Universiteit
Leiden
The Netherlands

Lab-on-a-tissue : optimization of on-tissue chemistry for improved mass spectrometry imaging

Heijs, B.P.A.M.

Citation

Heijs, B. P. A. M. (2018, February 1). *Lab-on-a-tissue : optimization of on-tissue chemistry for improved mass spectrometry imaging*. Retrieved from https://hdl.handle.net/1887/60212

Version:	Not Applicable (or Unknown)
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/60212

Note: To cite this publication please use the final published version (if applicable).

Cover Page

Universiteit Leiden

The handle http://hdl.handle.net/1887/60212 holds various files of this Leiden University dissertation.

Author: Heijs, B.P.A.M.
Title: Lab-on-a-tissue : optimization of on-tissue chemistry for improved mass spectrometry imaging
Issue Date: 2018-02-01

Chapter 4

Histology-guided high resolution MALDI mass spectrometry imaging

Bram Heijs‡; Walid M. Abdelmoula‡; Sha Lou; Inge H. Briaire-de Bruijn; Jouke Dijkstra; Judith V.M.G. Bovée; Liam A. McDonnell, "Histology-guided high resolution MALDI mass spectrometry imaging," *Analytical Chemistry* 87(24), pp. 11978-11983, 2015.

Mass spectrometry imaging (MSI) is widely used for clinical research because when combined with histopathological analysis the molecular signatures of specific cells/regions can be extracted from the often-complex histologies of pathological tissues. The ability of MSI to stratify patients according to disease, prognosis and response is directly attributable to this cellular specificity. MSI developments are increasingly focused on further improving specificity, through higher spatial resolution to better localize the signals or higher mass resolution to better resolve molecular ions. Higher spatial/mass resolution leads to increased data size and longer data acquisition times. For clinical applications, which analyze large series of patient tissues, this poses a challenge to keep data load and, acquisition time manageable. Here we report a new tool to perform histology guided MSI; instead of analyzing large parts of each tissue section the histology from adjacent tissue sections is used to focus the analysis on the areas of interest, e.g. comparable cell types in different patient tissues, thereby minimizing data acquisition time and data load. The histology tissue section is annotated and then automatically registered to the MSI-prepared tissue section; the registration transformation is then applied to the annotations, enabling them to be used to define the MSI measurement regions. Using series of formalin-fixed, paraffin-embedded human myxoid liposarcoma tissues we demonstrate an 80% reduction of data load and acquisition time, thereby enabling high resolution (mass or spatial) to be more readily applied to clinical research. The software is freely available for download.

4.1 Introduction

Histological analysis is arguably the most common tool in diagnostic pathology and clinical research. The use of molecular assays further aids in understanding, determining, and differentiating diseases (Aichler and Walch, 2015; Altelaar et al., 2012; Bovée and Hogendoorn, 2010). A technique that recently emerged is matrix-assisted laser desorption/ionization (MALDI) mass spectrometry imaging (MSI), which is an analytical technique that combines conventional microscopy with spatially resolved and label-free detection of hundreds of biomolecules directly from a tissue section (Caprioli et al., 1997). MALDI-MSI has proven itself a valuable tool in molecular pathology as it is capable of revealing the molecular histology of a tissue section and finding distinct molecular profiles that can predict the nature of the underlying tissue (Willems et al., 2010b; Balluff et al., 2015; Casadonte et al., 2014; Gustafsson et al., 2011; Meding et al., 2012).

The impact of MSI in clinical research is directly attributable to the cellular specificity of the analysis; signatures of specific cells/regions can be extracted from the often-complex histologies of pathological tissues. MSI developments are increasingly focused on further improving specificity, through higher spatial resolution to better localize the signals or higher mass resolution to better resolve ions with identical nominal masses.

