA

and RVO by means of project numbers ITEA151003/ITEA 14001. A.P.B, R.M.A.H, and S.R.E acknowledge financial support from Interreg V EMR and the Netherlands Ministry of Economic Affairs within the “EURLIPIDS” project (project number EMR23). We are grateful to the Glunde lab in the Department of Radiology at Johns Hopkins, Baltimore, for providing mouse kidney tissue.

# ***Chapter Three***

## **Strategies for managing multi-patient 3D mass spectrometry imaging data**

**D. R. Naomi Vos**, Ilaria Jansen, Marit Lucas, Martin R. L. Paine, Onne J. de Boer, Sybren L. Meijer, C Dilara Savci-Heijink, Henk A Marqeurig, D. Martijn de Bruin, Ron M. A. Heeren, Shane R. Ellis, Benjamin Balluff  
*Journal of Proteomics* **2019** 193, 184-191



MSI has become a powerful tool in biomedical research to investigate the distribution of a broad scale of compounds ranging from metabolites to proteins as well as drugs in diseased tissues, such as malignant tumors. MSI is most commonly applied in a 2D fashion but in recent years 3D-MSI has been more applied in the field to gain a deeper understanding of biological processes underlying diseases.

These 3D applications are potentially introducing a sampling bias on a sample or patient level, respectively. As more tissues are imaged per patient to obtain a 3D volume, and this amount is multiplied per patient included in the study, not only a large dataset is obtained but it requires also more sample preparation and measurements. This increase in sample handling can lead to sampling bias and discrepancies due to insufficient reproducibility.

In this chapter, we aim to investigate the consequences of sampling bias on sample representativeness. We, therefore, analyzed fourteen bladder cancer patients with varying histological grades with 3D MALDI-MSI by digesting 20 sections per patient with trypsin to obtain peptides. To obtain reproducible results we incorporated a quality control in the sample preparation. This quality control was used in one of the two novel outlier detection routines we developed for 3D-MSI data to evaluate digestion efficacy. The other outlier detection was based on z-directed regression and both outlier detections combined led to 20% of the data being classified as outliers and were subsequently removed.

After removing the data it was found that 33% of a sample has to be measured to obtain sufficient coverage of the existing biological variance within a tissue sample. Furthermore, we investigated what the consequence of sampling bias is on the precision of biomarker discovery for histological grading of human bladder cancers by MSI.

## Introduction

Mass spectrometry imaging (MSI) has emerged as a powerful tool in biomedical research to reveal the localization of a broad scale of compounds ranging from metabolites to proteins in tissues in a label-free manner.<sup>[113]</sup> Especially in cancer research, MSI enables scientists and pathologists to look at how molecules are distributed both between tumors and within tumors in order to shed light on tumor biology or find potential biomarkers for e.g. tumor staging.<sup>[128]</sup> So far, MSI is most often deployed as a two-dimensional (2D) imaging technique. However, this can pose limitations when studying typically heterogeneous tissues since a complete overview of the three-dimensional context of the tissue is not retained in a single section. With recent improvements in acquisition speed and data analysis, three-dimensional (3D) MSI is coming to the forefront whereby data from multiple consecutive 2D sections are combined into a 3D volume.<sup>[129]</sup>

The first report on 3D-MSI was by Crecelius et al.<sup>[56]</sup>, in which they visualized the distribution of myelin basic protein (MBP) in a mouse brain. In later years, articles have been published where neuropeptides and lipids were 3D visualized in a Crustacean brain<sup>[80]</sup> and neuropeptides and proteins in a mouse brain.<sup>[43]</sup> Furthermore, 3D-MSI has been used to visualize cancer drug distribution in tumors, which allowed the authors to confirm the relation of the structural heterogeneity of the microenvironment with the drug distribution inside tumors.<sup>[69]</sup> Patterson et al. also used 3D-MSI to study the molecular distribution of lipids in atherosclerotic plaques.<sup>[93]</sup> This paper clearly shows the added value of 3D over 2D, since the authors found several lipids that correlated with the degree of artery stenosis, the volumetric narrowing of the artery.

All these examples show the potential of 3D-MSI for biomedical research. However, all 3D-MSI studies to-date have so far only investigated a single fresh frozen tissue given the heavy experimental workload of 3D-MSI studies. This has two drawbacks: first, fresh frozen specimens are not readily available in a clinical setting; and second, the results from a single tissue lack representativeness for a group of patients in biomedical research.

Tissues that are available for this purpose are formalin-fixed paraffin-embedded (FFPE), which are stored in large biobanks worldwide. Even though MSI sample preparation methods exist for FFPE tissues<sup>[130-132]</sup>, they have not yet found usage for 3D-MSI. The main reason for this is the experimental workload and the challenge to guarantee reproducibility over multiple sections from the same sample and between different samples.<sup>[84]</sup>

In this work, we report on the usage of matrix-assisted laser desorption/ionization (MALDI)-MSI for the acquisition of 3D-MSI datasets from fourteen human bladder cancer resection specimens provided as FFPE samples. Bladder cancer is a common cancer type, especially amongst men.<sup>[133]</sup> In bladder cancer, the distinction between high- and low-grade urothelial carcinoma is important since high-grade bladder cancer has a high chance of recurrence and progression.<sup>[134]</sup> 2D-MSI can assist in differentiating these grades, as previously reported by Oezdemir et al. in 2012 by measuring proteins from fresh frozen tissues.<sup>[135]</sup> Taking into consideration tumor heterogeneity in bladder cancer, as in many cancers<sup>[23]</sup>, this method may fall short of precise representation of the disease state. Particularly, to assess the extent of the disease which mainly includes the level of invasion, if any, makes multi-site sampling or 3D analyses of the samples imperative for a comprehensive characterization of the tumor.<sup>[136]</sup>

In our study, we used FFPE tissue of bladder cancer resections to obtain molecular information on a peptide level by *in situ* proteolytic digestion on the tissues' surface which is widely used in clinical MSI studies.<sup>[20, 137, 138]</sup> A key aspect of our sample preparation protocol was the implementation of several quality controls designed to monitor the reproducibility over all datasets. The obtained data allowed an extensive investigation of the nature of 3D-MSI datasets including outlier detection, the effect of sampling bias, and the minimum set of representative sections. The presented study is the first report on 3D-MSI of tryptic peptides from multiple FFPE tissues and proposes a comprehensive data analysis workflow for this type of data.

## Experimental

### Patient material

FFPE bladder cancer resection specimens of fourteen patients were analyzed in this study. Per patient, 40 consecutive sections with a thickness of five  $\mu\text{m}$  were sampled of which every second section was provided on indium tin oxide-coated conductive (ITO) glass slides (Delta Technologies, Loveland, USA) by the Amsterdam Medical Center (Amsterdam, Netherlands) with two sections per slide. The section size varied per patient (Figure 1A) and was on average 24.29  $\text{mm}^2$  (Table 1). The patients gave their informed consent and the study was approved by the Institutional Review Board of the Amsterdam Medical Center (AMC), Amsterdam (W16\_088 # 16.108). Histological grading was available for seven subjects with four low-grade and three high-grade tumors.

Dataset	Minimum ( $\text{mm}^2$ )	Maximum ( $\text{mm}^2$ )	Mean ( $\text{mm}^2$ )
1	10.57	11.55	11.03
2	18.98	22.87	21.46
3	97.63	105.79	102.01
4	35.88	40.31	38.54
5	5.52	7.10	6.69
6	46.08	53.10	60.00
7	35.49	39.09	37.28
8	9.04	10.37	9.78
9	3.36	3.70	3.48
10	18.91	23.12	21.36
11	4.80	5.27	5.05
12	6.21	7.57	6.62
13	17.25	24.93	20.67
14	4.71	5.38	5.07

**Table 1.** Overview of section sizes per dataset. The minimum and maximum section sizes are given for each dataset as well as the mean of each dataset (20 sections).

## Sample preparation

All uneven-numbered sections of a patient were collected for hematoxylin and eosin (H&E) staining, followed by digitalization in a high-resolution slide scanner (Ultra Fast Scanner, Philips, Best, The Netherlands) and histological annotation. All even-numbered sections were used for 3D-MSI.

The sections were de-paraffinized with 3 xylene washes (5, 10, and 10 min), subsequently rinsed 3 times with HPLC-grade ethanol (each 2 min) followed by 2 UHPLC-grade water washes of 5 min (all solutions are from Biosolve Chimie SARL, Dieuze, France). To undo the crosslinking induced by formalin fixation, antigen retrieval was performed with citric acid at 10 mM and pH 6.0 (Sigma-Aldrich, Zwijndrecht, The Netherlands) at 121°C for 20 minutes using the Antigen Retriever 2100 (Aptum Biologics, UK). This was followed by applying 15 layers of 20 µg/mL water-dissolved porcine trypsin (Sigma-Aldrich) using a SunCollect pneumatic sprayer (SunChrom GmbH, Germany) with a constant 10 µL/min flowrate, track spacing of 1 mm, nozzle height of 25 mm, and a nozzle speed of 900 mm/min. Samples were then incubated at 37°C for 17 hours in an airtight box containing 100 mL of 1:1 water:methanol. Afterward, 7 layers of 5 mg/mL alpha-cyano-4-hydroxycinnamic acid (Sigma-Aldrich) in 50% acetonitrile (Biosolve Chimie SARL, Dieuze, France) and 0.2% trifluoroacetic acid (Sigma-Aldrich, Zwijndrecht, The Netherlands) was applied using the same SunCollect sprayer with increasing flowrates of 10, 20, 30, and 40 µL/min for the first four layers, and 40 µL/min for the remaining layers with a track spacing of 2 mm, nozzle height of 25 mm, and a nozzle speed of 1390 mm/min before performing MSI experiments.

On all ITO slides, 1 µL of 1 mg/mL intact cytochrome *c* (from equine heart, Sigma-Aldrich) was deposited manually before trypsin application to evaluate the efficacy of the digestion.

## MALDI-MSI measurement

All sections were imaged at 50 µm raster size on a rapiflex MALDI-ToF/ToF instrument<sup>[15]</sup> (Bruker Daltonik GmbH, Bremen, Germany) across an *m/z*

range of 800-3000 in positive-ion reflectron mode. The laser was focused for each tissue separately and scanned across a 50x50  $\mu\text{m}^2$  area with 1000 laser shots summed for each position. Laser intensity was adjusted off tissue to obtain similar intensities for the matrix across all measurements ( $\sim 4 \times 10^4$ ). Typical mass resolution was on average 20,000 for a single pixel (55,200 data points per spectrum). During acquisition the data was processed with centroid peak detection ( $S/N \geq 5$ , 0.5 Da peak width, 78% height), Savitzky-Golay smoothing (width of 0.01  $m/z$ , and 1 cycle), and TopHat baseline subtraction. The instrument was calibrated before each measurement by using red phosphorus as a calibration agent. At the end of acquisition, a fast-access-data-file was automatically created by resampling the spectra to 80% of the original size. The reduction led to a decrease in mass resolution to on average 14,000.

## Data analysis

In total, fourteen bladder cancer specimens each consisting of 20 sections spaced 10  $\mu\text{m}$  apart were measured (Figure 1A). The reduced MSI data were imported into SCiLS Lab 2016a (SCiLS, Bremen, Germany) without baseline removal for further analysis. In SCiLS Lab, the MSI data was also co-registered with the corresponding H&E stains that were previously annotated by trained pathologists (Figure 1C). All spectra were total-ion-count (TIC) normalized for all subsequent steps. Peak picking was performed on the exported overall average spectrum (Figure 1B) in mMass<sup>[117]</sup> (35 precision baseline correction, Gaussian smoothing with a window size of 0.5 Da and 2 cycles, deisotoping with an isotope mass tolerance of 0.1 Da and intensity tolerance of 50%, and at last peaks were picked with an  $S/N$  threshold of 3 using the complete height of the peaks) and matrix and trypsin related peaks were removed by mass matching, with 0.2 Da tolerance, a spectrum that has been acquired off-tissue. With this peak list, consisting of 298 masses, the maximum intensity in each peak interval (200 ppm) was exported for every patient and section in tabular form from SCiLS Lab for subsequent analysis in the R statistical environment (version 3.4.0).

Outliers were determined by employing two methods: the first uses the cytochrome *c* spots that were applied on each slide to monitor digest efficacy. Principal Component Analysis (PCA) was performed on the mean spectra from the cytochrome *c* spots of each dataset and a 95% error ellipse was calculated in the first two-component space. Any cytochrome *c* measurements that were outside the 95% error ellipse are considered outliers and the two sections belonging to this slide were taken out. The second outlier method is based on the assumption that the z-directed distribution of the intensity of a molecule has a certain consistent gradient that can be modeled and tested by regression (Figures 2C and 2F). A quadratic function (linear and higher-order polynomials showed lower Akaike information indices) is fitted for each *m/z* ( $x = z$ -position;  $y = \text{intensity}$ ) and the section that has the maximum absolute residual value is recorded. A section is considered an outlier if the residual frequency is higher than the 99% confidence interval ( $=2.58$  standard deviations from the mean) of the expected by-chance mean frequency (Formula 1):

$$\text{Formula 1: } \textit{residual frequency threshold} = \frac{\# \textit{ of peptide masses}}{\# \textit{ of sections}} + 2.58 * \frac{SD(\# \textit{ of residual masses per section})}{\sqrt{\# \textit{ of sections}}}$$

By combining these two methods a list of outlier sections is obtained that were not considered for further data analysis.

Assessment of the spatial correlation of MSI data with histological annotations was done using both an unsupervised and a supervised approach. For the unsupervised method, probabilistic Latent Semantic Analysis (pLSA) was applied in SCiLS Lab on all pixels of all datasets using ten components. The pLSA component coefficients for all pixels were then exported to R where they were investigated for their discriminatory power between tumor and non-tumor regions using the Receiver Operating Characteristic (ROC) analysis. The optimal cut-off for a pLSA component was derived from the ROC curve using Youden's index, a summary measure for the maximum potential effectiveness of a biomarker.<sup>[139]</sup>

For the supervised analysis, a linear discriminant analysis was carried out in SCiLS Lab using the six patients for whom both clinical data as well as annotations were available. A leave-one-out cross-validation was used to assess the accuracy of the model. This model was then applied to annotate previously not-annotated tumor regions within a seventh patient of whom the clinical data was also available.

All subsequent statistical analyses were performed on the log-transformed data in order to approximately conform it to normal distribution.

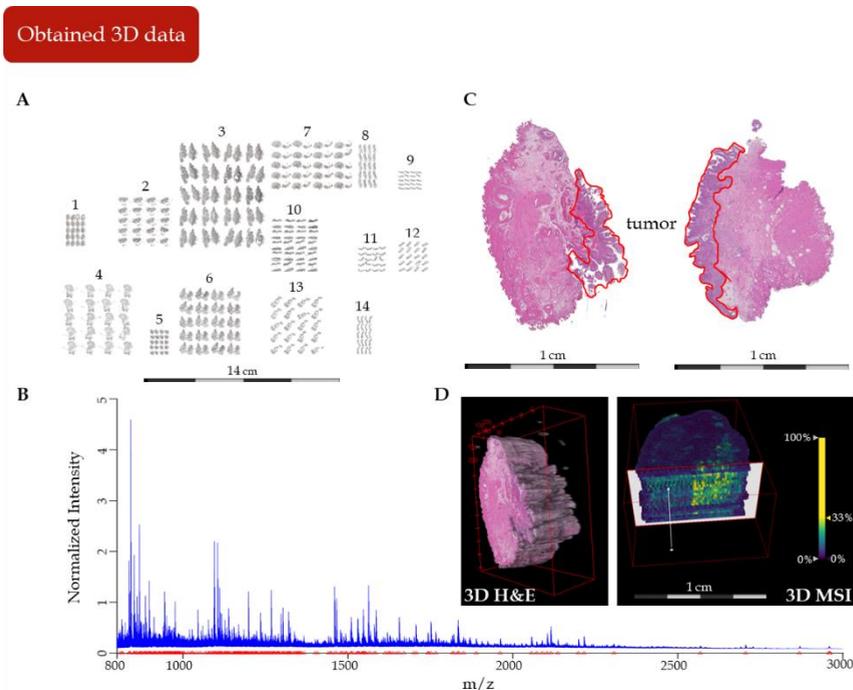
To investigate sampling bias for the detection of low- or high-grade specific peptides in tumor regions when only one section is considered per patient, a section is chosen randomly for each patient followed by statistical testing using a Student *t*-test for significant differences between both grades. Masses that had a *p*-value  $\leq 0.05$  were considered statistically significant. This was repeated 1000 times. A mass was considered statistically significant if its frequency was higher than 6.2% (a binomial test gave that threshold where the true probability of success was  $> 5\%$ ). To monitor also the directionality of the significant difference, the non-log scaled fold-changes in intensity were recorded as well.

To test the number of sections needed to cover the molecular information of a full dataset, the average spectrum of a complete dataset was determined first. Then, a defined number of sections were randomly selected 100 times, their average spectrum was calculated, and the Pearson correlation coefficient of the resulting spectrum to the full sample's average spectrum was determined. It was investigated how the correlation changed with an increasing number of randomly selected sections. The relevance of the sampling position was investigated by calculating the average distance between the selected sections.

Three-dimensional visualizations of MSI datasets and H&E stacks were created in SCiLS Lab and ImageJ, respectively (Figure 1D).

## Results and Discussion

Within this study, 20 consecutive sections have been measured for fourteen bladder cancer specimens by mass spectrometry imaging (MSI) with the aim to obtain a more representative molecular description of each patient (Figure 1A). The samples were provided as FFPE tissues and the molecular class of interest was the proteome that requires the use of antigen retrieval and proteolytic digestion, in order to obtain a peptide spectrum (Figure 1B). The efficacies, and therefore reproducibility, of these two sample preparation steps are still challenging to monitor.



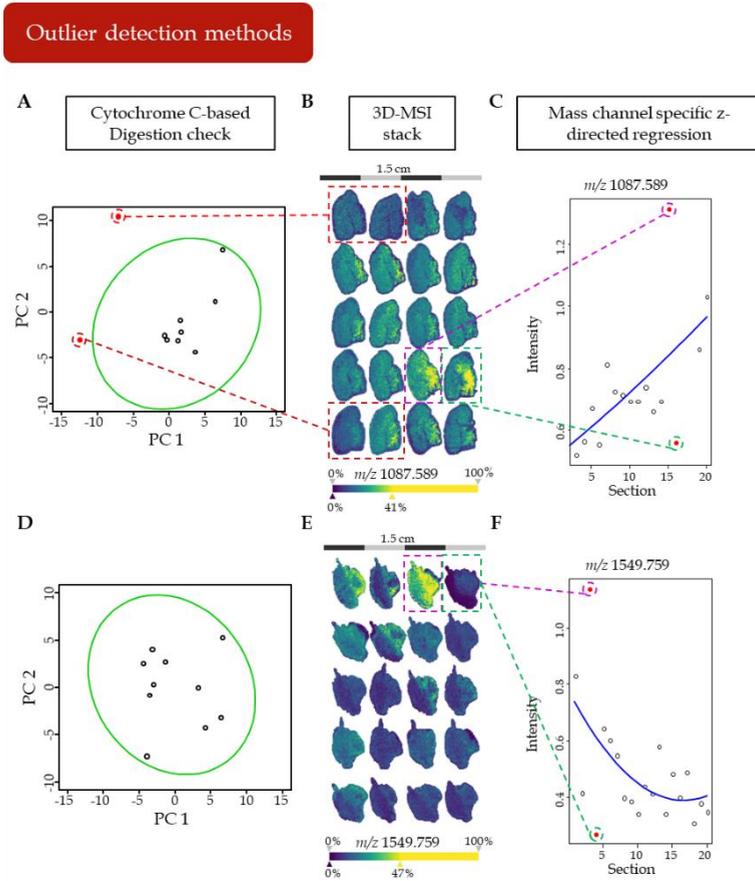
**Figure 1.** Overview of study data. MSI was used to measure 20 evenly spaced consecutive sections of 14 bladder cancer patients, labeled 1-14 (A). The average peptide spectrum across all datasets is shown in (B) with red triangles indicating the 298 peak-picked masses. Annotated H&E stains for dataset 1 and 5 are shown in (C). Three-dimensional visualizations of the H&E stains and the MSI-data of the same dataset are shown in (D).

## Outliers and reproducibility of sample preparation

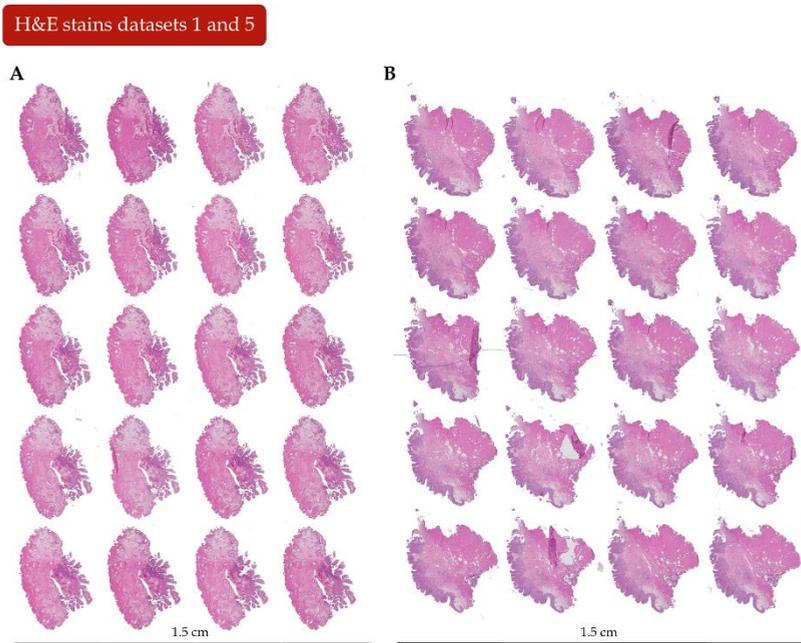
In this study, we developed two methods for outlier detection: one based on evaluating digestion efficacy and the second on checking data consistency within 3D-MSI datasets. The first criterion, proposed by Erich et al.<sup>[84]</sup>, uses a spot of intact cytochrome *c* being applied on each slide before digestion. We extend this idea by applying a principal component analysis on the resulting spectra to identify outliers as data points that lay outside the 95% error ellipse in the first two principal components space (Figures 2A and 2D). Applying this criterion on all datasets separately declared 11 slides, and hence a total of 22 sections out of the 280 sections as outliers – a dropout rate of < 8%.

The second criterion is based on the assumption that each molecular distribution follows a consistent gradient in the *z*-direction and can be modeled mathematically by regression. Consequently, sections whose mass signals consistently produce high deviations from that gradient can be considered outliers (Formula 1 and Figures 2C and 2F). For these calculations, the mean spectrum of the whole section was used from the datasets that had not previously been removed by the cytochrome *c* quality control. When applying this criterion, a total of 34 sections were classified as outliers from the remaining 258 sections, giving a reproducibility rate of 87%. An example that shows the complementarity of both approaches is given in Figures 2A and 2C, where sections were labeled as outliers from whom it was known that they adhered during antigen retrieval to the back of another slide. This would not have been detectable by the cytochrome *c* check only.

As stated before, these two criteria were combined to cover all the possible technical incidences. After cytochrome *c* (22 outlier sections) and regression (34) filtering, this resulted in a total of 56 outlier sections out of the 280 sections, and therefore in a reproducibility rate of 80%. Given that the samples are FFPE and therefore require more extensive sample preparation, we consider these dropout rates acceptable thereby allowing further analyses of the data.



**Figure 2.** Outlier detection strategies for two example datasets (patient 1 in B and patient 5 in E, corresponding H&E stains can be found in Figures 1C and 3). The outlier detection was based on evaluating digest efficacy (A and D) and a residual detection in a z-directed regression (C and F). The cytochrome *c* based outlier detection relies on the principal component analysis (PCA) of the digest spectra of cytochrome *c*, which was spotted on every slide before proteolytic digestion. Slides having scores in the first two-component PCA space outside the 95% error interval are considered outliers, shown as red dashed lines (A). Single sections are considered outliers if a critical number of mass channels show a deviating behavior from a quadratic regression function, shown as pink and green dashed lines (C and F).



**Figure 3.** H&E stains from dataset 1 (A) and dataset 5 (B) from which data is shown in Figures 2 and 5.

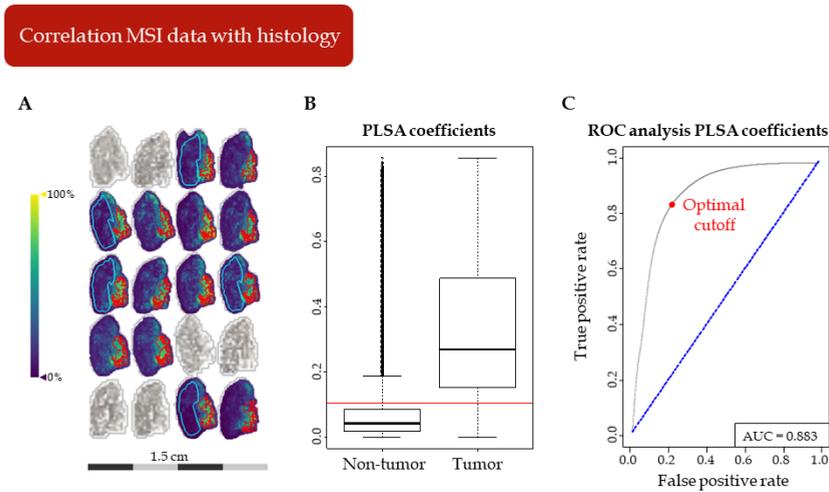
### **Spatial correlation of MSI data with histology**

The first question is related to the spatial correlation of the MSI data with the histology to ensure that any molecular signature for the tumor is indeed from the tumor. To this end, both unsupervised and supervised multivariate analyses were performed.

For the unsupervised approach, probabilistic Latent Semantic Analysis (pLSA) was performed with ten expected components. This analysis was performed on all datasets simultaneously, except for the previously eliminated outliers. To quantify the specificity of these components for representing tumor areas, five non-tumor regions were defined manually for each dataset as negative controls (Figure 4A, the dataset shown is one with the most outliers due to sample preparation problems. Corresponding H&E's can be found in Figures 1C and 3).

The pLSA data for all annotated regions was then exported to R where the discriminatory power of the components to discern tumor from non-tumor regions was investigated by a ROC analysis (Figures 4B and 4C). The area-under-the-curve of component 10 with a value of 0.883 was found to exhibit the highest discriminatory power. Youden's index was then used to determine the optimal cut-off for the pLSA coefficients, here 0.1, resulted in a sensitivity of 84%, a specificity of 79%, and overall accuracy of 79% (Figure 4C). This could help to determine whether there is a tumor present in a sample when annotations are not available.

The same purpose was pursued with the supervised analysis, especially since from the fourteen datasets measured there are only six where annotated H&E images were available. These six datasets were used to train and validate a linear discriminant analysis model to distinguish between tumor and non-tumor tissue. Leave-one-out cross-validation estimated the mean model's accuracy to be 88%, 9 percentage points more accurate than the unsupervised approach, with a sensitivity of 94% and a specificity of 68%. Similarly to Inglese et al.<sup>[42]</sup>, such classifiers can be extremely useful to annotate 3D-MSI datasets on a per-pixel basis when annotations are not available for all H&E images. The supervised classifier was used to annotate the tumor in these sections that are then incorporated in all subsequent analyses.

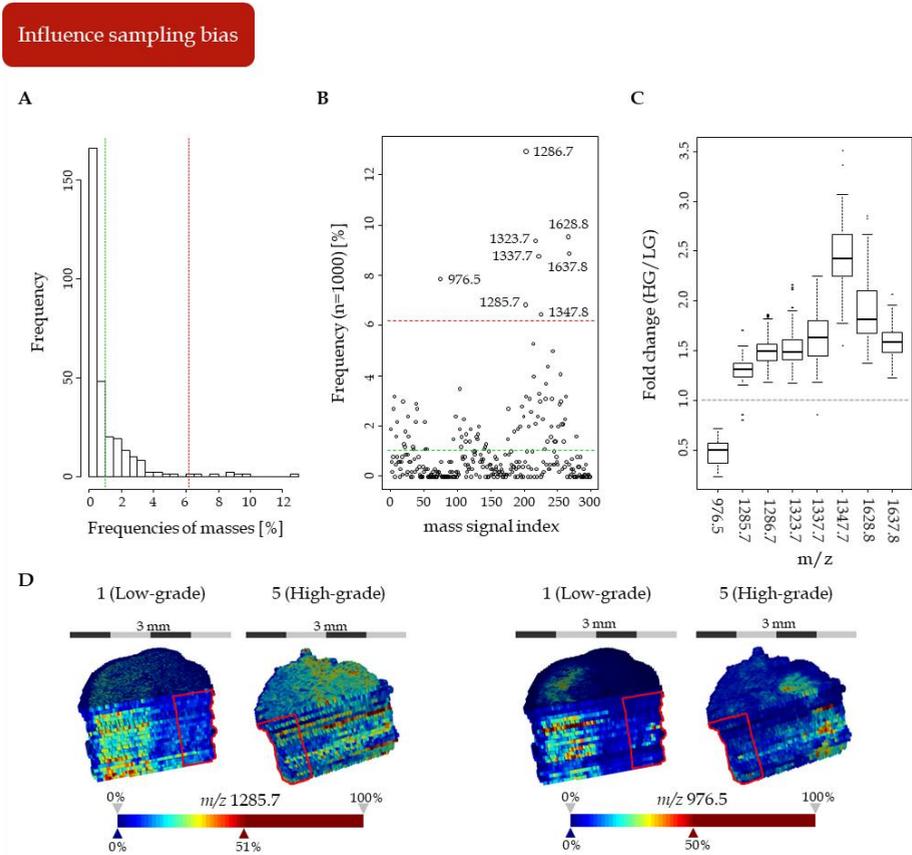


**Figure 4.** Unsupervised multivariate analysis and correlation with tumor regions. The unsupervised multivariate analysis was performed on the single pixels using probabilistic latent semantic analysis (pLSA). (A) Shows the visualization of the pLSA-component 10 in dataset 1 (corresponding H&E stains can be found in Figures 1C and 3) where high coefficients seem to co-localize well with tumor (red) and low coefficients with non-tumor regions (blue). (B) The coefficients of component 10 are shown for all samples in a boxplot illustrating the difference in coefficients for both histological entities. The red line depicts the optimal threshold at 0.1, which was determined using a ROC analysis and Youden’s index (C).

## Influence of sample bias on biomarker discovery

Histopathological annotations, resulting in increased cellular and histological specificity, have been key for success in MSI-based studies in order to increase statistical sensitivity for biomarker discovery.<sup>[27]</sup> However, only one section per patient has been usually used for this purpose and might affect the results due to a sampling bias.<sup>[140]</sup> In our case, a 3D volume has been collected per patient allowing for the investigation of how this sampling would influence biomarker discovery. To investigate this we randomly sampled a single section per patient a 1000 times and resolved which masses are found significantly different ( $P \leq 0.05$ ) in each run between the high- and low-grade patients. On average 3 masses were found significantly discriminating in each run. Although this number looks comparable to the sensitivity of other MSI studies<sup>[141]</sup>, the overlap and therefore the reproducibility between the runs was low. This is illustrated in Figure 5A which shows that the average consistency of a significant mass

across the 1000 runs was ~1%. However, from a statistical point of view, a frequency of > 6.2% is already significantly higher than by chance. Applying this frequency threshold to the found discriminating masses resulted in seven peptides being significantly discriminating (Figures 5A and 5B). Of these, six are more intense in high-grade and one in low-grade based on their fold changes (Figure 5C). Nevertheless, the overall low frequencies of masses indicate that there can be significant influences on biomarker discovery due to sampling bias, especially if the cohort is small. Inversely, if the cohort is larger, such as in classical 2D-MSI studies, the variance due to patient-related differences is expected to be much bigger than compared to sampling bias within a patient making such 2D-based results still valid.<sup>[142]</sup>

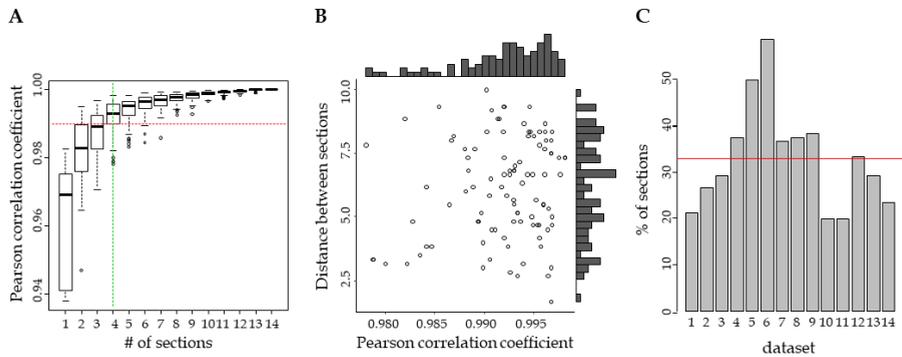


**Figure 5.** Investigation of sampling bias on biomarker discovery for high-grade cancers. To test sampling bias, one section per patient was randomly sampled and mass channels were tested for discriminating power between the two grades. This was repeated 1000 times and the frequency of significance for every mass channel was recorded. The histogram gives an overview of the range of frequencies (A). A binomial test found a frequency of  $> 6.2\%$  significantly constant (red line). (B) Shows the significantly frequent mass channels above the red dashed line. The fold changes of the frequent peptides indicate whether they are more present in high- or low-grade tumors (C). Example distributions of two frequent mass channels are shown in (D) for a high-grade specific (left) and a low-grade specific peptide (right) where tumor regions are indicated in red (corresponding H&E stains for datasets 1 and 5 can be found in Figures 1C and 3).

## Sampling extent and representativeness

To overcome the consequences of sampling bias based on one section, we examined for our samples how many sections need to be sampled to acquire a sufficiently accurate representation of the complete sample. Therefore, the average spectrum of the full dataset was calculated as a reference. From an increasing number of randomly sampled sections the average spectra were calculated and their correlation to the reference determined by the Pearson correlation coefficient. As expected, the results show that tissues are better represented as more sections are taken, shown in Figure 6A for dataset 1. If we consider a correlation coefficient of 0.99 sufficiently representative, then four sections for dataset 1, which is 29% of the sections, need to be measured. Overall, the minimum sampling percentage for all datasets was on average 33% (Figure 6C). Interestingly, we observed a batch of higher minimum percentages for datasets 4-9 in Figure 6C. Given a randomized experimental setup with respect to the clinical data of the patients and the numbering of the datasets in chronological order of their measurement, this means that the calculated minimum sampling percentage is most likely also influenced by time-related technical variation. Moreover, considering a 20% dropout rate according to our outlier detection (Figure 2), we recommend that a minimum of 41% of the volume of a sample is required to be measured. With respect to the sampling position, there is no correlation between the average distance between the randomly sampled sections and the achieved correlation to the ground truth, i.e., it does not matter from where in the 3D volume the sections are sampled (Figure 6B). This is most likely due to the fact that our sample volumes are relatively small (200  $\mu\text{m}$  in depth). If thicker tissues are being analyzed, it can be argued that representative sampling has to be done equally spread throughout the whole sample as the influence of tissue heterogeneity is larger in these cases. Even though our samples are small, this analysis does show the potential of determining for large patient studies how many sections are required for an accurate representation of the whole patient's sample.

## Sampling extent and representativeness



**Figure 6.** Relation between sampling extent and coverage of molecular information. For every patient, a ground truth spectrum was calculated by averaging all spectra across all sections. This is repeated for a randomly sampled number of sections ranging from 1 to the maximum number of sections available in a sample. The Pearson correlation coefficient was calculated between this and the ground truth spectrum for quantifying the deviation of the subsampled data from the full data. This was repeated 100 times. The result for dataset 1 is shown in (A). The green line shows the minimum number of sections needed to achieve a sufficient high similarity to the ground truth (correlation coefficient  $\geq 0.99$ , red dashed line). For the same dataset, (B) depicts the relation between the correlation coefficient and the physical distance in  $z$ -direction among the sampled sections. The minimum number of sampling extent varies between datasets but was on average 33% (horizontal red line) (C).

## Conclusion

Fourteen 3D-MSI datasets of FFPE human bladder cancer resections have been created by measuring the peptides produced following *in situ* tryptic digestion. These datasets have been used to set-up a data analysis pipeline for this type of data. In a first step, we proposed two methods for the detection of outliers: intact cytochrome *c* to monitor digest efficacy, and regression to evaluate consecutive section consistency. Both analyses show that our data is robust and as a consequence, we observed a consistently high spatial correlation with histology. The acquired 3D datasets allowed for the investigation of sampling bias on biomarker discovery and sample representativeness. We observed significant effects due to sampling bias on the final peak list of grade-associated peptides. To address sampling bias, we found out that a subset (on average 33%) of sections is needed to be

sufficiently representative for one sample (sparse 3D-MSI). This strategy can be transferred easily to other studies in order to determine a minimum sampling rate, which can significantly reduce the experimental workload of 3D-MSI studies. However, in order to obtain a more complete representation of one tumor, our approach has to be extended to samples from multiple sites of the tumor.<sup>[143]</sup> Both, multi-site sampling as well as sparse 3D-MSI could potentially be guided by a preceding non-invasive ex-vivo imaging of the whole tumor or individual subsamples (micro-CT, etc.).<sup>[144]</sup> Even then a reliable biomarker discovery has to be matched with a sufficient number of subjects involved. Although in our case the sampling pool was too small to provide results with high statistical power, the reported pipeline can be regarded as a starting point to further investigate and extend other multi-patient FFPE 3D-MSI studies.

## Acknowledgments

This work has been made possible with the financial support of the Dutch province of Limburg and has been funded by ITEA and RVO by means of project numbers ITEA151003/ITEA 14001. B.B acknowledges the support of the European Union (ERA-NET: TRANSCAN 2).



# ***Chapter Four***

## **2.5D N-glycan mass spectrometry imaging of esophageal adenocarcinogenesis**

**D. R. Naomi Vos**, Sybren L. Meijer, Ron M. A. Heeren, Shane R. Ellis, Benjamin Balluff

**Manuscript in preparation**



Esophageal adenocarcinoma (EAC) is a complex and heterogeneous disease that originates from Barret's esophagus. Progression from this precursor lesion via low- and high-grade dysplasia to EAC involves a large variety of genetic alterations.

These alterations do not occur throughout the whole tissue but are very localized events. This makes EAC both a molecularly and spatially complex disease. Both complexity scales have to be captured to gain more information about EAC progression and the accompanying molecular changes. One of these changes is in N-glycan expression as the glycosylation pathway is one of the mechanisms affected during esophageal adeno-carcinogenesis.

In this chapter, we employ mass spectrometry imaging (MSI) to spatially resolve N-glycan changes during EAC. To this end, 24 human FFPE esophageal resections spanning all stages of EAC progression were studied. The heterogeneous nature of the disease makes that both 3D-MSI, as well as 2.5D-MSI, have been employed. First, 3D-MSI was used to image one patient per stage to determine the number of sections needed to be representative of the whole volume. This analysis showed that a total of 4 sections out of 20 were needed. This subset was measured for the remaining patients and is called 2.5D-MSI.

Analyses of the 2.5D-MSI datasets revealed 3 differentially expressed N-glycans across the EAC stages. These N-glycans were classified as high-mannose and tetra-antennary N-glycans whose presence could be explained by publically available gene expression data. Combined these data confirm that N-glycosylation plays a vital role in EAC progression. As the role of glycosylation is gaining more interest in cancer research our results also show the added value of combining localized MSI analysis with gene expression data to gain a deeper understanding of N-glycan changes.

## Introduction

Esophageal cancer is a very aggressive disease that ranks sixth in mortality among all cancers, making it one of the deadliest cancers worldwide. Even though advances have been made in both diagnosis and therapeutics it still has a poor 5-year survival rate of 15-25%.<sup>[145-147]</sup> Amongst esophageal cancers, the incidence of esophageal adenocarcinoma (EAC) has risen dramatically over the last decades.<sup>[148]</sup>

EAC follows from Barrett's esophagus or Barrett's mucosa as a precursor lesion via low- and high-grade dysplasia.<sup>[124, 149, 150]</sup> Barrett's esophagus (BE) results from a chronic injury caused by acid and bile reflux whereby squamous epithelium of the esophageal mucosa undergoes a metaplastic transformation into columnar intestinal epithelium.<sup>[149-151]</sup> The progression from BE to EAC is accompanied by a large number of alterations in oncogenic pathways and tumor suppressor genes that play a role and vary individually in EAC development.<sup>[152]</sup> The complexity of EAC is also reflected in its molecular heterogeneity at the spatial level.<sup>[124, 153]</sup> Understanding the local molecular alterations underlying adeno-carcinogenesis in the context of tissue-based molecular and cellular heterogeneity will aid in understanding the mechanisms of cancer onset, its prevention, and the development of personalized therapy.<sup>[153]</sup> Especially, since EAC precursor lesions are very locally confined events, gaining insight into the molecular sequence of events that lead to EAC requires an *in situ* molecular analysis.