MSI analysis of proteins via on-tissue enzymatic digestion, metabolites, lipids and recently even intact proteins, all benefit from higher mass resolution by better resolving the many molecular species with the same nominal mass, so-called isobaric ions (Buck et al., 2015; Heijs et al., 2015; Muller et al., 2015; Spraggins et al., 2015). When analyzing large tissue series, a compromise must be sought between resolution and throughput, to keep the data load, acquisition time and computational challenges manageable. This is primarily because the established approach is to first record MSI data, and then use histopathological analysis of the same tissue section to define via a virtual microdissection which MSI data will be used in the statistical analysis. However only a fraction of the MSI data is typically used in the statistical analysis, as the cellular content of the compared regions must be well matched within the same tissue as well as within tissues from different patients. Accordingly, the data acquisition throughput and data load could be better optimized by adapting the histology-defined protein-profiling approach (Cornett et al., 2006) to high resolution MSI, namely limiting MSI data acquisition to those regions whose data will be used in the statistical analysis.

In the current work we have developed an automated image registration pipeline to register an annotated, histological image of an adjacent tissue section to a lower resolution image of the MSI-prepared tissue section. Subsequently, the annotation borders are propagated to the low resolution image, enabling the exclusive analysis of the annotated regions by MSI. This histology-guided MSI approach ensures that the MSI data is acquired solely from tissue regions with similar morphological makeup.
Using series of formalin-fixed, paraffin-embedded human myxoid liposarcoma tissue we demonstrate an 80% reduction of data load and acquisition time, thereby enabling high resolution (mass or spatial) to be more readily applied in clinical research.

4.2 Methodology

4.2.1 Chemicals and reagents

All chemicals were purchased from Sigma-Aldrich (Steinheim, Germany) except ethanol and xylene (Merck, Darmstadt, Germany). Trypsin was purchased from Promega (Madison, USA).

4.2.2 Tissues & sample collection

Myxoid liposarcoma (MLS) tissues were fixed in formalin and embedded in paraffin following standard histopathological procedures. All tissue samples were handled in a coded fashion according to Dutch national ethical guidelines (Code for proper secondary use of human tissue, Dutch Federation of Medical Scientific Societies). Adjacent tissue sections, 6 µm, were cut from the formalin-fixed, paraffin-embedded (FFPE) tissue blocks using a microtome. Tissue sections designated for MALDI-MSI analyses were mounted on poly-L-lysine coated indium-tin-oxide (ITO) glass slides (Bruker Daltonics, Bremen, Germany), dried overnight at 37°C and stored at 4°C until the analysis. Tissues designated for histological analysis were mounted on Starfrost

adhesive microscope slides (Light Labs, Dallas (TX), USA), stained with haematoxylin & eosin (H&E) and scanned using a digital slide scanner (IntelliSite Pathology Ultra-Fast Scanner, Philips, Eindhoven, The Netherlands). This tissue section is referred to as the histological-tissue-section in the subsequent text.

4.2.3 On-tissue digestion MALDI-MSI

FFPE tissue sections designated for MSI analysis were first put on a 60°C heating block to improve adherence of the tissue section to the slide. Deparaffination was performed by washing the slides twice in 100% xylene (I - 5 min, II - 10 min). Prior to heat-induced antigen retrieval (HIAR) in 10 mM citric acid pH6, the slides were washed twice in 100% ethanol (2 min each) and twice in milliQ water (5 min each). After HIAR two more washes in milliQ water (1 min each) were performed. The tissues were dried in a vacuum desiccator for 15 minutes. Fiducial markers were placed using water-based Tipp-Ex and a lower resolution (2400 dpi) optical image recorded using a flatbed scanner. This image is referred to as the *pre-MSI-optical-image*. Immediately thereafter trypsin (20 ng/µL in deionized water) was sprayed on the tissue using the SunCollect automatic sprayer (SunChrom, Friedrichsdorf, Germany) (5 layers at 5 µL/min). Following an 18 h incubation in a saturated (50% methanol) environment at 37°C, three layers (at 10, 35, and 35 µL/min) of MALDI matrix (20 mg/mL 2,5-dihydroxybenzoic acid in 50% acetonitrile and 0.1% trifluoroacetic acid) were sprayed on the tissue using the SunCollect automatic sprayer (Heijs et al., 2015).

MALDI-FTICR-MSI was performed on a 9.4 T SolariX XR mass spectrometer (Bruker Daltonics) in positive-ion mode, using 150 laser shots per spot and 100 x 100 µm pixel size, calibrated using Peptide Calibration Standard (Bruker Daltonics). Spectra were recorded in the *m/z* range 600–3,500 Da with a 512k data point transient (1.1 s duration), corresponding to an estimated resolution of 200,000 at *m/z* 400 Da. The average detected mass resolution over the analyzed mass range was 54,000, and the mass resolution in the center of the mass range, at *m/z* 1,500, was 65,000. Data acquisition was performed using ftmsControl (Bruker Daltonics), and visualizations were obtained from flexImaging 4.0 (Bruker Daltonics). Following the MALDI-MSI data acquisition, excess matrix was removed by washing in 70% ethanol (2x 2 min) and this MSI-analyzed-tissue-section also stained with H&E. The stained tissues were scanned using a digital slide scanner and used to assess the accuracy of the registration between this MSI-analyzed-tissue section and the histological tissue section. This image is referred to as the *post-MSI-histological-image*.

4.2.4 Image registration software tool

4.2.4.1 Pre-processing

Background noise was first removed from both the *histological-tissue-section* image and *pre-MSI-optical-image*. The pre-processing retains the tissue foreground information while setting the background pixels to zero. The intensity distribution of the

optical image follows a bimodal histogram that represent background and foreground pixel distribution, hence their separation is readily achieved by setting a global threshold using Otsu's method16. Morphological operations then followed to fill-in any gaps introduced by setting the image intensity threshold. The region with the maximum area of connected components was selected to form a mask that separates the tissue from the background. The high-resolution histology images were pre-processed using the pre-processing-pipeline presented by Abdelmoula *et al.* (Abdelmoula et al., 2014).

4.2.4.2 Image Registration

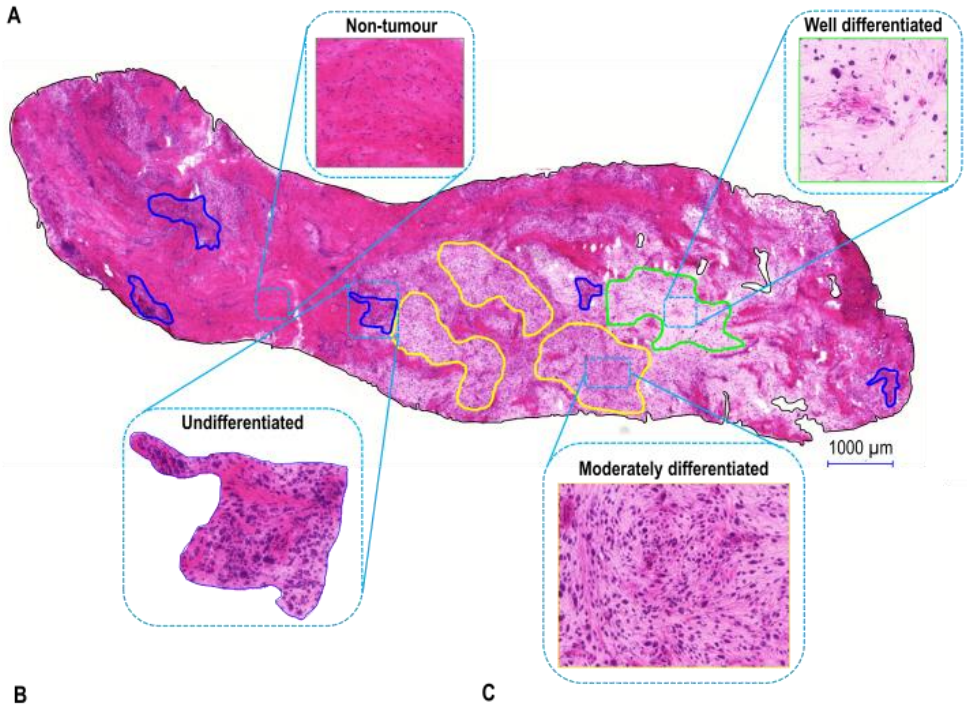
The registration process is performed between a fixed image (I_f), in this case the *pre-MSI-optical-image*, and a moving image (I_m), the *histological-tissue-section* image. The moving image is warped to be spatially aligned with the fixed image according to the standard registration optimization problem given in equation (1).