One technique that allows both the label-free detection of molecular species and records their spatial distribution is mass spectrometry imaging (MSI).<sup>[27, 154]</sup> It allows the simultaneous detection of a large number of compounds from a broad range of molecular classes (small molecules, lipids, peptides, proteins, and modifications thereof). In cancer research, it has already been used to determine metabolite and lipid profiles in therapy-resistant tumors<sup>[155]</sup>, to reveal tumor heterogeneity<sup>[24, 156]</sup>, as well as to aid in classifying tumors.<sup>[157]</sup>

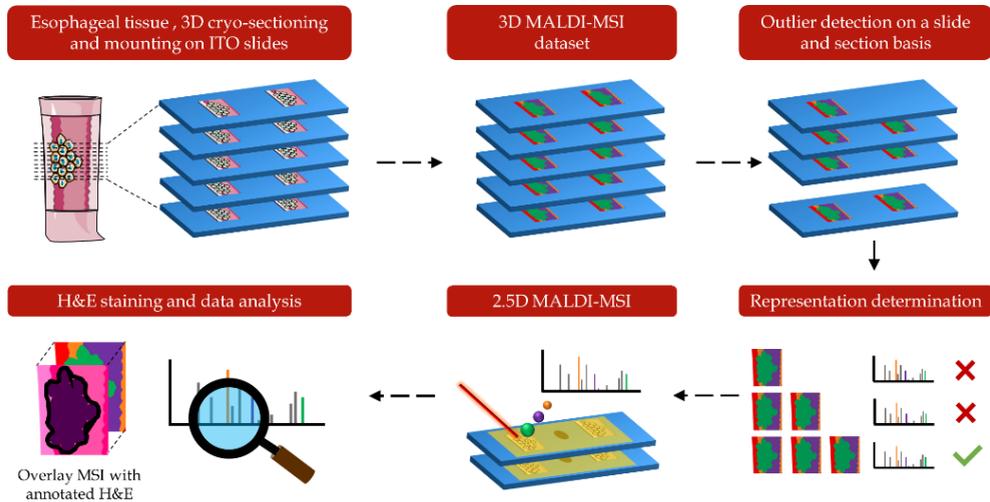
In EAC research, MSI has already been employed for the identification of specific peptide profiles for EAC stages in 300 patients assembled in tissue microarrays.<sup>[158]</sup> In two other studies on intact proteins, MSI identified several proteins that could be linked to EAC development, EAC metastases, patient prognosis, and chemotherapy response.<sup>[159, 160]</sup> A study using desorption ionization MSI investigated lipid profiles of metastatic and non-metastatic primary tumors, which allowed the metastatic status of patients to be predicted with 98% accuracy.<sup>[161]</sup>

Glycosylation is one of the most important posttranslational modifications of proteins<sup>[162]</sup> and aberrant glycosylation has been found implicated in various types of cancers and EAC during malignant transformation.<sup>[163]</sup> N-linked glycans have already been studied by MSI in ovarian and breast cancer. It allowed the differentiation between tumor and non-tumor regions<sup>[164]</sup> and identified N-glycans specific for late-stage compared to the early-stage ovarian cancer like sialylated structures and oligomannose N-glycans.<sup>[165]</sup> In breast cancer, MSI found the specific presence of polylactosamine glycans in triple-negative, metastatic, and advanced HER2+ patients<sup>[166]</sup> and F(6)A4G4Lac1, a core-fucosylated tetra-antennary glycan, correlated with poor clinical outcome.<sup>[167]</sup>

The aforementioned studies provided evidence for the added benefit of MSI in differentiating different stages as well as finding new potential markers. Although these results are encouraging, most MSI studies limit themselves to one section per sample. However, EAC is a heterogeneous disease, and analyzing only a single section can lead to discrepancies with a full tissue analysis.<sup>[39]</sup> It is important to overcome the sampling bias and analyze either the whole 3D volume or measure a representative subset of sections of the whole tissue (2.5D), as we have shown previously.<sup>[68, 85]</sup>

In this research, we have employed 3D matrix-assisted laser/desorption ionization MSI (MALDI-MSI) followed by 2.5D MALDI-MSI to investigate N-glycan changes in human esophageal adeno-carcinogenesis (Figure 1). A total of 24 formalin-fixed paraffin-embedded (FFPE) esophageal resections of low-grade dysplasia (n=2), high-grade dysplasia (n=3), low- and high-grade dysplasia (n=1), intra-mucosal cancer (n=8), adenocarcinoma (n=4),

and adenocarcinoma with metastases (n=6) were analyzed to ensure a complete overview of all steps of EAC adeno-carcinogenesis.



**Figure 1.** Schematic overview of the different steps in the experimental set-up for a single EAC stage. First, 3D-MSI is employed on one patient sample, and outliers are removed on a slide and section basis. After outlier removal the number of sections needed to be representative of the whole volume is determined. This subset is used for subsequent 2.5D-MSI of the remaining patients and used for further analyses where the resulting images are overlaid with their annotated H&E stains. This process is repeated for all the different EAC stages.

## Experimental

### Patient material

Human formalin-fixed, paraffin-embedded (FFPE) esophagus tissues of 24 patients were collected and analyzed in this study. Of these 24 patients, 2 had low-grade dysplasia, 3 had high-grade dysplasia, 1 had low- and high-grade dysplasia, 8 had intra-mucosal cancer, 4 had adenocarcinoma, and 6 had adenocarcinoma with metastasis. From each patient, 20 consecutive sections with a thickness of 5  $\mu\text{m}$  and spaced 10  $\mu\text{m}$  apart were provided by the Amsterdam Medical Center (Amsterdam, Netherlands) on indium tin oxide-coated conductive (ITO) glass slides ( $R_s$  4-8  $\Omega/\text{sq}$ , Delta Technologies

Ltd, Loveland, Colorado, USA) with two sections per slide. For all patients, histological annotations and grading were available and did not require approval from the relevant Institutional Ethics Committee under applicable local regulatory law ('Code of conduct', FEDERA).

## Materials

Xylene, HPLC grade ethanol, UHPLC grade water, ULC/MS - CC/SFC acetonitrile, and ULC/MS - CC/SFC grade methanol were purchased from Biosolve Chimie SARL (Dieuze, France). Citric acid monohydrate,  $\alpha$ -cyano-4-hydroxycinnamic acid (CHCA,  $C_{10}H_7NO_3$ ), and trifluoroacetic acid (TFA, Chromasolv®, for HPLC), and RNase B (Glycoprotein Standard from bovine pancreas) were purchased from Sigma-Aldrich (Zwijndrecht, The Netherlands). PNGase F was purchased from N-Zyme Scientifics, LLC (Doylestown, Pennsylvania, USA). Eosine-Y, Alcoholic was purchased from Avantor® Performance Materials B.V. (Deventer, the Netherlands). Hematoxylin solution modified according to Gill and entellan® new was purchased from Merck KGaA (Darmstadt, Germany), and coverslips were purchased from Thermo Scientific (Waltham, Massachusetts, USA).

## Sample preparation

Sections were pre-heated on a heating plate at 60°C for an hour to melt the paraffin. Immediately after, the sections were washed with xylene for 1 x 5 minutes and 2 x 10 minutes to remove the paraffin, followed by subsequent washes with ethanol (3 x 2 minutes) to remove the lipids and with water (2 x 5 minutes) to rehydrate the sections. Antigen retrieval was subsequently performed to undo the cross-linking induced by the formalin fixation with a 10 mM citric acid solution at pH 6.0 using the Antigen Retriever 2100 (Aptum Biologics, Southampton, UK) at 121°C for 20 minutes. Sections were then washed with water (2 x 1 minute) and after drying of the slides 1  $\mu$ L of 1  $\mu$ g/ $\mu$ L RNase B was spotted on the slides as quality control. This was followed by application of 15 layers of 200 ng/ $\mu$ L water-dissolved PNGase F using a SunCollect pneumatic sprayer (SunChrom Wissenschaftliche Geräte GmbH, Friedrichsdorf, Germany) at a flowrate of 10  $\mu$ L/min, nozzle

speed of 900 mm/min, nozzle height of 25 mm, and a track spacing of 1 mm where the slides were randomized. This was followed by incubation of the slides for 3 hours in an incubation box containing 50% methanol at 37°C. Afterward, the sections were coated in the same randomized order with 8 layers of 5 mg/mL CHCA in a 50% acetonitrile solution containing 0.2% TFA using an automated TM-sprayer (HTX Technologies, LLC, North Carolina, USA) at 30°C with a flowrate of 0.05 mL/min with no drying time in between the layers.

Following MSI analysis, sections were H&E stained by first washing off the matrix from the slides in 70% ethanol for 2 x 3 minutes, followed by a 3-minute wash with Milli Q. Slides were stained with hematoxylin (3 minutes), washed for 3 minutes with tap water to remove excess hematoxylin, then stained with eosin (30 seconds), washed again for 3 minutes to remove excess eosin, followed by a 1-minute ethanol wash and a 30 seconds xylene wash before attaching coverslips to the slides using Entellan as a mounting medium. Optical images of the H&E tissue section were acquired using an optical microscope scanner (Mirax Desk, Zeiss, Jena, Germany). For each patient, one unmeasured section was stained and annotated by a pathologist. These annotations were then transferred via image co-registration onto the scanned and H&E stained MSI sections.

## **MALDI-MSI measurement**

All sections were measured on a rapifleX MALDI-ToF/ToF instrument (Bruker Daltonik GmbH, Bremen, Germany) in positive-ion mode at a spatial resolution of 50  $\mu\text{m}$  in the  $m/z$  range of 800-4000 by scanning the laser across a 45x45  $\mu\text{m}^2$  area and summing 200 shots per pixel. To obtain reproducible results, detector checks were performed at the start of each measurement day and laser intensity and height were optimized on the matrix intensity off tissue to obtain similar intensities for the matrix ( $\sim 3\text{--}4 \times 10^4$ ). Before each measurement, the instrument was calibrated using red phosphorus as a calibration agent. During acquisition, the data were automatically processed with Savitzky-Golay smoothing (0.01  $m/z$  width, and 1 cycle), TopHat baseline subtraction, and centroid peak detection (0.05  $m/z$  peak width,  $S/N \geq 5$ , and height 78%).

## Data analysis

In total four complete 3D-MSI datasets were acquired (one for each EAC stage) and imported into separate SCiLS Lab MVS (SCiLS, Bremen, Germany) files for outlier detection and representativity analysis using a 200 data points precision for the TopHat baseline subtraction. Spectra were TIC normalized and a  $\pm 0.2$  Da interval width per peak was chosen for further analysis. The overall average mass spectrum was exported to mMass<sup>[117]</sup> for peak picking to reduce the amount of data prior to processing. Peak picking in mMass was preceded by 35 precision baseline correction, Gaussian smoothing with a 0.3 Da window size and 1 cycle, crop from 920 to 3000  $m/z$ , de-isotoping with a 70% intensity tolerance and 0.2 Da isotope mass tolerance. Peaks were picked that had an  $S/N \geq 5$  at maximum height in the mean spectrum and matrix peaks were manually removed. The resulting list was imported back into SCiLS and used for further analysis. Maximum intensities for each peak interval (0.2 Da) in the peak list for every section was afterward exported from SCiLS in tabular form for sequential outlier detection and representation analysis in RStudio (version 1.2.1335).

MSI data from the RNase B quality control were imported into separate (per 3D dataset) SCiLS Lab MVS files for outlier detection employing a 200 precision TopHat baseline subtraction. All spectra were normalized to TIC for subsequent analysis and peak picking was performed in mMass on the exported overall average mass spectrum employing the same settings as for the 3D-MSI data. Peaks were picked that had an  $S/N \geq 3$  and a relative intensity threshold  $\geq 1\%$  (with respect to the base peak) at maximum peak height and matrix peaks were removed by expelling those peaks in the 800-2500  $m/z$  range that have a mass defect  $\leq 0.3$ . The resulting list was imported back into SCiLS and used for further analysis.

Outlier slides or sections were identified based on two outlier strategies that were previously reported by our group.<sup>[85]</sup> For identifying outlier slides, Principal Component Analysis (PCA) was performed on the TIC normalized mean spectra of the RNase B quality control spots per slide using a 95% error ellipse in the first two-component space. All RNase B measurements outside this ellipse were considered outliers and corresponding slides were removed from further analysis. To identify outlier sections, we use a method that

investigates the intensity gradient of single  $m/z$  values in the z-directionality within a 3D model. Per  $m/z$  value, a regression is fitted to a linear or quadratic function, depending on which one fits best according to the Akaike information index, and the section with the maximum residual is recorded. This is repeated for all  $m/z$  values and the sections that have a residual count higher than a set threshold (usually,  $\geq 99\%$  confidence interval of expected by-chance mean frequency, formula 1) are considered outliers. A combination of these two outlier detection methods resulted in a list with outlier sections that were not considered for further analysis.

$$\text{Formula 1: } \textit{residual frequency threshold} = \frac{\# \textit{ of N-glycan masses}}{\# \textit{ of sections}} + 2.58 \times \frac{\textit{SD}(\# \textit{ of residual masses per section})}{\sqrt{\# \textit{ of sections}}}$$

Based on the 3D-MSI data after outlier detection, it was determined how many sections per patient are sufficiently representative of the whole volume. This was done as described previously by our group.<sup>[85]</sup> Briefly, an average spectrum is created for the whole 3D volume and each section. The mean spectrum of the whole 3D volume is then compared to the average spectra obtained for a subset of  $n$  random sections and repeated 100 times. Subsequently, this is repeated with  $n+1$  sections until all sections have been used. For each combination, a Pearson correlation coefficient is calculated and the first number of sections that reaches 0.99 defines the minimum number of sections required.

In total 20 2.5D-MSI datasets were acquired that included RNase B quality controls for each measured slide. All sections were individually recalibrated by selecting one spectrum per section containing N-glycans covering the whole spectral range. These spectra were then manually recalibrated in flexAnalysis (Bruker Daltonik GmbH, Bremen, Germany) using a mass list containing known molecular signals up to  $m/z$  2540 (Table 1) and a 500 ppm tolerance to be able to recalibrate the low-intensity masses at the higher end of the spectrum. The calibration constants of each recalibrated spectrum were copied and pasted to all other spectra from the same respective dataset. The recalibrated MSI sections were manually co-registered with their

annotated H&E images to obtain average mass spectra for the annotated regions of interest and imported into a SCiLS Lab MVS file for analysis using 200 precision TopHat baseline subtraction. Spectra were TIC normalized and a  $\pm 0.2$  Da interval width per peak was chosen for further analysis. The overall average mass spectrum was exported to mMass<sup>[117]</sup> for peak picking to reduce the amount of data prior to processing. Peak picking in mMass was preceded by 35 precision baseline correction, Gaussian smoothing with a 0.3 Da window size and 1 cycle, crop from 920 to 3000  $m/z$ , de-isotoping with a 70% intensity tolerance and 0.2 Da isotope mass tolerance. Peaks were picked that had an  $S/N \geq 5$  at maximum height in the mean spectrum and matrix peaks were manually removed. The resulting list was imported back into SCiLS and used for further analysis.

Recalibration list	
$m/z$	Molecule / N-glycan
861.11	Matrix cluster: M4KNa3-H
1072.10	
1257.44	H5N2
1419.46	H6N2
1647.61	H4N4F1
1663.56	H5N4
1809.62	H5N4F1
1905.63	H9N2
1976.65	H5N4S1
2100.76	H5N4S1F1
2122.77	H5N4S1F1
2540.00	H7N6F1

**Table 1.** Overview of the masses used for recalibration of the 20 2.5D-MSI datasets.

MSI data from the RNase B quality control was also imported into SCiLS Lab MVS files for outlier detection employing a 200 precision TopHat baseline subtraction. All spectra were normalized to TIC for subsequent analysis and peak picking was performed in mMass on the exported overall average mass spectrum with the same settings as for the 3D-MSI RNase B

data (see above). The resulting list was imported back into SCiLS and used for further analysis. The RNase B quality controls were checked for outliers in the 2.5D-MSI data as described above. Potential outliers were removed from the data and annotations for each section were grouped to obtain an average spectrum for every patient and section.

The SCiLS data was exported using the maximum mean value for every peak to investigate if N-glycans show significant differences in abundance over the different stages. Statistical analyses were performed in RStudio (version 1.2.1335). Analysis of Variance was employed, for which the data were log-transformed to comply with the requirement of normally distributed data. As the data is composed of 20 patients with technical replicates, mixed effect models were calculated with the EAC stage as a fixed effect, the patient as a random effect, and the measured glycan-signal abundance as a dependent variable. The obtained results were corrected for multiple testing using the Benjamini-Hochberg method with a false-discovery-rate threshold of 0.1. Post-hoc pairwise comparisons were corrected using Tukey's method.

N-glycans were tentatively identified by mass-matching them with a maximum tolerance of 60 ppm to N-glycans reported in the literature (Table 2).<sup>[21, 168-170]</sup>

Tentatively identified N-glycans					
<i>m/z</i>	Heijs <sup>[169]</sup>	<i>m/z</i>	Heijs <sup>[169]</sup>	<i>m/z</i>	Heijs <sup>[169]</sup>
933.37	H3N2	1663.555	H5N4	2174.8	H6N5F1
1079.385	H3N2F1	1688.599	H3N5F1	2215.841	H5N6F1
1095.381	H4N2	1704.596	H4N5	2268.797	H5N4S2
1257.394	H5N2	1743.541	H8N2	2289.759	H5N4S2
1282.438	H3N3F1	1793.628	H4N4F2	2304.873	H5N5F3
1298.435	H4N3	1809.625	H5N4F1	2320.871	H6N5F2
1339.476	H3N4	1825.622	H6N4	2325.835	H5N5S1F1
1419.461	H6N2	1850.666	H4N5F1	2361.911	H5N6F2
1444.505	H4N3F1	1866.663	H5N5	2377.909	H6N6F1
1460.502	H5N3	1891.707	H3N6F1	2393.906	H7N6
1485.491	H3N4F1	1905.608	H9N2	2435.884	H5N4S2F1
1501.488	H4N4	1955.695	H5N4F2	2466.941	H6N5F3
1542.529	H3N5	1976.657	H5N4S1	2487.902	H6N5S1F1
1581.528	H7N2	1996.736	H4N5F2	2507.982	H5N6F3
1606.517	H5N3F1	2012.733	H5N5F1	2523.979	H6N6F2
1622.514	H6N3	2028.73	H6N5	2539.976	H7N6F1
1647.558	H4N4F1	2122.727	H5N4S1F1	2581.017	H6N7F1

**Table 2.** Overview of the tentatively identified N-glycan masses by mass-matching with reported N-glycans by Heijs et al.<sup>[169]</sup>

The publically available gene expression datasets from the NCBI GEO Datasets repository (<https://www.ncbi.nlm.nih.gov/gds>) on esophageal cancer GSE1240<sup>[171]</sup> and GSE2444<sup>[172]</sup> were searched for known enzymes to be involved in N-glycan synthesis (Tables 3 and 4 respectively).<sup>[173, 174]</sup> In case of dataset GSE1240, the enzymes were re-coded to be 0 (absent), 0.5 (marginal), or 1 (present).

GSE1420							
Protein	# ID ref's	Normal (n=8)			Adenocarcinoma (n=8)		
		A	M	P	A	M	P
MAN1A1	2	10	1	5	5	0	11
MAN1A2	4	15	1	16	12	4	16
MAN1B1	2	8	0	8	3	0	13
MAN2A1	1	4	0	4	1	1	6
MAN2A2	2	6	2	8	0	1	15
MGAT1	1	0	0	8	0	0	8
MGAT2	3	8	0	16	8	0	16
MGAT4A	1	4	0	4	3	2	3
MGAT4B	1	0	0	8	0	0	8
MGAT4C	1	6	2	0	8	0	0
MGAT5	2	8	0	8	8	0	8
B3GALT1	1	8	0	0	8	0	0
B3GALT2	2	16	0	0	15	0	1
B3GALT5	1	8	0	0	7	0	1
B4GALT1	4	11	6	15	16	2	14
B4GALT2	1	0	2	6	0	0	8
B4GALT3	1	0	0	8	0	0	8
B4GALT4	2	5	0	11	1	0	15
B4GALT5	2	0	0	16	0	0	16

**Table 3.** Overview of the enzymes involved in N-glycan synthesis in publically available dataset GSE1420. Per protein, there were either one or multiple reference ID's found and per ID it was stated if it was absent (A), marginal present (M), or present (P) for each sample.

GSE2444		
Protein	Adenocarcinoma (n=22)	Adenocarcinoma with metastases (n=4)
MAN1A1	0.843326448	0.883569487
MAN1A2	0.918323302	0.835617837
MAN2A2	0.810151274	0.923930214
MGAT1	1.158070038	1.23564253
MGAT2	1.015546371	1.042224968
MGAT4B	1.06440933	1.062230245
B4GALT1	0.888545504	1.021724706
B4GALT2	0.931151213	0.97537298
B4GALT3	0.956965641	0.910412448
B4GALT5	0.852269253	0.821324874

**Table 4.** Overview of the enzymes involved in N-glycan synthesis in publicly available dataset GSE2444. Per protein, log-ratio values were given between the sample and healthy control. These values are positive if they are upregulated in the sample or negative if they are down-regulated compared to healthy tissue.

## Results

### Determination of the minimum number of sections per patient using 3D MALDI-MSI

Imaging the tissues of all patients in 3D is time-consuming work that merits the question if it provides sufficiently more information over a random subset of sections to be worthwhile. We set out to determine the minimum number of sections statistically representative of a whole 3D volume. For this purpose, one full 3D-MSI dataset was acquired for every EAC stage (Figures 1 and 2).

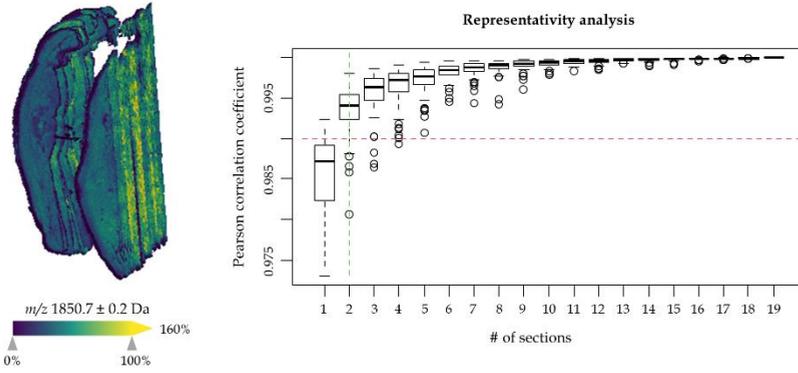
Possible outlier slides were removed based on RNase B quality controls that were present on each slide. To do so, a PCA was performed on the RNase B spectra and a 95% error ellipse was calculated in the first two principal component space for each dataset. The results show that the enzyme worked similarly on almost all slides (Figure 3) except for one slide in the low- +

high-grade dysplasia dataset, which exhibited a lower digestion efficiency and was therefore removed. Further outlier sections were determined based on the intensities in a 3D volume that are assumed to follow a certain gradient in the  $z$ -direction.<sup>[85]</sup> This was tested by a regression analysis based on either a linear or quadratic function. Sections that show the highest residual value over all  $m/z$  values more than by chance (formula 1) are considered outliers. On average two sections were classified as outliers per 3D-MSI dataset through the application of this criterion (Figure 3).

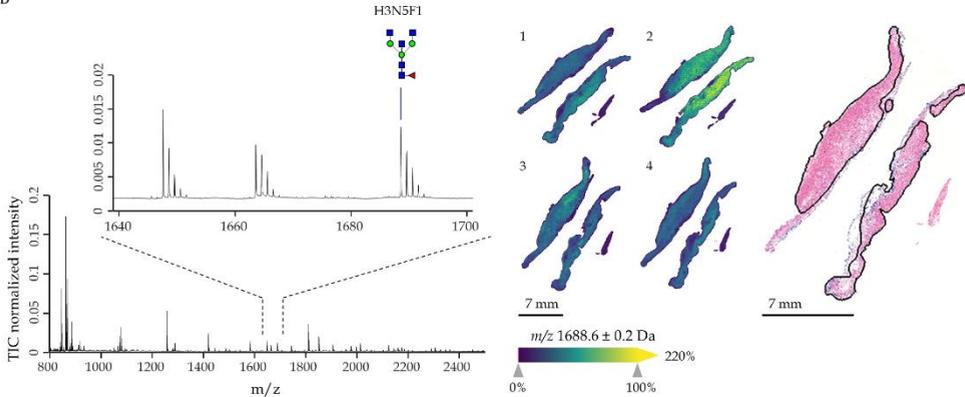
For each of these 3D-MSI datasets, the aim was to subsequently find a set of representative sections (2.5D) for the whole tissue. An average spectrum was calculated as a reference for every 3D-MSI dataset for this purpose. Then this reference spectrum was compared to the average spectrum calculated from an increasing number of randomly chosen sections by determining the similarity between these two profiles using Pearson correlation. Taking a minimum coefficient of 0.99, 2-3 sections were found to be sufficiently representative for a 3D-MSI dataset (Figure 3), which is 10-20% of the total sample size of 20 sections. To have a consistent amount of sections for all EAC stages, we decided to measure 4 sections for all remaining patients, as 2 sections were placed on one single slide. For the 2.5D-MSI, the two slides chosen were slide 3 (sections 5-6), and slide 8 (sections 15-16).

## Obtained 3D and 2.5D data

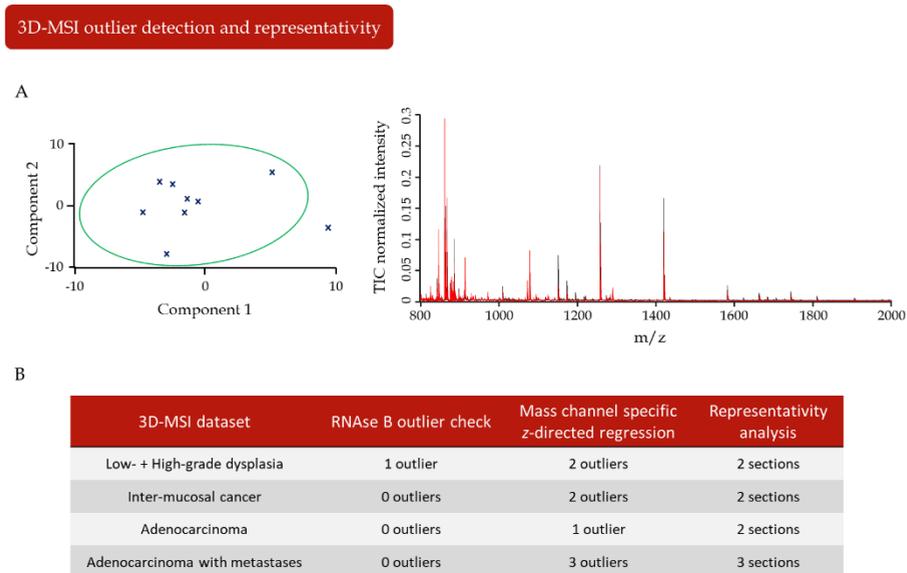
A



B



**Figure 2.** Examples of 3D and 2.5D mass spectrometry imaging data. (A) 3D visualization of the adenocarcinoma dataset consisting of 20 sections with the corresponding representativity analysis. (B) Spectra and visualization of one 2.5 MSI dataset consisting of 4 sections along with the pathologist annotated H&E.



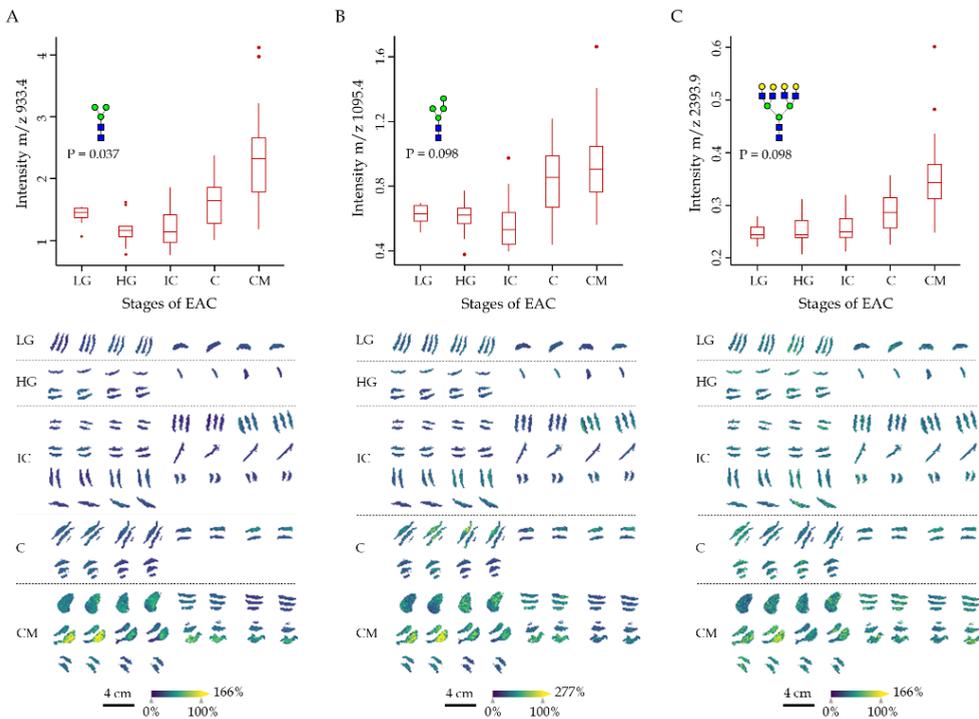
**Figure 3.** Overview of the 3D-MSI outlier and representativity analysis results. (A) Example of the RNAse B-based PCA-based outlier detection with a 95% error ellipse where the crosses indicate each measured RNAse B spot. Next to it is the RNAse B spectrum of the outlier (red) shown as an overlay with the normal RNAse B spectrum (black). (B) Table summarizing the outlier detection results of all 3D-MSI datasets using the different outlier detection mechanisms as well as the representativity analysis results.

## Mapping N-glycan changes during esophageal adeno-carcinogenesis

Glycosylation plays an important role in disease progression during esophageal adenocarcinogenesis (EAC).<sup>[163]</sup> Mapping these changes in glycosylation could provide useful insights into disease progression and offer a new approach to EAC staging. We, therefore, performed 2.5D N-glycan MSI by measuring 4 sections from 20 patients with different stages of esophageal adenocarcinogenesis: 2 low-grade dysplasias, 3 high-grade dysplasias, 7 intra-mucosal cancers, 3 adenocarcinomas, and 5 adenocarcinomas with metastases. Annotations by an experienced pathologist were used to extract the N-glycan profiles from dysplastic and neoplastic tissue areas.

After peak picking, the N-glycan profile contained 108 signals. A mixed-effect model was calculated for every  $m/z$  species of the N-glycan profile followed by multiple testing corrections to determine if there are differentially expressed N-glycans between the different EAC stages. This analysis identified three  $m/z$  species all of which were found at higher abundance in carcinoma compared to the precursor lesions (Figure 4).

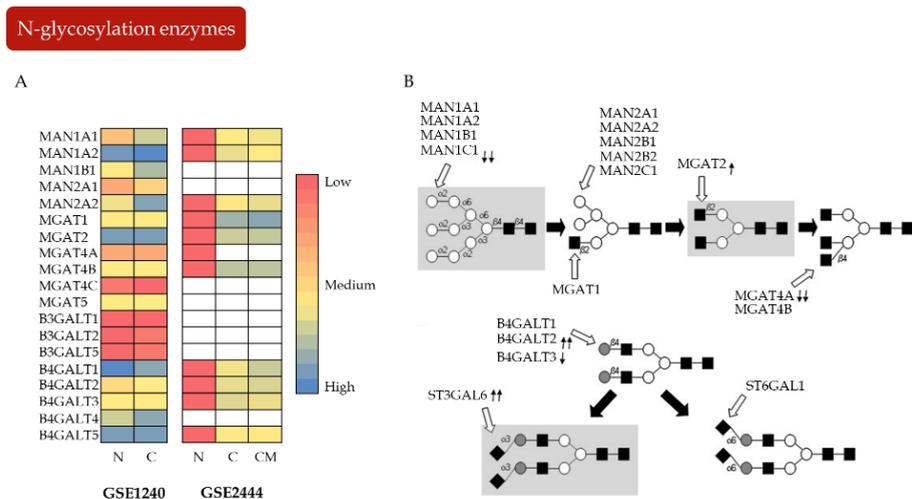
### Differentially expressed N-glycans



**Figure 4.** Boxplots of the three N-glycans found to be more abundant in the adenocarcinomas with the corresponding visualization of said molecule in the 2.5D-MSI dataset below. The different stages have been abbreviated to LG (low-grade dysplasia), HG (high-grade dysplasia), IC (inter-mucosal cancer), C (adenocarcinoma), and CM (adenocarcinoma with metastases). (A) Boxplot of  $m/z$  933.4, tentatively assigned to man3 (H3N2), with its corresponding structure and p-value. (B) Boxplot of  $m/z$  1095.4, tentatively assigned to man4 (H4N2), with its corresponding structure and p-value. (C) Boxplot of  $m/z$  2393.9, tentatively assigned to a tetra-antennary N-glycan (H7N6), with its corresponding structure and p-value.

One of these,  $m/z$  933.4 ( $P = 0.037$ ), was found to distinguish EAC patients with metastases from those without. It was tentatively assigned to Man3 (H3N2, Hex3HexNAc2 + 1Na,  $\Delta 53.6$  ppm) based on prior literature.<sup>[169, 170]</sup> The other two N-glycans that were more present in adenocarcinomas compared to dysplasia and intramucosal carcinoma were  $m/z$  1095.4 and  $m/z$  2393.9 (both  $P = 0.098$ ), tentatively assigned to Man4 (H4N2, Hex4HexNAc2 + 1Na,  $\Delta 9.1$  ppm)<sup>[169, 170]</sup> and tetra-antennary N-glycan H7N6 (Hex7HexNAc6 + 1Na,  $\Delta 20.9$  ppm)<sup>[21, 168, 169]</sup>, respectively.

Next, we investigated two publically available gene expression datasets from esophageal cancer, GSE1420<sup>[171]</sup> and GSE2444<sup>[172]</sup> for alterations in enzymes that play a role in the synthesis of the found N-glycans<sup>[173, 174]</sup> in EAC onset or progression (Figure 5A). In this meta-analysis, we found almost all of these enzymes increased in cancer compared to normal esophagus (Figure 5A).



**Figure 5.** Meta-data analyses of the enzymes involved in the synthesis of N-glycans H3N2, H4N2, and H7N6. The different stages have been abbreviated to N (normal), C (adenocarcinoma), and CM (adenocarcinoma with metastases). (A) Heatmap showing the up- or down-regulation of enzymes involved in the synthesis of N-glycans H3N2, H4N2, and H7N6 in publically available datasets GSE1240 and GSE2444. The heatmap shows that almost all enzymes are up-regulated in adenocarcinoma. (B) Overview of which enzymes play a role in N-glycan synthesis and their role.

## Discussion

In this study, we aimed to find N-glycans related to the carcinogenesis of esophageal adenocarcinoma (EAC) using MALDI mass spectrometry imaging (MSI). To achieve this we analyzed 24 human samples covering the whole sequence of EAC progression. We had 20 consecutive sections available for every patient to account for the spatial heterogeneity of this process in all 3 dimensions.

As EAC is a heterogeneous disease, one patient per stage was imaged in 3D first to determine the number of sections needed to be representative for the whole volume.

We performed a previously established outlier detection procedure<sup>[85]</sup>, which determined that on average 2.5 sections needed to be removed per MSI dataset or in total 10 sections of the 80 sections measured giving an 87.5% reproducibility rate. This was considered an acceptable dropout rate since previous and similar experiments on bladder cancer (also FFPE and use of enzymatic digest) gave a lower reproducibility rate of 80%.<sup>[85]</sup>

The representativity analysis determined that 2–3 sections could be considered sufficiently representative (Pearson correlation > 0.99) for the stack of 20 consecutive sections. It is clear that this stack is itself not fully representative for the entire tissue sample or the patient, but it can be considered an approximation. Sampling bias will continue persisting as long as invasive sampling methods are used.

These data were subsequently used to identify differentially expressed N-glycans between the different EAC stages. A total of 108 signals were detected in the mass range typical for N-glycans. Of these, 52 (48%) could be tentatively assigned to N-glycans (Table 2), demonstrating the reliability of the employed protocol to detect N-glycans.

Three N-glycans (H4N2, H7N6, Man3) were found statistically altered between some of the stages of EAC with all of them being more abundant in advanced stages (carcinoma or carcinoma with metastasis) compared to

dysplasia or intra-mucosal carcinoma. Man3 is a pauci-mannose and has, like man4, been already reported in breast and colorectal cancer, and at higher levels in malignant colorectal tumors.<sup>[170, 175]</sup> In ovarian cancer tissues it was however found that this N-glycan was specific for adipose tissue instead of the tumor region.<sup>[164]</sup> H4N2 is a high-mannose N-glycan that has already been found at higher levels in breast and colorectal cancer<sup>[170, 175]</sup>, which is in line with our results. This N-glycan is usually the lowest in abundance of the high-mannose group making this an interesting finding since we did not observe other high-mannose N-glycans to be altered. The other N-glycan, H7N6, was found at higher levels in stage IA3 lung adenocarcinoma and hepatocellular carcinoma with almost no presence in normal tissue.<sup>[21, 168]</sup>

While this confirms the role of these glycans in carcinogenesis, we strived for shedding light on the mechanisms of glycosylation in esophageal cancers. Based on the aforementioned results, we investigated two publically available gene expression datasets for alterations in enzymes that play a role in the synthesis of the found N-glycans. In line with our observations, the involved enzymes MAN, MGAT1, and MGAT2 were up-regulated in esophageal cancer compared to healthy tissues.

In particular, higher levels of the MAN enzymes, which are responsible for trimming down N-glycans starting from a structure with 9 mannoses to Man3 (Figure 5B), could explain the higher levels of Man4 and Man3 observed by us in EAC tissues. Several enzymes, including the MGAT1 and MGAT2 enzymes, facilitate the synthesis of H7N6. These enzymes also exhibit an increase from healthy to cancer, with MGAT1 being further increased in cancer with metastases (Figure 5A). These enzymes add a GlcNAc to the core of N-glycans ending up with a Man3 with two GlcNAc residues forming the base for all multi-antennary N-glycans, where MGAT4 and MGAT5 add the final two GlcNAc to form tetra-antennary N-glycans (Figure 5B). To get to the final structure, the terminal GlcNacs need to be galactosylated, which is done by the B3GAL and B4GAL enzymes.

This study shows that N-glycan MSI data can be contextualized with gene expression data from studies with similar research questions. Relating these two levels of molecular information can, therefore, help to gain a deeper

understanding of the role of glycosylation in EAC progression and gives evidence at multiple levels that specific N-glycosylation plays an important role during EAC.

## Conclusion

MALDI-MSI of N-glycans was successfully applied to 24 FFPE human esophageal resections spanning the different stages of esophageal adenocarcinogenesis to gain more insight into N-glycosylation changes during disease progression. As EAC is a heterogeneous disease, one patient per stage was imaged in 3D first to determine the number of sections needed to be representative for the whole volume. We found that 2-3 sections were needed to be representative leading to 4 sections being used for the 2.5D-MSI of the remaining patients. These data were subsequently used to identify differentially expressed N-glycans between the different EAC stages. In total 3 N-glycans were found to be more abundant in the carcinoma stages whose presence could be explained by meta-analyses of publically available gene expression data. This evidence at multiple levels confirms that specific N-glycosylation plays an important role during EAC.

## Acknowledgments

This work has been made possible with the financial support of the Dutch province of Limburg through the LINK program, ITEA, and RVO by means of project numbers ITEA151003/ITEA14001 as well as funding from the National Institutes of Health (NIH) of the United States of America, Grant No. NIH R01 CA213492. B.B acknowledges the financial support from the European Union (ERA-NET TRANSCAN 2; Grant No. 643638).



# ***Chapter Five***

## **Valorization**



Biological tissues are molecularly complex volumetric systems. To understand these tissues techniques are needed that can capture this complexity while maintaining the spatial information. This demands innovative new ways to increase dimensionality, both molecular and spatial, of existing techniques.

Increasing dimensionality is not only needed for research to gain a deeper understanding of complex living systems, in the industry this increase is needed to provide a better diagnosis to patients. A complete tissue gives a better view of the spread of the disease and its severity based on the molecular information. The work presented in this thesis formed a central part of the ITEA 3D Pathology project where the aim was to determine if we can valorize this increased dimensionality by moving 3D pathology to the clinic. In this light, we set out to not only investigate the potential of moving from 2D H&E, which is the standard cell-based analysis used in pathology, to 3D H&E but also the addition of molecular information using 3D-MSI. For this purpose, the consortium consisted of a variety of members focusing on either acquiring the data, data analysis, visualization of the data, as well as setting up a platform to teach pathologists in training how to work and interpret 3D data. All these aspects are of importance and require the collaboration of scientists, companies, and pathologists if 3D pathology is to be established in the clinic. This is how the impact of MSI can be increased beyond the fundamental insights it can provide and provides the possibility to valorize the research described in this thesis.

The scope of this thesis was to investigate the 3D capabilities of MSI and its potential for use in the clinic. Several steps have been added to the valorization potential of this work. The first important step towards clinical translation is to increase the compatibility of MSI with clinical tissues. These tissues are often FFPE tissues that require new sample preparation methods to be developed to allow molecules to be imaged that are still reflective of the biology. Preferably, these sample preparations not only retain the biological information but are also fast to conform to the standard pathology times of analyzing tissues and provide a diagnosis to the patient. In **chapter 2** we, therefore, investigated the effect formalin fixation has on the lipidome of tissues. These results will aid clinicians in better understanding the lipid data acquired from formalin-fixed tissues.

The next phase towards the translation of MSI to the clinic is having established data analysis methods that allow a uniform analysis of the data across different clinics. This is needed to prevent a patient from obtaining diagnosis X at clinic A while with the same data gets diagnosis Y at clinic B. Within M4I we therefore set-out to develop strategies to manage multi-patient data that are presented in **chapter 3**. Our proposed methods ensure that the data used for analysis are without outliers. This is important in a clinic as outliers can steer model-based diagnosis into a false direction thereby reducing the accuracy of diagnosis using this type of data. In addition, the pipeline can be easily transferred to other types of 3D-MSI data sets acquired at different labs. As it is also applicable to single patient 3D-MSI studies or even multi-patient 2D-MSI studies, it makes an ideal starting point towards a uniform data analysis pipeline for clinical 3D-MSI data. This universal application of our method, which is demonstrated in **chapter 4**, is not only interesting in a clinical setting but also for companies when they are considering providing 3D-MSI services.

The previously mentioned aspects are all important to translate 3D-MSI to the clinic. Its true implementation, however, requires integration with the pathologist standard method, which is H&E staining. Within our consortium, we have closely interacted with several pathologists and they also stated that the only way they envision 3D-MSI in their labs is when it can be integrated with their H&E's placed next to it and co-registered to allow regions of interest to be quickly defined.