$$\hat{\mu} = \arg\min_{\mu} C[I_f, T(I_m)] \tag{1}$$

The optimized transformation parameters are captured in the vector μ in such a way the cost function C is minimized with respect to the transformation model T using the adaptive stochastic gradient descent optimizer (Klein et al., 2011). First, I_m was linearly transformed, using affine transformation, to capture the global deformation parameters (i.e. translation, rotation, scaling, and shearing), followed by a non-linear transformation model of B-Spline to correct for local deformations. The registration was performed in a multi-resolution scheme (4 levels) using a Gaussian pyramid model and the mutual information was used as a cost function to assess the registration quality. The multi-resolution registration has proven faster and more robust than single-scale techniques with a convergence that is less likely to be trapped in a local optimum (Unser et al., 1993). The mutual information has been introduced for multi-modality image registration, as it measures the statistical dependence between two images with the same geometrical information but different intensity distributions and maximizing their statistical measure implies they are geometrically aligned (Maes et al., 1997; Pluim et al., 2003). The Elastix software package was used to implement the reported registration algorithms (Klein et al., 2010).

4.2.4.3 Evaluation

The accuracy of the image registration was assessed by comparing expert pathological annotations of the *post-MSI-histological-image* (JVMGB) with those provided by registering the annotations of the *histological-tissue-section* to the *pre-MSI-histological-image*. The accuracy was quantified by using the Dice similarity coefficient (DSC), which is a measure for the similarity between two samples (Dice, 1945).



B

Region	Surface [µm²]	Surface [%]
— Full tissue section	44,379,119	100
— Well differentiated	1,574,331	3.6
— Moderately differentiated	4,117,824	9.3
— Undifferentiated	1,853,890	4.2

C

Tumour type	Annotated [%]	σ [%]
Osteosarcoma (n = 17)	24	4
Myxofibrosarcoma (n = 12)	16	2
Leiomyosarcoma (n = 13)	8	3
Undifferentiated Pleomorphic Sarcoma (n = 13)	38	5
Total tissue series (n = 55)	22	11

Figure 4-1: (A) Scanned image of an H&E stained myxofibrosarcoma tissue with annotated regions of clinical interest. Pixels present in the ROIs were selected for data analysis. Various tumour grades were selected based on their cellular differentiation, green, yellow and blue annotations represent well differentiated, moderately differentiated and undifferentiated tissue regions respectively. **(B)** The (relative) surface of the differently diagnosed tissue regions compared to the total surface of the full tissue section. **(C)** The average percentage of pixels used for data analysis per tumour subtype for a total of 55 soft tissue sarcoma samples.

4.3 Results and discussion

An established approach in clinical MALDI-MSI is to analyze large areas of a series of tissue sections at a predefined spatial resolution. Subsequently, excess matrix is removed and a histological staining of the analyzed section allows to perform virtual microdissection, and to export the pixels from regions with similar cellular makeup for further data analysis (Figure 4-1A).