In view of the above, we collaborated with our consortium partners at Phillips to visualize our MSI data together with the H&E data in their digital pathology system. This was successful and allows both data to be investigated at the same time, however, the data is visualized as single sections instead of a complete 3D volume. As the goal was to integrate both modalities as 3D volumes, we additionally collaborated with PS-Medtech who was also a part of the ITEA 3D Pathology consortium. They provide 3D visualization workstations that are already implemented in hospitals to visualize and interact with 3D CT and MRI data. This interaction is an added value as it will allow the radiologist to explore the 3D image and see what changes. We first set-out to read MSI data into their workstation and

interactive visualize it. Once this was in place we worked towards the end goal of the ITEA 3D pathology, namely visualize a 3D H&E volume next to a 3D MSI volume for the clinical pathology community. This was challenging as two different modalities had to be placed into one coordinate system. In addition, they have to be in the same orientation and coordinate system to ensure the H&E section and the MSI section are the same ones. If these challenges could be overcome we would have everything in place to bring 3D Pathology, and thus 3D-MSI, to the clinic. To this end, a 3D H&E volume was generated at the AMC and with M4I we provided the corresponding 3D-MSI volume. At the end of the ITEA project, we were able to visualize both volumes in one space and interactively go through both volumes simultaneously.

This shows that 3D Pathology can be translated to the clinic but it requires interdisciplinary cooperation of scientists, pathologists, and companies. Only when they all work together across the boundaries of their own disciplines and interests, 3D-MSI analysis can become applicable in clinical diagnostic and provided to pathologists to improve their diagnostic toolbox.



# ***Epilogue***

## **General discussion and Summary (Samenvatting)**



Mass spectrometry imaging (MSI) has seen widespread use in clinical research to study the molecular complexity of tissue on a spatial level. These have contributed to new insights into tumor heterogeneity and molecular changes that happen before their effect is visible on a cellular level. This latter can aid in obtaining clean margins during tumor resection. Although these studies all provide valuable information they are often performed on a two-dimensional scale where one or a few sections are used. These 2D studies are prone to sampling bias as the molecular complexity of biomedical tissues cannot be captured in a few sections, especially when they are all from the same sampling site. To thoroughly capture the molecular complexity of a single tissue it should be investigated as a whole. The development of faster instrumentation and automated sample preparation devices has led MSI to do exactly this by the emergence of 3D-MSI. Current 3D-MSI studies are mostly done on a single biomedical tissue as these studies require often extensive sample preparation and measurement time, even with the recent advancements. However, to gain a deeper understanding of biological processes underlying diseases, multi-patient 3D-MSI studies are needed. The work presented in this thesis is focused on the sample preparation and data analysis strategies that are needed to enable 3D-MSI of biomedical tissues to be applied to multi-patient studies.

## Sample preparation

Biomedical tissues are often formalin-fixed and paraffin-embedded (FFPE) as this allows for long-term storage. This required new sample preparation methods to make MSI compatible with these types of tissues. Over the years protocols have been developed and applied for 2D-MSI studies. However, these protocols are often long and require enzymatic digestion of the tissue. Applying this to 3D-MSI studies, and especially multi-patient studies, make that acquisition of the data will take a long time maybe even up to months. A less extensive sample preparation, and consequently quicker, is therefore preferred.

The introduction of a new method that allows for lipid detection and only requires a quick washing step and matrix application seems to be promising for speeding up 3D-MSI of FFPE tissues. To be able to apply this for 3D-MSI,

or even 2D-MSI studies, it has to be certain that the obtained lipidome is reflective of the tissues. We found by doing an in-depth study that class-specific changes are happening during formalin-fixation. In particular the cross-linking of amine-containing lipids resulting in depletion of these molecules. Taken into account that these tissues would also undergo paraffin-embedding, which can result in a loss of lipids, it is hard to say how reflective the lipidome still is. Even though it can still be used to differentiate different disease states the biological meaning of the lipids are hard to discern and need to be rationalized. When antigen retrieval can be optimized for lipids, thereby undoing the cross-linking, the amine-containing lipids will be detectable again and the lipidome will be (more) reflective of the tissue biology.

As lipids are not yet suitable for 3D-MSI studies, enzymatic digestion of the tissues is needed to obtain biological reflective molecules that can be imaged. These digestions involve a lot of manual sample handling which can lead to technical variation. In addition, the efficiency of the enzymatic digestion can vary per section. This has already been shown in 2D-MSI studies and makes reproducibility a challenge, though is often achieved as 2D-MSI allows for multiple sections from different patients to be prepared in a single day.

### Quality controls and data analysis

Reproducibility is an even bigger impediment in 3D-MSI studies, where at most one single tissue can be prepared in one go given that the amount of sections is not too high and different tissues are prepared on different days. When the 3D dataset is too large a single tissue will even be prepared on different days. This spread in preparation and acquisition can lead to technical variations and consequential outliers. It is therefore imperative that sample preparation and data analysis strategies are adapted to ensure reproducibility and removal of outliers. One way to achieve this is by incorporating quality controls during sample preparation to monitor important steps in the protocol prone to variation. This predicament was solved by incorporating a quality control before trypsin application which was the enzyme used to digest 14 FFPE bladder cancer resections. The quality control allowed us to track the digestion efficiency on a slide basis

and the removal of those slides with a too high variance in digestion efficiency.

Incorporation of these quality controls is a great step forward in realizing multi-patient 3D-MSI studies. It allows the removal of outliers thereby increasing the confidence that variations found in a single 3D volume or between patients are biological. As sample preparation is paving the way, new data analysis strategies are needed to finish it. Established data analyses are insufficient to handle the large amounts of data generated in 3D-MSI studies. To extract the wealth of information in these studies, data analysis strategies are needed that first remove outliers and next reduce the amount of data to those molecules that are biologically interesting, e.g. those molecules that are different between disease states, or specific for a certain histological region.

One way to incorporate outlier removal during data analysis can be by making use of incorporated quality controls during sample preparation as has been shown by the usage of cytochrome *c*. These kinds of controls are useful for tracking digestion outliers, however, they do not indicate outliers due to other technical variations. Data analysis should, therefore, incorporate other strategies as well to pick these out. We have overcome this conundrum by developing an outlier detection method that uses the *z*-regression of single masses as each mass will follow a certain trend in the *z*-direction. Beneficial of this kind of method is that it depends only on the intensities of the molecules in the single sections, thereby picking out both outliers due to sample preparation variation as well as those due to instrument variation. Although this is a robust strategy, along with quality controls, in removing all possible outliers the downside is that it only works when molecules are present throughout the whole volume. When volume sizes get bigger or more heterogeneous, and molecules are present in only parts of the tissues, this strategy will not work and different detection methods need to be established.

## Necessity 3D-MSI

As strategies are slowly developing to analyze 3D-MSI data, and potentially venture into combination with other imaging modalities, it is becoming the new trend in MSI. This increase in 3D-MSI data also begs the question of whether 3D-MSI is needed and is worth the effort. Yes, biomedical tissues are molecularly complex and heterogeneous volumes but is imaging of the whole volume the only option to gain a complete overview of this complexity and avoid sampling bias. An investigation into the sampling bias indeed shows that restricting an investigation to one single tissue section induces sampling bias and can greatly affect the discovery of new biomarkers, highlighting once again the need to measure more sections per sample to overcome this. However, this does not necessarily mean the whole tissue as was observed after a representativeness analysis which gave a minimum number of sections needed to be molecularly representative of the whole tissue. This number was less than 50% of the total volume for two different datasets consisting of two different types of tissue (bladder cancer and esophageal adenocarcinoma) and molecules of interest (peptides and N-glycans) providing evidence that 3D-MSI might not be needed to map the molecular heterogeneity and something in between, like 2.5D MSI, is sufficient. Indeed there might be reasons for 3D-MSI, other than mapping the molecular complexity that does require a complete volume to be investigated. However, this has so far not been shown. Although 3D-MSI is thus a hot and shiny new thing, its added value over a more extensive 2D-MSI, like 2.5D-MSI, remains to be seen.

## Summary

Mass spectrometry imaging (MSI) has become an established tool to spatially resolve the molecular make-up of biomedical tissues. Advancements in sample preparation and instrumentation facilitated the development of 3D-MSI to gain a deeper understanding of the spatially and molecular complexity of biological processes in tissues. In this thesis, the challenges in applying 3D-MSI in biomedical research are appraised and addressed. The newly developed methods and insights are subsequently applied to multi-patient cohort studies.

In **chapter 1** the readers are introduced to the field of 3D-MSI and its current challenges. These challenges are focused on two major fields: sample preparation and data analysis. For both, aspects are discussed that have to be considered when performing 3D-MSI experiments as well as the obstacles that still need to be resolved. One of these aspects is how to prepare single-cells for 3D-SIMS or serial sections for 3D-DESI and 3D-MALDI MSI to enable 3D reconstruction. The experimental design is also discussed and how to ensure reproducibility, which follows into outlier detection as one of the challenges to be addressed during data analysis. In addition, 3D image reconstruction and multimodal 3D-imaging are reviewed. Overall, this chapter provides considerations that aid in setting-up optimal (multi-patient) 3D-MSI studies.

**Chapter 2** focuses on the sample preparation of FFPE samples and whether lipids can be a possible new mine to dig for 3D-MSI of these biomedical tissues. An in-depth analysis of the lipid classes detectable from fresh frozen and formalin-fixed mouse and rat kidney tissues gave noticeable differences. Amine-containing lipids are completely depleted upon fixation which is most likely a result of cross-linking of these lipids with each other and with amine-containing proteins. As a result of this depletion, an increase in sulfatides and cholesterol is observed. These class-specific lipid changes upon formalin fixation show that care has to be taken when biological statements are made based on lipid data from FFPE tissues.

In **chapter 3** some of the challenges in the analysis of 3D-MSI data are being tackled by setting up a pipeline for handling multi-patient 3D-MSI data. We first acquired 14 3D-MSI datasets of human FFPE bladder cancer resections by enzymatic digestion with trypsin. This data was used to develop novel outlier detection methods that are based on quality controls incorporated in the sample preparation as well as the acquired data itself. Using cytochrome *c* on every slide as quality control to monitor digestion efficiency enables the identification of those slides on which digestion performed insufficiently. The other outlier method is based on the z-directed regression of the single masses thereby identifying those outliers based on other technical variations. Combined, these two methods resulted in 56 sections being marked as outliers out of the total 280 sections measured, giving a realistic reproducibility rate of 80%. Besides these, a method was developed to determine the minimum amount of sections needed to be molecularly representative of the whole tissue. This showed that for this type of sample 33% of the sections is sufficient to be representative of the whole tissue. This strategy can be transferred to other studies and aid in reducing the workload of 3D-MSI studies.

This expertise is applied in **chapter 4** where N-glycan changes are investigated during esophageal adeno-carcinogenesis. To this end, 24 human FFPE esophageal resections have been digested using PNGase F to obtain N-glycosylation profiles. Due to the heterogeneous behavior of disease progression, one patient per stage was first imaged in 3D and the number of representative sections was determined. This showed that 4 sections are required and subsequently used for the remaining patients and called 2.5D-MSI. Analysis of the 2.5D-MSI data marked 3 N-glycans as more abundant in adenocarcinoma whose presence could be explained by publically available gene expression data. This combination of data demonstrates that specific N-glycosylation plays an important role and both MSI and gene data should be combined to gain a deeper understanding of this process.

The work presented in this thesis all contributes to one major valorization aspect portrayed in **chapter 5**. The potential of implementing 3D-MSI in the clinic as part of 3D pathology is described and requires collaboration between researchers, pathologists, and companies.

## Samenvatting

Massaspectrometrie imaging (MSI) is een gangbare techniek geworden om de moleculaire samenstelling van biomedische weefsels ruimtelijk vast te leggen. Verbeteringen in monster voorbereiding en instrumentatie bevorderde de ontwikkeling van 3D-MSI om een dieper inzicht te verkrijgen in de ruimtelijke en moleculaire complexiteit van biologische processen in weefsels. In dit proefschrift worden de uitdagingen bij het toepassen van 3D-MSI in biomedisch onderzoek beoordeeld en aangepakt. De nieuw ontwikkelde methoden en vergaarde inzichten zijn vervolgens toegepast in cohortonderzoeken met meerdere patiënten.

In **hoofdstuk 1** maken de lezers kennis met het veld van 3D-MSI en zijn huidige uitdagingen. Deze uitdagingen zijn gericht op twee grote gebieden; monster voorbereiding en data-analyse. Voor beide worden die aspecten behandeld die in consideratie genomen moeten worden wanneer een 3D-MSI experiment wordt uitgevoerd als mede de knelpunten die nog moeten worden opgelost. Eén van die aspecten is hoe een enkele cel voor 3D-SIMS of seriële secties voor 3D-DESI en 3D-MALDI geprepareerd moet worden om 3D reconstructie mogelijk te maken. De experimentele opzet en hoe reproduceerbaarheid gewaarborgd kan worden wordt ook bediscussieerd. Dit vloeit uit in detectie van uitbijters als één van de uitdaging die geadresseerd moet worden tijdens data-analyse. Daarnaast worden 3D beeldreconstructie en multimodaal 3D-imaging beoordeeld. Globaal biedt dit hoofdstuk overwegingen die helpen bij het opzetten van (multi-patiënt) 3D-MSI onderzoeken.

**Hoofdstuk 2** focust zich op monster voorbereiding van FFPE monsters en de vraag of een mogelijk nieuwe manier om 3D-MSI van deze biomedische weefsels te versnellen de detectie van lipiden is. Een uitgebreide analyse van de detecteerbare lipidenklassen van vers ingevroren en formaline gefixeerde nierweefsels van muis en rat leverde aanmerkelijke verschillen op. Amine-bevattende lipiden zijn volledig afgenomen en is waarschijnlijk het gevolg van verweving van deze lipiden met zichzelf als mede met amine-bevattende eiwitten. Door deze afname is een toename in sulfaten en cholesterol geconstateerd. Deze klassen-specifieke lipide verandering door

formaline fixatie laat zien dat voorzichtigheid in acht moet worden genomen wanneer biologische verklaringen worden genomen gebaseerd op lipide data van FFPE weefsels.

In **hoofdstuk 3** worden enkele uitdagingen in de data-analyse van 3D-MSI data aangepakt door het opzetten van een procedure voor het aanpakken van multi-patiënt 3D-MSI data. Eerst hebben we 14 3D-MSI gegevensbestanden van menselijk FFPE blaaskanker resecties vergaard door enzymatische digestie met trypsine. De verkregen data werd gebruikt om nieuwe detectiemethoden voor uitbijters te ontwikkelen die gebaseerd zijn op kwaliteitscontroles geïntrigeerd in de monster voorbereiding als ook de verweven data zelf. Gebruik maken van cytochroom *c* op elk glaasje als kwaliteitscontrole voor digestie efficiëntie faciliteerde de identificatie van die glaasjes waar dit ontoereikend was. De andere uitbijter detectiemethode is gebaseerd op de regressie van individuele massa's in de z-directie waarbij uitbijters, ontstaan door technische variatie, worden aangemerkt. Gecombineerd wezen deze methoden 56 secties van de in totaal gemeten 280 secties aan als uitbijter wat resulteert in een reproduceerbaarheid van 80%. Daarnaast was er een methode ontwikkeld om het minimale aantal secties te bepalen dat nodig is om op een moleculair niveau representatief te zijn voor het gehele weefsel. Dit gaf aan dat voor onze monsters 33% van de secties voldoende is om representatief te zijn voor het gehele weefsel. Deze strategie kan worden overgebracht naar andere onderzoeken en zo bijdragen aan het verminderen van de werklust van 3D-MSI onderzoeken.

**Hoofdstuk 4** ziet de toepassing van de eerder ontwikkelde data-analyse methoden in een biologisch onderzoek naar N-glycan verandering tijdens slokdarm adeno-carcinogenese. Hiervoor zijn 24 menselijke FFPE slokdarm resecties enzymatisch gedigesteerd met PNGase F om N-glycans te verkrijgen. Vanwege het heterogene karakter van de ziekteprogressie is één patiënt per stadium eerst gemeten in 3D en het aantal representatieve secties bepaald. Dit wees uit dat 4 secties nodig zijn en dit is vervolgens gebruikt voor de resterende patiënten en wordt 2.5D MSI genoemd. Analyse van de 2.5D-MSI data duidde 3 N-glycans aan die meer aanwezig waren in adenocarcinomen en waarvan de aanwezigheid verklaard kon worden door publiek toegankelijk genexpressie data. Deze combinatie van data toont aan dat specifieke N-glycosylering een belangrijke rol speelt en dat MSI en

genetische data gecombineerd moet worden om een dieper inzicht te vergaren in dit proces.

Het werk dat gepresenteerd is in dit proefschrift draagt allemaal bij aan één belangrijk valorisatie aspect weergegeven in hoofdstuk 5. De realiseerbaarheid om 3D-MSI in de kliniek te implementeren als onderdeel van 3D pathologie is beschreven en vereist de samenwerking van onderzoekers, pathologen, en bedrijven.



# ***References***



1. Sabatini, U., K. Boulanouar, N. Fabre, F. Martin, C. Carel, C. Colonnese, L. Bozzao, I. Berry, J.L. Montastruc, F. Chollet, and O. Rascol, *Cortical motor reorganization in akinetic patients with Parkinson's disease: a functional MRI study*. *Brain*, 2000. **123 ( Pt 2)**: p. 394-403.
2. Staals, J., S.D. Makin, F.N. Doubal, M.S. Dennis, and J.M. Wardlaw, *Stroke subtype, vascular risk factors, and total MRI brain small-vessel disease burden*. *Neurology*, 2014. **83(14)**: p. 1228-34.
3. Donners, M.M., M.J. Verluyten, F.G. Bouwman, E.C. Mariman, B. Devreese, F. Vanrobaeys, J. van Beeumen, L.H. van den Akker, M.J. Daemen, and S. Heeneman, *Proteomic analysis of differential protein expression in human atherosclerotic plaque progression*. *J Pathol*, 2005. **206(1)**: p. 39-45.
4. Li, C., Z. Chen, Z. Xiao, X. Wu, X. Zhan, X. Zhang, M. Li, J. Li, X. Feng, S. Liang, P. Chen, and J.Y. Xie, *Comparative proteomics analysis of human lung squamous carcinoma*. *Biochem Biophys Res Commun*, 2003. **309(1)**: p. 253-60.
5. Hicks, D.G., B.J. Yoder, J. Pettay, E. Swain, S. Tarr, M. Hartke, M. Skacel, J.P. Crowe, G.T. Budd, and R.R. Tubbs, *The incidence of topoisomerase II-alpha genomic alterations in adenocarcinoma of the breast and their relationship to human epidermal growth factor receptor-2 gene amplification: a fluorescence in situ hybridization study*. *Hum Pathol*, 2005. **36(4)**: p. 348-56.
6. Wang, S.X., G. Cai, and S. Sui, *Intrinsic fluorescence study of the interaction of human apolipoprotein H with phospholipid vesicles*. *Biochemistry*, 1999. **38(29)**: p. 9477-84.
7. van Hove, E.R.A., D.F. Smith, and R.M.A. Heeren, *A concise review of mass spectrometry imaging*. *Journal of Chromatography A*, 2010. **1217(25)**: p. 3946-3954.
8. Watrous, J.D., T. Alexandrov, and P.C. Dorrestein, *The evolving field of imaging mass spectrometry and its impact on future biological research*. *Journal of Mass Spectrometry*, 2011. **46(2)**: p. 209-222.
9. McDonnell, L.A. and R.M.A. Heeren, *Imaging mass spectrometry*. *Mass Spectrometry Reviews*, 2007. **26(4)**: p. 606-643.
10. Bodzon-Kulakowska, A. and P. Suder, *IMAGING MASS SPECTROMETRY: INSTRUMENTATION, APPLICATIONS, AND COMBINATION WITH OTHER VISUALIZATION TECHNIQUES*. *Mass Spectrometry Reviews*, 2016. **35(1)**: p. 147-169.
11. Hagenhoff, B., *High Resolution Surface Analysis by TOF-SIMS*. *Microchimica Acta*, 2000. **132(2)**: p. 259-271.

12. Weibel, D., S. Wong, N. Lockyer, P. Blenkinsopp, R. Hill, and J.C. Vickerman, *A C60 primary ion beam system for time of flight secondary ion mass spectrometry: its development and secondary ion yield characteristics*. *Anal Chem*, 2003. **75**(7): p. 1754-64.
13. Kollmer, F., *Cluster primary ion bombardment of organic materials*. *Applied Surface Science*, 2004. **231-232**: p. 153-158.
14. Eberlin, L.S., K. Margulis, I. Planell-Mendez, R.N. Zare, R. Tibshirani, T.A. Longacre, M. Jalali, J.A. Norton, and G.A. Poultsides, *Pancreatic Cancer Surgical Resection Margins: Molecular Assessment by Mass Spectrometry Imaging*. *PLoS Med*, 2016. **13**(8): p. e1002108.
15. Ogrinc Potocnik, N., T. Porta, M. Becker, R.M.A. Heeren, and S.R. Ellis, *Use of advantageous, volatile matrices enabled by next-generation high-speed matrix-assisted laser desorption/ionization time-of-flight imaging employing a scanning laser beam*. *Rapid communications in mass spectrometry : RCM*, 2015. **29**(23): p. 2195-203.
16. Spengler, B. and M. Hubert, *Scanning microprobe matrix-assisted laser desorption ionization (SMALDI) mass spectrometry: Instrumentation for sub-micrometer resolved LDI and MALDI surface analysis*. *Journal of the American Society for Mass Spectrometry*, 2002. **13**(6): p. 735-748.
17. Heeren, R.M.A., B. Kukrer-Kaletas, I.M. Taban, L. MacAleese, and L.A. McDonnell, *Quality of surface: The influence of sample preparation on MS-based biomolecular tissue imaging with MALDI-MS and (ME-)SIMS*. *Applied Surface Science*, 2008. **255**(4): p. 1289-1297.
18. Thavarajah, R., V.K. Mudimbaimannar, J. Elizabeth, U.K. Rao, and K. Ranganathan, *Chemical and physical basics of routine formaldehyde fixation*. *J Oral Maxillofac Pathol*, 2012. **16**(3): p. 400-5.
19. Goodwin, R.J.A., *Sample preparation for mass spectrometry imaging: Small mistakes can lead to big consequences*. *Journal of Proteomics*, 2012. **75**(16): p. 4893-4911.
20. Groseclose, M.R., P.P. Massion, P. Chaurand, and R.M. Caprioli, *High-throughput proteomic analysis of formalin-fixed paraffin-embedded tissue microarrays using MALDI imaging mass spectrometry*. *Proteomics*, 2008. **8**(18): p. 3715-24.
21. Powers, T.W., B.A. Neely, Y. Shao, H. Tang, D.A. Troyer, A.S. Mehta, B.B. Haab, and R.R. Drake, *MALDI Imaging Mass Spectrometry Profiling of N-Glycans in Formalin-Fixed Paraffin Embedded Clinical Tissue Blocks and Tissue Microarrays*. *PLOS ONE*, 2014. **9**(9): p. e106255.
22. Buck, A., A. Ly, B. Balluff, N. Sun, K. Gorzolka, A. Feuchtinger, K.P. Janssen, P.J. Kuppen, C.J. van de Velde, G. Weirich, F. Erlmeier, R. Langer, M. Aubele, H. Zitzelsberger, M. Aichler, and A. Walch, *High-*

- resolution MALDI-FT-ICR MS imaging for the analysis of metabolites from formalin-fixed, paraffin-embedded clinical tissue samples.* J Pathol, 2015. **237**(1): p. 123-32.
23. Lawrence, M.S., P. Stojanov, P. Polak, G.V. Kryukov, K. Cibulskis, A. Sivachenko, S.L. Carter, C. Stewart, C.H. Mermel, S.A. Roberts, A. Kiezun, P.S. Hammerman, A. McKenna, Y. Drier, L. Zou, A.H. Ramos, T.J. Pugh, N. Stransky, E. Helman, J. Kim, C. Sougnez, L. Ambrogio, E. Nickerson, E. Shefler, M.L. Cortes, D. Auclair, G. Saksena, D. Voet, M. Noble, D. DiCara, P. Lin, L. Lichtenstein, D.I. Heiman, T. Fennell, M. Imielinski, B. Hernandez, E. Hodis, S. Baca, A.M. Dulak, J. Lohr, D.A. Landau, C.J. Wu, J. Melendez-Zajgla, A. Hidalgo-Miranda, A. Koren, S.A. McCarroll, J. Mora, B. Crompton, R. Onofrio, M. Parkin, W. Winckler, K. Ardlie, S.B. Gabriel, C.W.M. Roberts, J.A. Biegel, K. Stegmaier, A.J. Bass, L.A. Garraway, M. Meyerson, T.R. Golub, D.A. Gordenin, S. Sunyaev, E.S. Lander, and G. Getz, *Mutational heterogeneity in cancer and the search for new cancer-associated genes.* Nature, 2013. **499**(7457): p. 214-218.
  24. Balluff, B., C.K. Frese, S.K. Maier, C. Schone, B. Kuster, M. Schmitt, M. Aubele, H. Hofler, A.M. Deelder, A. Heck, Jr., P.C. Hogendoorn, J. Morreau, A.F. Maarten Altelaar, A. Walch, and L.A. McDonnell, *De novo discovery of phenotypic intratumour heterogeneity using imaging mass spectrometry.* J Pathol, 2015. **235**(1): p. 3-13.
  25. Pacholski, M.L. and N. Winograd, *Imaging with Mass Spectrometry.* Chemical Reviews, 1999. **99**(10): p. 2977-3006.
  26. Paine, M.R.L., P.C. Kooijman, G.L. Fisher, R.M.A. Heeren, F.M. Fernández, and S.R. Ellis, *Visualizing molecular distributions for biomaterials applications with mass spectrometry imaging: a review.* Journal of Materials Chemistry B, 2017. **5**(36): p. 7444-7460.
  27. Vaysse, P.M., R.M.A. Heeren, T. Porta, and B. Balluff, *Mass spectrometry imaging for clinical research - latest developments, applications, and current limitations.* Analyst, 2017. **142**(15): p. 2690-2712.
  28. Heeren, R.M.A., *Getting the picture: The coming of age of imaging MS.* International Journal of Mass Spectrometry, 2015. **377**: p. 672-680.
  29. McPhail, D.S., *Applications of Secondary Ion Mass Spectrometry (SIMS) in Materials Science.* Journal of Materials Science, 2006. **41**(3): p. 873-903.
  30. Yang, J. and I. Gilmore, *Application of secondary ion mass spectrometry to biomaterials, proteins and cells: a concise review.* Materials Science and Technology, 2015. **31**(2): p. 131-136.

31. Goor, O.J., H.M. Keizer, A.L. Bruinen, M.G. Schmitz, R.M. Versteegen, H.M. Janssen, R.M. Heeren, and P.Y. Dankers, *Efficient Functionalization of Additives at Supramolecular Material Surfaces*. *Adv Mater*, 2017. **29**(5).
32. Brunner, C., H. Hutter, P. Wilhartitz, and M. Grasserbauer, *Investigation of the formation and properties of protective oxide layers on high purity chromium with SIMS imaging techniques*. *Microchimica Acta*, 1997. **125**(1): p. 69-72.
33. Wilhartitz, P., R. Krismer, H. Hutter, M. Grasserbauer, S. Weinbruch, and H.M. Ortner, *3D-SIMS analysis of ultra high purity molybdenum and tungsten: a characterisation of different manufacturing techniques and products*. *Fresenius' Journal of Analytical Chemistry*, 1995. **353**(5): p. 524-532.
34. Fletcher, J.S. and J.C. Vickerman, *A new SIMS paradigm for 2D and 3D molecular imaging of bio-systems*. *Anal Bioanal Chem*, 2010. **396**(1): p. 85-104.
35. Clerc, J., C. Fourre, and P. Fragu, *SIMS microscopy: methodology, problems and perspectives in mapping drugs and nuclear medicine compounds*. *Cell Biol Int*, 1997. **21**(10): p. 619-33.
36. John, C.M., R.W. Odom, L. Salvati, A. Annapragada, and M.Y. Fu Lu, *XPS and TOF-SIMS microanalysis of a peptide/polymer drug delivery device*. *Anal Chem*, 1995. **67**(21): p. 3871-8.
37. Brown, D.L., *Practical Stereology Applications for the Pathologist*. *Vet Pathol*, 2017. **54**(3): p. 358-368.
38. Brown, D.L., *Bias in image analysis and its solution: unbiased stereology*. *J Toxicol Pathol*, 2017. **30**(3): p. 183-191.
39. Guo, T., L. Li, Q. Zhong, N.J. Rupp, K. Charmpi, C.E. Wong, U. Wagner, J.H. Rueschoff, W. Jochum, C.D. Fankhauser, K. Saba, C. Poyet, P.J. Wild, R. Aebersold, and A. Beyer, *Multi-region proteome analysis quantifies spatial heterogeneity of prostate tissue biomarkers*. *Life Sci Alliance*, 2018. **1**(2).
40. Lambin, P., E. Rios-Velazquez, R. Leijenaar, S. Carvalho, R.G. van Stiphout, P. Granton, C.M. Zegers, R. Gillies, R. Boellard, A. Dekker, and H.J. Aerts, *Radiomics: extracting more information from medical images using advanced feature analysis*. *Eur J Cancer*, 2012. **48**(4): p. 441-6.
41. Lein, E.S., M.J. Hawrylycz, N. Ao, M. Ayres, A. Bensinger, A. Bernard, A.F. Boe, M.S. Boguski, K.S. Brockway, E.J. Byrnes, L. Chen, L. Chen, T.M. Chen, M.C. Chin, J. Chong, B.E. Crook, A. Czaplinska, C.N. Dang, S. Datta, N.R. Dee, A.L. Desaki, T. Desta, E. Diep, T.A.

- Dolbeare, M.J. Donelan, H.W. Dong, J.G. Dougherty, B.J. Duncan, A.J. Ebbert, G. Eichele, L.K. Estin, C. Faber, B.A. Facer, R. Fields, S.R. Fischer, T.P. Fliss, C. Frensley, S.N. Gates, K.J. Glattfelder, K.R. Halverson, M.R. Hart, J.G. Hohmann, M.P. Howell, D.P. Jeung, R.A. Johnson, P.T. Karr, R. Kawal, J.M. Kidney, R.H. Knapik, C.L. Kuan, J.H. Lake, A.R. Laramee, K.D. Larsen, C. Lau, T.A. Lemon, A.J. Liang, Y. Liu, L.T. Luong, J. Michaels, J.J. Morgan, R.J. Morgan, M.T. Mortrud, N.F. Mosqueda, L.L. Ng, R. Ng, G.J. Orta, C.C. Overly, T.H. Pak, S.E. Parry, S.D. Pathak, O.C. Pearson, R.B. Puchalski, Z.L. Riley, H.R. Rockett, S.A. Rowland, J.J. Royall, M.J. Ruiz, N.R. Sarno, K. Schaffnit, N.V. Shapovalova, T. Sivities, C.R. Slaughterbeck, S.C. Smith, K.A. Smith, B.I. Smith, A.J. Sodt, N.N. Stewart, K.R. Stumpf, S.M. Sunkin, M. Sutram, A. Tam, C.D. Teemer, C. Thaller, C.L. Thompson, L.R. Varnam, A. Visel, R.M. Whitlock, P.E. Wohnoutka, C.K. Wolkey, V.Y. Wong, M. Wood, M.B. Yaylaoglu, R.C. Young, B.L. Youngstrom, X.F. Yuan, B. Zhang, T.A. Zwingman and A.R. Jones, *Genome-wide atlas of gene expression in the adult mouse brain*. *Nature*, 2007. **445**(7124): p. 168-76.
42. Inglese, P., J.S. McKenzie, A. Mroz, J. Kinross, K. Veselkov, E. Holmes, Z. Takats, J.K. Nicholson, and R.C. Glen, *Deep learning and 3D-DESI imaging reveal the hidden metabolic heterogeneity of cancer*. *Chem Sci*, 2017. **8**(5): p. 3500-3511.
43. Jones, E.A., R. Shyti, R.J. van Zeijl, S.H. van Heiningen, M.D. Ferrari, A.M. Deelder, E.A. Tolner, A.M. van den Maagdenberg, and L.A. McDonnell, *Imaging mass spectrometry to visualize biomolecule distributions in mouse brain tissue following hemispheric cortical spreading depression*. *J Proteomics*, 2012. **75**(16): p. 5027-35.
44. Wucher, A., S. Sun, C. Szakal, and N. Winograd, *Molecular Depth Profiling of Histamine in Ice Using a Buckminsterfullerene Probe*. *Analytical Chemistry*, 2004. **76**(24): p. 7234-7242.
45. Cheng, J., A. Wucher, and N. Winograd, *Molecular Depth Profiling with Cluster Ion Beams*. *The Journal of Physical Chemistry B*, 2006. **110**(16): p. 8329-8336.
46. Lu, C., A. Wucher, and N. Winograd, *Molecular Depth Profiling of Buried Lipid Bilayers Using C60-Secondary Ion Mass Spectrometry*. *Analytical Chemistry*, 2011. **83**(1): p. 351-358.
47. Winograd, N., *Imaging mass spectrometry on the nanoscale with cluster ion beams*. *Anal Chem*, 2015. **87**(1): p. 328-33.

48. Lee, J.W., K.J. Kim, H.K. Kim, and D.W. Moon, *Deconvolution of SIMS depth profiles of As multiple delta layers in silicon*. Surface and Interface Analysis, 2005. **37**(2): p. 176-180.
49. Fletcher, J.S., J.C. Vickerman, and N. Winograd, *Label free biochemical 2D and 3D imaging using secondary ion mass spectrometry*. Curr Opin Chem Biol, 2011. **15**(5): p. 733-40.
50. Breitenstein, D., C.E. Rommel, R. Mollers, J. Wegener, and B. Hagenhoff, *The chemical composition of animal cells and their intracellular compartments reconstructed from 3D mass spectrometry*. Angew Chem Int Ed Engl, 2007. **46**(28): p. 5332-5.
51. Chaurand, P., *Imaging mass spectrometry of thin tissue sections: A decade of collective efforts*. Journal of Proteomics, 2012. **75**(16): p. 4883-4892.
52. Chughtai, K. and R.M.A. Heeren, *Mass Spectrometric Imaging for Biomedical Tissue Analysis*. Chemical Reviews, 2010. **110**(5): p. 3237-3277.
53. Spraggins, J.M. and R.M. Caprioli, *High-Speed MALDI-TOF Imaging Mass Spectrometry: Rapid Ion Image Acquisition and Considerations for Next Generation Instrumentation*. Journal of the American Society for Mass Spectrometry, 2011. **22**(6): p. 1022-1031.
54. Huizing, L.R.S., S.R. Ellis, B.W.A.M.M. Beulen, F.P.Y. Barré, P.B. Kwant, R.J. Vreeken, and R.M.A. Heeren, *Development and evaluation of matrix application techniques for high throughput mass spectrometry imaging of tissues in the clinic*. Clinical Mass Spectrometry, 2019. **12**: p. 7-15.
55. Campbell, D.I., C.R. Ferreira, L.S. Eberlin, and R.G. Cooks, *Improved spatial resolution in the imaging of biological tissue using desorption electrospray ionization*. Anal Bioanal Chem, 2012. **404**(2): p. 389-98.
56. Crecelius, A.C., D.S. Cornett, R.M. Caprioli, B. Williams, B.M. Dawant, and B. Bodenheimer, *Three-dimensional visualization of protein expression in mouse brain structures using imaging mass spectrometry*. J Am Soc Mass Spectrom, 2005. **16**(7): p. 1093-9.
57. Fletcher, J.S., N.P. Lockyer, S. Vaidyanathan, and J.C. Vickerman, *TOF-SIMS 3D biomolecular imaging of Xenopus laevis oocytes using buckminsterfullerene (C60) primary ions*. Anal Chem, 2007. **79**(6): p. 2199-206.
58. Eberlin, L.S., D.R. Ifa, C. Wu, and R.G. Cooks, *Three-dimensional visualization of mouse brain by lipid analysis using ambient ionization mass spectrometry*. Angew Chem Int Ed Engl, 2010. **49**(5): p. 873-6.
59. Fletcher, J.S., *Latest applications of 3D ToF-SIMS bio-imaging*. Biointerphases, 2015. **10**(1): p. 018902.

60. Fisher, G.L., A.M. Belu, C.M. Mahoney, K. Wormuth, and N. Sanada, *Three-dimensional time-of-flight secondary ion mass spectrometry imaging of a pharmaceutical in a coronary stent coating as a function of elution time*. *Anal Chem*, 2009. **81**(24): p. 9930-40.
61. Starr, N.J., K. Abdul Hamid, J. Wibawa, I. Marlow, M. Bell, L. Perez-Garcia, D.A. Barrett, and D.J. Scurr, *Enhanced vitamin C skin permeation from supramolecular hydrogels, illustrated using in situ ToF-SIMS 3D chemical profiling*. *Int J Pharm*, 2019. **563**: p. 21-29.
62. Seeley, E.H. and R.M. Caprioli, *3D imaging by mass spectrometry: a new frontier*. *Anal Chem*, 2012. **84**(5): p. 2105-10.
63. Scott, A.J., J.M. Post, R. Lerner, S.R. Ellis, J. Lieberman, K.A. Shirey, R.M.A. Heeren, L. Bindila, and R.K. Ernst, *Host-based lipid inflammation drives pathogenesis in Francisella infection*. *Proc Natl Acad Sci U S A*, 2017. **114**(47): p. 12596-12601.
64. Jones, E.E., C. Quiason, S. Dale, and S.K. Shahidi-Latham, *Feasibility Assessment of a MALDI FTICR Imaging Approach for the 3D Reconstruction of a Mouse Lung*. *J Am Soc Mass Spectrom*, 2017.
65. Verbeeck, N., J. Yang, B. De Moor, R.M. Caprioli, E. Waelkens, and R. Van de Plas, *Automated anatomical interpretation of ion distributions in tissue: linking imaging mass spectrometry to curated atlases*. *Anal Chem*, 2014. **86**(18): p. 8974-82.
66. Abdelmoula, W.M., R.J. Carreira, R. Shyti, B. Balluff, R.J. van Zeijl, E.A. Tolner, B.F. Lelieveldt, A.M. van den Maagdenberg, L.A. McDonnell, and J. Dijkstra, *Automatic registration of mass spectrometry imaging data sets to the Allen brain atlas*. *Anal Chem*, 2014. **86**(8): p. 3947-54.
67. Mallah, K., J. Quanico, D. Trede, F. Kobeissy, K. Zibara, M. Salzet, and I. Fournier, *Lipid Changes Associated with Traumatic Brain Injury Revealed by 3D MALDI-MSI*. *Anal Chem*, 2018.
68. Paine, M.R.L., J. Liu, D. Huang, S.R. Ellis, D. Trede, J.H. Kobarg, R.M.A. Heeren, F.M. Fernandez, and T.J. MacDonald, *Three-Dimensional Mass Spectrometry Imaging Identifies Lipid Markers of Medulloblastoma Metastasis*. *Sci Rep*, 2019. **9**(1): p. 2205.
69. Giordano, S., L. Morosi, P. Veglianese, S.A. Licandro, R. Frapolli, M. Zucchetti, G. Cappelletti, L. Falciola, V. Pifferi, S. Visentin, M. D'Incalci, and E. Davoli, *3D Mass Spectrometry Imaging Reveals a Very Heterogeneous Drug Distribution in Tumors*. *Sci Rep*, 2016. **6**: p. 37027.
70. de Rijke, E., D. Hooijerink, S.S. Sterk, and M.W. Nielen, *Confirmation and 3D profiling of anabolic steroid esters in injection sites using imaging desorption electrospray ionisation (DESI) mass spectrometry*. *Food Addit*

- Contam Part A Chem Anal Control Expo Risk Assess, 2013. **30**(6): p. 1012-9.
71. Malm, J., D. Giannaras, M.O. Riehle, N. Gadegaard, and P. Sjovall, *Fixation and drying protocols for the preparation of cell samples for time-of-flight secondary ion mass spectrometry analysis*. *Anal Chem*, 2009. **81**(17): p. 7197-205.
72. Parry, S. and N. Winograd, *High-resolution TOF-SIMS imaging of eukaryotic cells preserved in a trehalose matrix*. *Anal Chem*, 2005. **77**(24): p. 7950-7.
73. Rabbani, S., J.S. Fletcher, N.P. Lockyer, and J.C. Vickerman, *Exploring subcellular imaging on the buncher-ToF J105 3D chemical imager*. *Surface and Interface Analysis*, 2011. **43**(1-2): p. 380-384.
74. Fletcher, J.S., S. Rabbani, A. Henderson, N.P. Lockyer, and J.C. Vickerman, *Three-dimensional mass spectral imaging of HeLa-M cells--sample preparation, data interpretation and visualisation*. *Rapid Commun Mass Spectrom*, 2011. **25**(7): p. 925-32.
75. Winograd, N. and A. Bloom, *Sample preparation for 3D SIMS chemical imaging of cells*. *Methods in molecular biology (Clifton, N.J.)*, 2015. **1203**: p. 9-19.
76. Thiele, H., S. Heldmann, D. Trede, J. Strehlow, S. Wirtz, W. Dreher, J. Berger, J. Oetjen, J.H. Kobarg, B. Fischer, and P. Maass, *2D and 3D MALDI-imaging: Conceptual strategies for visualization and data mining*. *Biochimica Et Biophysica Acta-Proteins and Proteomics*, 2014. **1844**(1): p. 117-137.
77. Norris, J.L. and R.M. Caprioli, *Analysis of Tissue Specimens by Matrix-Assisted Laser Desorption/Ionization Imaging Mass Spectrometry in Biological and Clinical Research*. *Chemical Reviews*, 2013. **113**(4): p. 2309-2342.
78. Schwartz, S.A., M.L. Reyzer, and R.M. Caprioli, *Direct tissue analysis using matrix-assisted laser desorption/ionization mass spectrometry: practical aspects of sample preparation*. *Journal of Mass Spectrometry*, 2003. **38**(7): p. 699-708.
79. Schwartz, S.A., M.L. Reyzer, and R.M. Caprioli, *Direct tissue analysis using matrix-assisted laser desorption/ionization mass spectrometry: practical aspects of sample preparation*. *J Mass Spectrom*, 2003. **38**(7): p. 699-708.
80. Chen, R., L. Hui, R.M. Sturm, and L. Li, *Three dimensional mapping of neuropeptides and lipids in crustacean brain by mass spectral imaging*. *J Am Soc Mass Spectrom*, 2009. **20**(6): p. 1068-77.

81. Chughtai, K., L. Jiang, T.R. Greenwood, I. Klinkert, E.R. Amstalden van Hove, R.M. Heeren, and K. Glunde, *Fiducial markers for combined 3-dimensional mass spectrometric and optical tissue imaging*. *Anal Chem*, 2012. **84**(4): p. 1817-23.
82. Anderson, D.M., R. Van de Plas, K.L. Rose, S. Hill, K.L. Schey, A.C. Solga, D.H. Gutmann, and R.M. Caprioli, *3-D imaging mass spectrometry of protein distributions in mouse Neurofibromatosis 1 (NF1)-associated optic glioma*. *J Proteomics*, 2016. **149**: p. 77-84.
83. Ogrinc Potocnik, N., K. Skraskova, B. Flinders, P. Pelicon, and R.M. Heeren, *Gold sputtered fiducial markers for combined secondary ion mass spectrometry and MALDI imaging of tissue samples*. *Anal Chem*, 2014. **86**(14): p. 6781-5.
84. Erich, K., D.A. Sammour, A. Marx, and C. Hopf, *Scores for standardization of on-tissue digestion of formalin-fixed paraffin-embedded tissue in MALDI-MS imaging*. *Biochim Biophys Acta*, 2017. **1865**(7): p. 907-915.
85. Vos, D.R.N., I. Jansen, M. Lucas, M.R.L. Paine, O.J. de Boer, S.L. Meijer, C.D. Savci-Heijink, H.A. Marquering, D.M. de Bruin, R.M.A. Heeren, S.R. Ellis, and B. Balluff, *Strategies for managing multi-patient 3D mass spectrometry imaging data*. *J Proteomics*, 2019. **193**: p. 184-191.
86. Buck, A., B. Heijs, B. Beine, J. Schepers, A. Cassese, R.M.A. Heeren, L.A. McDonnell, C. Henkel, A. Walch, and B. Balluff, *Round robin study of formalin-fixed paraffin-embedded tissues in mass spectrometry imaging*. *Anal Bioanal Chem*, 2018. **410**(23): p. 5969-5980.
87. Patterson, N.H., A. Thomas, and P. Chaurand, *Monitoring time-dependent degradation of phospholipids in sectioned tissues by MALDI imaging mass spectrometry*. *J Mass Spectrom*, 2014. **49**(7): p. 622-7.
88. Dill, A.L., L.S. Eberlin, A.B. Costa, D.R. Ifa, and R.G. Cooks, *Data quality in tissue analysis using desorption electrospray ionization*. *Analytical and Bioanalytical Chemistry*, 2011. **401**(6): p. 1949.
89. Duenas, M.E., J.J. Essner, and Y.J. Lee, *3D MALDI Mass Spectrometry Imaging of a Single Cell: Spatial Mapping of Lipids in the Embryonic Development of Zebrafish*. *Sci Rep*, 2017. **7**(1): p. 14946.
90. Yeager, A.N., P.K. Weber, and M.L. Kraft, *Three-dimensional imaging of cholesterol and sphingolipids within a Madin-Darby canine kidney cell*. *Biointerphases*, 2016. **11**(2): p. 02a309.
91. Fornai, L., A. Angelini, I. Klinkert, F. Giskes, A. Kiss, G. Eijkel, E.A. Amstalden-van Hove, L.A. Klerk, M. Fedrigo, G. Pieraccini, G. Moneti, M. Valente, G. Thiene, and R.M.A. Heeren, *Three-dimensional*

- molecular reconstruction of rat heart with mass spectrometry imaging*. Analytical and Bioanalytical Chemistry, 2012. **404**(10): p. 2927-2938.
92. Morosi, L., S. Giordano, F. Falcetta, R. Frapolli, S.A. Licandro, C. Matteo, M. Zucchetti, P. Ubezio, E. Erba, S. Visentin, M. D'Incalci, and E. Davoli, *Application of 3D Mass Spectrometry Imaging to TKIs*. Clin Pharmacol Ther, 2017. **102**(5): p. 748-751.
  93. Patterson, N.H., R.J. Doonan, S.S. Daskalopoulou, M. Dufresne, S. Lenglet, F. Montecucco, A. Thomas, and P. Chaurand, *Three-dimensional imaging MS of lipids in atherosclerotic plaques: Open-source methods for reconstruction and analysis*. PROTEOMICS, 2016. **16**(11-12): p. 1642-51.
  94. Schneider, C.A., W.S. Rasband, and K.W. Eliceiri, *NIH Image to ImageJ: 25 years of image analysis*. Nature Methods, 2012. **9**(7): p. 671-675.
  95. Jiang, L., T.R. Greenwood, E.R. van Hove, K. Chughtai, V. Raman, P.T. Winnard, Jr., R.M. Heeren, D. Artemov, and K. Glunde, *Combined MR, fluorescence and histology imaging strategy in a human breast tumor xenograft model*. NMR Biomed, 2013. **26**(3): p. 285-98.
  96. Lotz, J.M., F. Hoffmann, J. Lotz, S. Heldmann, D. Trede, J. Oetjen, M. Becker, G. Ernst, P. Maas, T. Alexandrov, O. Guntinas-Lichius, H. Thiele, and F. von Eggeling, *Integration of 3D multimodal imaging data of a head and neck cancer and advanced feature recognition*. Biochim Biophys Acta Proteins Proteom, 2017. **1865**(7): p. 946-956.
  97. Abdelmoula, W., M. Regan, B. Lopez, E.C. Randall, S. Lawler, A. Mladek, M. Nowicki, B. Marin, J. Agar, K. Swanson, T. Kapur, J.N. Sarkaria, W. Wells, and N.Y.R. Agar, *Automatic 3D Non-linear Registration of Mass Spectrometry Imaging and Magnetic Resonance Imaging Data*. Anal Chem, 2019.
  98. Alexandrov, T. and J.H. Kobarg, *Efficient spatial segmentation of large imaging mass spectrometry datasets with spatially aware clustering*. Bioinformatics, 2011. **27**(13): p. i230-8.
  99. Cassese, A., S.R. Ellis, N.O. Potocnik, E. Burgermeister, M. Eber, A. Walch, A. van den Maagdenberg, L.A. McDonnell, R.M.A. Heeren, and B. Balluff, *Spatial Autocorrelation in Mass Spectrometry Imaging*. Analytical Chemistry, 2016. **88**(11): p. 5871-5878.
  100. Picard de Muller, G., R. Ait-Belkacem, D. Bonnel, R. Longuespee, and J. Stauber, *Automated Morphological and Morphometric Analysis of Mass Spectrometry Imaging Data: Application to Biomarker Discovery*. J Am Soc Mass Spectrom, 2017. **28**(12): p. 2635-2645.

101. Abdelmoula, W.M., N. Pezzotti, T. Holt, J. Dijkstra, A. Vilanova, L.A. McDonnell, and B.P.F. Lelieveldt, *Interactive Visual Exploration of 3D Mass Spectrometry Imaging Data Using Hierarchical Stochastic Neighbor Embedding Reveals Spatiomolecular Structures at Full Data Resolution*. J Proteome Res, 2018. **17**(3): p. 1054-1064.
102. Krishna, C.M., G.D. Sockalingum, R.A. Bhat, L. Venteo, P. Kushtagi, M. Pluot, and M. Manfait, *FTIR and Raman microspectroscopy of normal, benign, and malignant formalin-fixed ovarian tissues*. Anal Bioanal Chem, 2007. **387**(5): p. 1649-56.
103. Krishna, C.M., G.D. Sockalingum, B.M. Vadhiraaja, K. Maheedhar, A.C. Rao, L. Rao, L. Venteo, M. Pluot, D.J. Fernandes, M.S. Vidyasagar, V.B. Kartha, and M. Manfait, *Vibrational spectroscopy studies of formalin-fixed cervix tissues*. Biopolymers, 2007. **85**(3): p. 214-21.
104. Kelly, A.D., S.B. Breitkopf, M. Yuan, J. Goldsmith, D. Spentzos, and J.M. Asara, *Metabolomic profiling from formalin-fixed, paraffin-embedded tumor tissue using targeted LC/MS/MS: application in sarcoma*. PLoS One, 2011. **6**(10): p. e25357.
105. Steiner, C., J.C. Tille, J. Lamerz, S. Kux van Geijtenbeek, T.A. McKee, M. Venturi, L. Rubbia-Brandt, D. Hochstrasser, P. Cutler, P. Lescuyer, and A. Ducret, *Quantification of HER2 by Targeted Mass Spectrometry in Formalin-Fixed Paraffin-Embedded (FFPE) Breast Cancer Tissues*. Mol Cell Proteomics, 2015. **14**(10): p. 2786-99.
106. Wojakowska, A., M. Chekan, L. Marczak, K. Polanski, D. Lange, M. Pietrowska, and P. Widlak, *Detection of metabolites discriminating subtypes of thyroid cancer: Molecular profiling of FFPE samples using the GC/MS approach*. Mol Cell Endocrinol, 2015. **417**: p. 149-57.
107. Cacciatore, S., G. Zadra, C. Bango, K.L. Penney, S. Tyekucheva, O. Yanes, and M. Loda, *Metabolic Profiling in Formalin-Fixed and Paraffin-Embedded Prostate Cancer Tissues*. Mol Cancer Res, 2017. **15**(4): p. 439-447.
108. Ly, A., A. Buck, B. Balluff, N. Sun, K. Gorzolka, A. Feuchtinger, K.P. Janssen, P.J. Kuppen, C.J. van de Velde, G. Weirich, F. Erlmeier, R. Langer, M. Aubele, H. Zitzelsberger, L. McDonnell, M. Aichler, and A. Walch, *High-mass-resolution MALDI mass spectrometry imaging of metabolites from formalin-fixed paraffin-embedded tissue*. Nat Protoc, 2016. **11**(8): p. 1428-43.
109. Gaudin, M., M. Panchal, S. Ayciriex, E. Werner, A. Brunelle, D. Touboul, C. Boursier-Neyret, N. Auzeil, B. Walther, C. Duyckaerts, and O. Laprevote, *Ultra performance liquid chromatography - mass*

- spectrometry studies of formalin-induced alterations of human brain lipidome*. *Journal of Mass Spectrometry*, 2014. **49**(10): p. 1035-1042.
110. Carter, C.L., J.W. Jones, A.M. Farese, T.J. MacVittie, and M.A. Kane, *Inflation-Fixation Method for Lipidomic Mapping of Lung Biopsies by Matrix Assisted Laser Desorption/Ionization-Mass Spectrometry Imaging*. *Analytical Chemistry*, 2016. **88**(9): p. 4788-4794.
  111. Carter, C.L., C.W. McLeod, and J. Bunch, *Imaging of Phospholipids in Formalin Fixed Rat Brain Sections by Matrix Assisted Laser Desorption/Ionization Mass Spectrometry*. *Journal of The American Society for Mass Spectrometry*, 2011. **22**(11): p. 1991-1998.
  112. Pietrowska, M., M. Gawin, J. Polanska, and P. Widlak, *Tissue fixed with formalin and processed without paraffin embedding is suitable for imaging of both peptides and lipids by MALDI-IMS*. *Proteomics*, 2016. **16**(11-12): p. 1670-1677.
  113. Addie, R.D., B. Balluff, J.V. Bovee, H. Morreau, and L.A. McDonnell, *Current State and Future Challenges of Mass Spectrometry Imaging for Clinical Research*. *Anal Chem*, 2015. **87**(13): p. 6426-33.
  114. Belov, M.E., S.R. Ellis, M. Dilillo, M.R.L. Paine, W.F. Danielson, G.A. Anderson, E.L. de Graaf, G.B. Eijkel, R.M.A. Heeren, and L.A. McDonnell, *Design and Performance of a Novel Interface for Combined Matrix-Assisted Laser Desorption Ionization at Elevated Pressure and Electrospray Ionization with Orbitrap Mass Spectrometry*. *Anal Chem*, 2017. **89**(14): p. 7493-7501.
  115. Ellis, S.R., M.R.L. Paine, G.B. Eijkel, J.K. Pauling, P. Husen, M.W. Jervelund, M. Hermansson, C.S. Ejsing, and R.M.A. Heeren, *Automated, parallel mass spectrometry imaging and structural identification of lipids*. *Nat Methods*, 2018. **15**(7): p. 515-518.
  116. Palmer, A., P. Phapale, I. Chernyavsky, R. Lavigne, D. Fay, A. Tarasov, V. Kovalev, J. Fuchser, S. Nikolenko, C. Pineau, M. Becker, and T. Alexandrov, *FDR-controlled metabolite annotation for high-resolution imaging mass spectrometry*. *Nature Methods*, 2016. **14**: p. 57.
  117. Strohm, M., M. Hassman, B. Kosata, and M. Kodicek, *mMass data miner: an open source alternative for mass spectrometric data analysis*. *Rapid Commun Mass Spectrom*, 2008. **22**(6): p. 905-8.
  118. He, L., J. Diedrich, Y.Y. Chu, and J.R. Yates, 3rd, *Extracting Accurate Precursor Information for Tandem Mass Spectra by RawConverter*. *Anal Chem*, 2015. **87**(22): p. 11361-7.
  119. Pauling, J.K., M. Hermansson, J. Hartler, K. Christiansen, S.F. Gallego, B. Peng, R. Ahrends, and C.S. Ejsing, *Proposal for a common*

- nomenclature for fragment ions in mass spectra of lipids*. PLoS One, 2017. **12**(11): p. e0188394.
120. Hsu, F.F. and J. Turk, *Studies on sulfatides by quadrupole ion-trap mass spectrometry with electrospray ionization: structural characterization and the fragmentation processes that include an unusual internal galactose residue loss and the classical charge-remote fragmentation*. J Am Soc Mass Spectrom, 2004. **15**(4): p. 536-46.
  121. Almeida, R., J.K. Pauling, E. Sokol, H.K. Hannibal-Bach, and C.S. Ejsing, *Comprehensive Lipidome Analysis by Shotgun Lipidomics on a Hybrid Quadrupole-Orbitrap-Linear Ion Trap Mass Spectrometer*. Journal of The American Society for Mass Spectrometry, 2015. **26**(1): p. 133-148.
  122. Husen, P., K. Tarasov, M. Katafiasz, E. Sokol, J. Vogt, J. Baumgart, R. Nitsch, K. Ekroos, and C.S. Ejsing, *Analysis of Lipid Experiments (ALEX): A Software Framework for Analysis of High-Resolution Shotgun Lipidomics Data*. PLOS ONE, 2013. **8**(11): p. e79736.
  123. Vens-Cappell, S., I.U. Kouzel, H. Kettling, J. Soltwisch, A. Bauwens, S. Porubsky, J. Müthing, and K. Dreisewerd, *On-Tissue Phospholipase C Digestion for Enhanced MALDI-MS Imaging of Neutral Glycosphingolipids*. Analytical Chemistry, 2016. **88**(11): p. 5595-5599.
  124. Walker, R.C. and T.J. Underwood, *Molecular pathways in the development and treatment of oesophageal cancer*. Best Pract Res Clin Gastroenterol, 2018. **36-37**: p. 9-15.
  125. Wang, H.Y., C.B. Liu, and H.W. Wu, *A simple desalting method for direct MALDI mass spectrometry profiling of tissue lipids*. J Lipid Res, 2011. **52**(4): p. 840-9.
  126. Angel, P.M., J.M. Spraggins, H.S. Baldwin, and R. Caprioli, *Enhanced sensitivity for high spatial resolution lipid analysis by negative ion mode matrix assisted laser desorption ionization imaging mass spectrometry*. Anal Chem, 2012. **84**(3): p. 1557-64.
  127. Han, X. and R.W. Gross, *Structural determination of picomole amounts of phospholipids via electrospray ionization tandem mass spectrometry*. J Am Soc Mass Spectrom, 1995. **6**(12): p. 1202-10.
  128. McDonnell, L.A., P.M. Angel, S. Lou, and R.R. Drake, *Mass Spectrometry Imaging in Cancer Research: Future Perspectives*. Adv Cancer Res, 2017. **134**: p. 283-290.
  129. Palmer, A.D. and T. Alexandrov, *Serial 3D Imaging Mass Spectrometry at Its Tipping Point*. Analytical Chemistry, 2015. **87**(8): p. 4055-4062.
  130. Gustafsson, J.O.R., M.K. Oehler, S.R. McColl, and P. Hoffmann, *Citric Acid Antigen Retrieval (CAAR) for Tryptic Peptide Imaging Directly on*

- Archived Formalin-Fixed Paraffin-Embedded Tissue*. Journal of Proteome Research, 2010. **9**(9): p. 4315-4328.
131. Schober, Y., S. Guenther, B. Spengler, and A. Rompp, *High-resolution matrix-assisted laser desorption/ionization imaging of tryptic peptides from tissue*. Rapid Commun Mass Spectrom, 2012. **26**(9): p. 1141-6.
  132. De Sio, G., A.J. Smith, M. Galli, M. Garancini, C. Chinello, F. Bono, F. Pagni, and F. Magni, *A MALDI-Mass Spectrometry Imaging method applicable to different formalin-fixed paraffin-embedded human tissues*. Mol Biosyst, 2015. **11**(6): p. 1507-14.
  133. Antoni, S., J. Ferlay, I. Soerjomataram, A. Znaor, A. Jemal, and F. Bray, *Bladder Cancer Incidence and Mortality: A Global Overview and Recent Trends*. Eur Urol, 2017. **71**(1): p. 96-108.
  134. Nepple, K.G. and M.A. O'Donnell, *The optimal management of T1 high-grade bladder cancer*. Can Urol Assoc J, 2009. **3**(6 Suppl 4): p. S188-92.
  135. Oezdemir, R.F., N.T. Gaisa, K. Lindemann-Docter, S. Gostek, R. Weiskirchen, M. Ahrens, K. Schwamborn, C. Stephan, D. Pfister, A. Heidenreich, R. Knuechel, and C. Henkel, *Proteomic tissue profiling for the improvement of grading of noninvasive papillary urothelial neoplasia*. Clin Biochem, 2012. **45**(1-2): p. 7-11.
  136. Thomsen, M.B.H., I. Nordentoft, P. Lamy, S. Vang, L. Reinert, C.K. Mapendano, S. Hoyer, T.F. Orntoft, J.B. Jensen, and L. Dyrskjot, *Comprehensive multiregional analysis of molecular heterogeneity in bladder cancer*. Sci Rep, 2017. **7**(1): p. 11702.
  137. Casadonte, R. and R.M. Caprioli, *Proteomic analysis of formalin-fixed paraffin-embedded tissue by MALDI imaging mass spectrometry*. Nat Protoc, 2011. **6**(11): p. 1695-709.
  138. Mascini, N.E., J. Teunissen, R. Noorlag, S.M. Willems, and R.M.A. Heeren, *Tumor classification with MALDI-MSI data of tissue microarrays: A case study*. Methods, 2018.
  139. Ruopp, M.D., N.J. Perkins, B.W. Whitcomb, and E.F. Schisterman, *Youden Index and optimal cut-point estimated from observations affected by a lower limit of detection*. Biom J, 2008. **50**(3): p. 419-30.
  140. Heinmoller, E., B. Renke, K. Beyser, W. Dietmaier, C. Langner, and J. Ruschoff, *Pitfalls in diagnostic molecular pathology--significance of sampling error*. Virchows Arch, 2001. **439**(4): p. 504-11.
  141. Steurer, S., J.M. Singer, M. Rink, F. Chun, R. Dahlem, R. Simon, E. Burandt, P. Stahl, L. Terracciano, T. Schlomm, W. Wagner, W. Hoppner, M. Omid, O. Kraus, M. Kwiatkowski, O. Doh, M. Fisch, A. Soave, G. Sauter, M. Wurlitzer, H. Schluter, and S. Minner, *MALDI imaging-based identification of prognostically relevant signals in bladder*

- cancer using large-scale tissue microarrays*. *Urol Oncol*, 2014. **32**(8): p. 1225-33.
142. Steurer, S., C. Borkowski, S. Odinga, M. Buchholz, C. Koop, H. Huland, M. Becker, M. Witt, D. Trede, M. Omidi, O. Kraus, A.S. Bahar, A.S. Seddiqi, J.M. Singer, M. Kwiatkowski, M. Trusch, R. Simon, M. Wurlitzer, S. Minner, T. Schlomm, G. Sauter, and H. Schluter, *MALDI mass spectrometric imaging based identification of clinically relevant signals in prostate cancer using large-scale tissue microarrays*. *Int J Cancer*, 2013. **133**(4): p. 920-8.
  143. Lopez, J.I. and J.M. Cortes, *Multisite tumor sampling: a new tumor selection method to enhance intratumor heterogeneity detection*. *Hum Pathol*, 2017. **64**: p. 1-6.
  144. Ehling, J., B. Theek, F. Gremse, S. Baetke, D. Mockel, J. Maynard, S.A. Ricketts, H. Grull, M. Neeman, R. Knuechel, W. Lederle, F. Kiessling, and T. Lammers, *Micro-CT imaging of tumor angiogenesis: quantitative measures describing micromorphology and vascularization*. *Am J Pathol*, 2014. **184**(2): p. 431-41.
  145. Enzinger, P.C. and R.J. Mayer, *Esophageal Cancer*. *New England Journal of Medicine*, 2003. **349**(23): p. 2241-2252.
  146. Domper Arnal, M.J., Á. Ferrández Arenas, and Á. Lanás Arbeloa, *Esophageal cancer: Risk factors, screening and endoscopic treatment in Western and Eastern countries*. *World journal of gastroenterology*, 2015. **21**(26): p. 7933-7943.
  147. Lin, E.W., T.A. Karakasheva, P.D. Hicks, A.J. Bass, and A.K. Rustgi, *The tumor microenvironment in esophageal cancer*. *Oncogene*, 2016. **35**: p. 5337.
  148. Runge, T.M., J.A. Abrams, and N.J. Shaheen, *Epidemiology of Barrett's Esophagus and Esophageal Adenocarcinoma*. *Gastroenterol Clin North Am*, 2015. **44**(2): p. 203-31.
  149. Smyth, E.C., J. Lagergren, R.C. Fitzgerald, F. Lordick, M.A. Shah, P. Lagergren, and D. Cunningham, *Oesophageal cancer*. *Nat Rev Dis Primers*, 2017. **3**: p. 17048.
  150. Testa, U., G. Castelli, and E. Pelosi, *Esophageal Cancer: Genomic and Molecular Characterization, Stem Cell Compartment and Clonal Evolution*. *Medicines (Basel)*, 2017. **4**(3).
  151. Iyer, P.G. and V. Kaul, *Barrett Esophagus*. *Mayo Clin Proc*, 2019. **94**(9): p. 1888-1901.
  152. Caspa Gokulan, R., M.T. Garcia-Buitrago, and A.I. Zaika, *From genetics to signaling pathways: molecular pathogenesis of esophageal*

- adenocarcinoma*. *Biochim Biophys Acta Rev Cancer*, 2019. **1872**(1): p. 37-48.
153. Talukdar, F.R., M. di Pietro, M. Secrier, M. Moehler, K. Goepfert, S.S.C. Lima, L.F.R. Pinto, D. Hendricks, M.I. Parker, and Z. Herceg, *Molecular landscape of esophageal cancer: implications for early detection and personalized therapy*. *Ann N Y Acad Sci*, 2018. **1434**(1): p. 342-359.
  154. Cressman, E.N.K. and J.M. Spraggins, *Viewing the Future of IR through Molecular Histology: An Overview of Imaging Mass Spectrometry*. *J Vasc Interv Radiol*, 2018. **29**(11): p. 1543-1546.e1.
  155. Barre, F.P.Y., B.S.R. Claes, F. Dewez, C. Peutz-Kootstra, H.F. Munch-Petersen, K. Gronbaek, A.H. Lund, R.M.A. Heeren, C. Come, and B. Cillero-Pastor, *Specific Lipid and Metabolic Profiles of R-CHOP-Resistant Diffuse Large B-Cell Lymphoma Elucidated by Matrix-Assisted Laser Desorption Ionization Mass Spectrometry Imaging and in Vivo Imaging*. *Anal Chem*, 2018. **90**(24): p. 14198-14206.
  156. Schone, C., H. Hofler, and A. Walch, *MALDI imaging mass spectrometry in cancer research: combining proteomic profiling and histological evaluation*. *Clin Biochem*, 2013. **46**(6): p. 539-45.
  157. Mascini, N.E., J. Teunissen, R. Noorlag, S.M. Willems, and R.M.A. Heeren, *Tumor classification with MALDI-MSI data of tissue microarrays: A case study*. *Methods*, 2018. **151**: p. 21-27.
  158. Quaas, A., A.S. Bahar, K. von Loga, A.S. Seddiqi, J.M. Singer, M. Omid, O. Kraus, M. Kwiatkowski, M. Trusch, S. Minner, E. Burandt, P. Stahl, W. Wilczak, M. Wurlitzer, R. Simon, G. Sauter, A. Marx, and H. Schluter, *MALDI imaging on large-scale tissue microarrays identifies molecular features associated with tumour phenotype in oesophageal cancer*. *Histopathology*, 2013. **63**(4): p. 455-62.
  159. Elsner, M., S. Rauser, S. Maier, C. Schone, B. Balluff, S. Meding, G. Jung, M. Nipp, H. Sarioglu, G. Maccarrone, M. Aichler, A. Feuchtinger, R. Langer, U. Jutting, M. Feith, B. Kuster, M. Ueffing, H. Zitzelsberger, H. Hofler, and A. Walch, *MALDI imaging mass spectrometry reveals COX7A2, TAGLN2 and S100-A10 as novel prognostic markers in Barrett's adenocarcinoma*. *J Proteomics*, 2012. **75**(15): p. 4693-704.
  160. Aichler, M., M. Elsner, N. Ludyga, A. Feuchtinger, V. Zangen, S.K. Maier, B. Balluff, C. Schone, L. Hierber, H. Braselmann, S. Meding, S. Rauser, H. Zischka, M. Aubele, M. Schmitt, M. Feith, S.M. Hauck, M. Ueffing, R. Langer, B. Kuster, H. Zitzelsberger, H. Hofler, and A.K. Walch, *Clinical response to chemotherapy in oesophageal adenocarcinoma*

- patients is linked to defects in mitochondria. *J Pathol*, 2013. **230**(4): p. 410-9.
161. Abbassi-Ghadi, N., O. Golf, S. Kumar, S. Antonowicz, J.S. McKenzie, J. Huang, N. Strittmatter, H. Kudo, E.A. Jones, K. Veselkov, R. Goldin, Z. Takats, and G.B. Hanna, *Imaging of Esophageal Lymph Node Metastases by Desorption Electrospray Ionization Mass Spectrometry*. *Cancer Res*, 2016. **76**(19): p. 5647-5656.
  162. Wang, M., J. Zhu, D.M. Lubman, and C. Gao, *Aberrant glycosylation and cancer biomarker discovery: a promising and thorny journey*. *Clin Chem Lab Med*, 2019. **57**(4): p. 407-416.
  163. Mechref, Y., A. Hussein, S. Bekesova, V. Pungpapong, M. Zhang, L.E. Dobrolecki, R.J. Hickey, Z.T. Hammoud, and M.V. Novotny, *Quantitative serum glycomics of esophageal adenocarcinoma and other esophageal disease onsets*. *J Proteome Res*, 2009. **8**(6): p. 2656-66.
  164. Everest-Dass, A.V., M.T. Briggs, G. Kaur, M.K. Oehler, P. Hoffmann, and N.H. Packer, *N-glycan MALDI Imaging Mass Spectrometry on Formalin-Fixed Paraffin-Embedded Tissue Enables the Delineation of Ovarian Cancer Tissues*. *Mol Cell Proteomics*, 2016. **15**(9): p. 3003-16.
  165. Briggs, M.T., M.R. Condina, Y.Y. Ho, A.V. Everest-Dass, P. Mittal, G. Kaur, M.K. Oehler, N.H. Packer, and P. Hoffmann, *MALDI Mass Spectrometry Imaging of Early- and Late-Stage Serous Ovarian Cancer Tissue Reveals Stage-Specific N-Glycans*. *Proteomics*, 2019: p. e1800482.
  166. Scott, D.A., R. Casadonte, B. Cardinali, L. Spruill, A.S. Mehta, F. Carli, N. Simone, M. Kriegsmann, L. Del Mastro, J. Kriegsmann, and R.R. Drake, *Increases in Tumor N-Glycan Polylactosamines Associated with Advanced HER2-Positive and Triple-Negative Breast Cancer Tissues*. *Proteomics Clin Appl*, 2019. **13**(1): p. e1800014.
  167. Herrera, H., T. Dilday, A. Uber, D. Scott, J.N. Zambrano, M. Wang, P.M. Angel, A.S. Mehta, R.R. Drake, E.G. Hill, and E.S. Yeh, *Core-Fucosylated Tetra-Antennary N-Glycan Containing A Single N-Acetylglucosamine Branch Is Associated with Poor Survival Outcome in Breast Cancer*. *Int J Mol Sci*, 2019. **20**(10).
  168. Powers, W.T., S. Holst, M. Wuhrer, S.A. Mehta, and R.R. Drake, *Two-Dimensional N-Glycan Distribution Mapping of Hepatocellular Carcinoma Tissues by MALDI-Imaging Mass Spectrometry*. *Biomolecules*, 2015. **5**(4).
  169. Heijs, B., S. Holst, I.H. Briaire-de Bruijn, G.W. van Pelt, A.H. de Ru, P.A. van Veelen, R.R. Drake, A.S. Mehta, W.E. Mesker, R.A. Tollenaar, J.V.M.G. Bovée, M. Wuhrer, and L.A. McDonnell, *Multimodal Mass Spectrometry Imaging of N-Glycans and Proteins from*

- the Same Tissue Section*. Analytical Chemistry, 2016. **88**(15): p. 7745-7753.
170. Scott, D.A., K. Norris-Caneda, L. Spruill, E. Bruner, Y. Kono, P.M. Angel, A.S. Mehta, and R.R. Drake, *Specific N-linked glycosylation patterns in areas of necrosis in tumor tissues*. International Journal of Mass Spectrometry, 2019. **437**: p. 69-76.
171. Kimchi, E.T., M.C. Posner, J.O. Park, T.E. Darga, M. Kocherginsky, T. Karrison, J. Hart, K.D. Smith, J.J. Mezhir, R.R. Weichselbaum, and N.N. Khodarev, *Progression of Barrett's metaplasia to adenocarcinoma is associated with the suppression of the transcriptional programs of epidermal differentiation*. Cancer Res, 2005. **65**(8): p. 3146-54.
172. Gomes, L.I., G.H. Esteves, A.F. Carvalho, E.B. Cristo, R. Hirata, Jr., W.K. Martins, S.M. Marques, L.P. Camargo, H. Brentani, A. Pelosof, C. Zitron, R.A. Sallum, A. Montagnini, F.A. Soares, E.J. Neves, and L.F. Reis, *Expression profile of malignant and nonmalignant lesions of esophagus and stomach: differential activity of functional modules related to inflammation and lipid metabolism*. Cancer Res, 2005. **65**(16): p. 7127-36.
173. Hemmoranta, H., T. Satomaa, M. Blomqvist, A. Heiskanen, O. Aitio, J. Saarinen, J. Natunen, J. Partanen, J. Laine, and T. Jaatinen, *N-glycan structures and associated gene expression reflect the characteristic N-glycosylation pattern of human hematopoietic stem and progenitor cells*. Experimental Hematology, 2007. **35**(8): p. 1279-1292.
174. Stanley, P., N. Taniguchi, and M. Aebi, *N-Glycans*, in *Essentials of Glycobiology*, rd, et al., Editors. 2015: Cold Spring Harbor (NY). p. 99-111.
175. Kaprio, T., T. Satomaa, A. Heiskanen, C.H. Hokke, A.M. Deelder, H. Mustonen, J. Hagström, O. Carpen, J. Saarinen, and C. Haglund, *N-glycomic Profiling as a Tool to Separate Rectal Adenomas from Carcinomas*. Molecular & Cellular Proteomics, 2015. **14**(2): p. 277.

# ***List of Publications***



## This thesis is based on the following publications

### **Experimental and data analysis considerations for three-dimensional mass spectrometry imaging in biomedical research**

**D. R. Naomi Vos**, Shane R. Ellis, Benjamin Balluff, Ron M. A. Heeren  
Molecular Imaging and Biology **2020**

### **Class-specific depletion of lipid ion signals in tissues upon formalin fixation**

**D. R. Naomi Vos**, Andrew P. Bowman, Ron M. A. Heeren, Benjamin Balluff, Shane R. Ellis  
International Journal of Mass Spectrometry **2019** 446, 116212

### **Strategies for managing multi-patient 3D mass spectrometry imaging data**

**D. R. Naomi Vos**, Ilaria Jansen, Marit Lucas, Martin R. L. Paine, Onne J. de Boer, Sybren L. Meijer, C Dilara Savci-Heijink, Henk A Marqeurung, D. Martijn de Bruin, Ron M. A. Heeren, Shane R. Ellis, Benjamin Balluff  
Journal of Proteomics **2019** 193, 184-191

### **2.5D N-glycan mass spectrometry imaging of esophageal adenocarcinogenesis**

**D. R. Naomi Vos**, Sybren L. Meijer, Ron M. A. Heeren, Shane R. Ellis, Benjamin Balluff  
Manuscript in preparation

## Other publications

### **Simultaneous lipidomic and transcriptomic profiling in mouse brain punches of acute epileptic seizure model compared to controls**

Raissa Lerner, Julia M. Post, Shane R. Ellis, **D. R. Naomi Vos**, Ron M. A. Heeren, Beat Lutz, Laura Bindila

The Journal of Lipid Research **2018** 59, 283-297

# ***Acknowledgments***



Congratulations, you made it to the acknowledgments part of this thesis. I never thought I would make it this far and while I'm writing this with our pup Boomer on my lap I still can't fathom it.

The road was long and dark at times with no light at the end of the tunnel in sight. But with willpower, faith, and support from the people around me, I made it through it all. Even if things didn't always go the way I wanted, like the ending of my PhD in corona-times, I am grateful to have made it here. To have an end-product reflective of my scientific achievements as well as my personality.

As said before, this wouldn't have been possible without the people supporting and guiding me along the way, both scientifically and personally. I would like to take the opportunity to thank them in the following pages.

**Ron**, thank you for believing in me and for giving me the opportunity to do my PhD in your lab. Your guidance not only made me a better scientist it also helped me grow as a person. I also want to convey my gratitude for supporting me and giving me the time and space I needed to work on my mental health.

**Shane**, thank you for supervising me these four years. As one of my two supervisors you have taught me a lot about (imaging) mass spectrometry in general and specifically related to lipids. These insights you have given me I will never forget. You have also helped me to become a better writer and more independent. When you told me that you would be moving back to Australia and not be around for the last part of my PhD, I wasn't sure how it would go and if I would manage. That I am here now, writing this part, is a testimony to how well you have guided me these years that allowed me to finish (relatively) smoothly while you are back in Australia.

**Benjamin**, as my other supervisor, you have also guided me well these years. When I started I did not know much about mass spectrometry imaging let alone the data analysis that comes with it. Over these four years, you have taught me a great deal about how to not only analyze MSI data but also judge the quality of the data. I want to thank you for all that you have taught me

and your guidance throughout these four years. I will never forget our “road trip” to Belgium for an ITEA meeting.

**Nina**, four years ago when I started my PhD I didn't know you yet. Since then a lot has happened to where I now get to call you one of my best friends and paranymp. Your support has seen me through some dark times and your “kick in the butt” got me to finally get help for my mental health. I will forever be grateful for this. The discussions we had about science have broadened my view on things beyond my own and helped me shape into the scientist I am today. I will never forget our dinners and drink dates which were filled with joy and where we talked about life and everything under the sun. Also the fun we had going out dancing, at birthday parties, and conferences were quite memorable.

**Maxime**, our PhD's started around the same time and as it developed so did our friendship. Over the years it blossomed into something special that people (Jian-Hua) even thought we were dating. I want to thank you for your unrelenting support, both as a friend and paranymp, and the joy and lightheartedness you have given me, especially in times I was struggling. We also had a lot of fun together, like OurCon in Saint-Malo, watching the world cup and other football games together, carnival and so much more.

**Britt**, what a lot of fun did we have together. From dancing and singing in the lab, thereby chasing Brynn away, to long coffee breaks, and going out for dinner and drinks. At times it seemed we were joined to the hip and people thought we were not working, but almost always we were actually talking science. We have had some very interesting and fruitful discussions and you were the best H&E buddy. I will never forget celebrating carnival together and all the fun we had going out dancing.

**Florian**, we have experienced a lot together, especially new things for me. Like the first time on a plane, first time swimming in the sea, and my first cocktail. And that all in the first two months of me starting my PhD. The rest of the years we have shared many more things and our going out dancing after the Christmas dinner is something I'll remember for years to come. Our meaningful discussions about science and life gave me new perspectives.

**Tiff**, you have been my conference (bed) buddy. It was great fun sharing these experiences with you. Our ASMS San Diego trip and stay in La Jolla was epic, even though getting there was a bit hectic. In the lab, we worked together on courses and workshops which was a real pleasure. As well as going for “one drink” with Britt.

**Fred**, I got to know you when you were still interning under Florian and we were in the 3<sup>rd</sup>-floor office. When you started your PhD our friendship really started with our trip to Lille as one of the highlights. Doing Omnia and core together with you and Jian-Hua was also something else. I have also never met someone that was so happy to know that McDonalds was delivering.

**Marty**, we both started to work around the same time on the ITEA project. Sharing those meetings with you helped make them more bearable and fun. I still remember our trip to Amsterdam together with Benjamin. You were also the first person to show and teach me how to do MSI. Thank you for all the wisdom and experience you shared with me. They helped me become a better scientist. We have also shared many memorable moments outside the lab that I will never forget like waking you up with Feliz Navidad in the car on our way to the Munster X-mass meeting.

**Jian-Hua**, when you started I remember that you had trouble with everyone's name except for mine (still not Chinese) much to my pleasure and the Frenchies' dismay. You quickly became our adopted Chinese and took us with you into your culture. Learning how to make dumplings from you was frightening and fun at the same time as well as going with you to get your ears pierced.

To our Spanish ladies, **Bea** and **Marta**, it was a delight knowing you and having you in the lab. Bea, I hope you're staying upright and not making any more starfishes on the floor. We are still having difficulties planning things for Nesquick. Marta, drinking sangria, and celebrating carnival was a great combination.

**Pi-max**, sharing an office with you was amazing. You're sometimes chaotic running in and out of the office was always amusing to see. The 3<sup>rd</sup>-floor office trash party you organized when we had to move was legendary.

## Acknowledgments

**Ana** and **Lucia**, you two are a crazy duo. Eating dinner outside when it's cold, really?? It was great fun going climbing together with you girls. The dancing is also something I will never forget.

To **Darya**, **Isabeau**, and **Fabian** it was a real pleasure sharing birthdays, dinners, and carnivals with you guys.

I would also like to thank my M4I colleagues that made working at M4I a great learning experience. Thank you for all your support and the fun we shared during picnics, dinners, and game nights. **Berta**, **Helen**, **Maarten**, **Joel**, singing Taylor Swift together in the lab was awesome, **Fanny**, for being my digestion guru when I started, **Anne**, **Klára**, **Lieke**, getting Shane drunk was so worth the middle finger, **Bryn**, **Andrew**, **Martha**, **Philippe**, **Eva**, **Tristan**, **Felix**, **Yuandi**, **Peiliang**, **Mirella**, **Lennart**, **Brenda**, **Anjusha**, **Pieter**, **Michele**, **Stephanie**, **Jo**, **Ian**, **Roel**, **Sylvia**, **Christel**, **Jonah**, **Arnoud**, **Greg**, **Keely**, **Ronny**, **Livia**, **Rob**, **Gert**, and **Frans**.

During my time at M4I, I also had the pleasure to meet guests from around the world. **Rachelle**, it was great having you in the lab and sharing dinners. I hope you're not ripping any more pants by dropping down too low. **Maria**, it was a pleasure getting to know you and sharing good times both in and outside the lab. **Allison**, **Piia**, **Jaanus**, **Chiara**, **Anton**, **Silvia**, and **Kamil** thanks for the fun times we shared and interesting discussions.

An mien oale leu, aan mijn lieve ouders, die mij door alles heen gesteund hebben, bedankt voor alles. Mijn broers, zus en uitgebreide familie, bedankt voor alle gezelligheid. Onze Tommie, bedankt voor alle liefde en rust die je mij hebt gegeven. En ten slotte onze pup Boomer, bedankt voor de welkome afleiding en stress verlichting tijdens het afronden van mijn PhD.

# ***Curriculum Vitae***



## Dina Rebecca Naomi Vos

I was born on June 20th, 1991 in Hengelo (O) the Netherlands. I've grown up in Delden where I completed my elementary school and obtained my secondary education diploma at the Grundel in Hengelo.



I continued studying at Wageningen University and Research where I followed the bachelor molecular life science. This enabled me to explore my interest in science, especially on a molecular level, while not having to choose a defined area. In addition, I gained a broad knowledge of different areas of research. During my bachelor, I discovered my love for biomedical research in particular after composing a minor in health biology. This led me to follow the master molecular life science with a specialization in biomedical research.

In the second year of my master, I did my 7-month thesis research at the Laboratory of BioNanoTechnology where I could work on a multidisciplinary research project. In the last 7 months of my master, I went to the Eindhoven University of Technology to the lab of dr. ir. Ilja Voets, which allowed me to learn small-angle x-ray scattering (SAXS).

After obtaining my master's degrees in 2015 I applied for a PhD position at Ron Heeren's lab in 2016. Under the supervision of Ron Heeren, Shane Ellis, and Benjamin Balluf I was given the opportunity to work on the advancement of 3D-MSI for tissue diagnosis within the ITEA 3D Pathology consortium. My research was focused on acquiring 3D MALDI-MSI data from human FFPE bladder cancer resection for the integration and visualization with other modalities by project partners. The obtained data was also used to develop a data analysis pipeline for 3D MALDI-MSI that I then employed in another 3D-MSI study to study esophageal adenocarcinogenesis using human FFPE material.

In parallel with my PhD I organized M4I-MSI courses that were given by PI's to increase knowledge among the PhD's at the M4I-IMS department and which were open for collaborating departments.