A previous analysis to study intact proteins from a patient series of high grade sarcomas (n = 55) revealed that, after virtual microdissection, an average of just 22% (±11%) of the MALDI-MSI data was used in the subsequent data analysis (Figure 4-1B-C). This dataset was acquired at a 100 µm spatial resolution on a relatively low

mass resolution MALDI-TOF/TOF platform (UltrafleXtreme, Bruker Daltonics) but already resulted in a total dataset size of 232 GB for the 55 analyzed tissue sections. This post-MSI virtual microdissection approach is sub optimal for high resolution studies, because of the greatly increased data load and data acquisition times associated with higher mass/spatial resolution.

We recognized that we could increase the measurement throughput and, decrease data load if MSI data acquisition was limited to histologically comparable regions, in an MSI analogue of the histology defined protein profiling approach (Cornett et al., 2006). Therefore, we have developed histology-guided MSI (HG-MSI) software that makes use of automated image registration to limit data acquisition to specific histologically pre-defined regions (Figure 4-2).

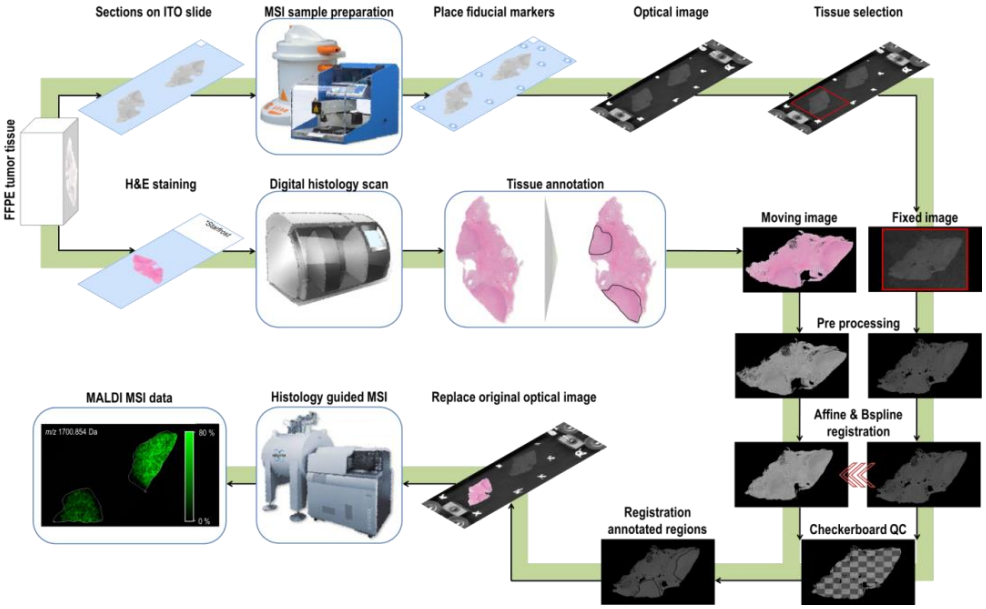


Figure 4-2: A schematic overview of the most important steps in the HG-MSI workflow. The workflow enables the selection of MALDI-MSI measurement regions to be guided by the tissue's histology.

The *Histology_Guided_MS* software performs an automatic registration of images of adjacent tissue sections: (i) the moving image (I_m), which is represented by an annotated high resolution image of an H&E stained tissue section, here referred to as the *histological-tissue section* (Figure 4-3A), and (ii) the fixed image (I_f), a lower resolution optical image required for the MALDI-MSI analysis, here referred to as the *pre-MSI-optical-image* (Figure 4-3B-top). In this manner the histological annotations are mapped onto the *pre-MSI-optical-image* and can be used to define the measurement regions (Figure 4-3B-bottom).

The adjacent tissue sections were manually mounted onto different glass slides and scanned at different resolutions; and so the registration encompasses a linear, affine

registration to correct for global deformations (e.g. translation, scaling, rotation, and shearing) as well as a non-linear, B-spline transform to compensate for local deformations (e.g. differences in sizes of regions, stretches introduced during mounting) (Abdelmoula et al., 2014).

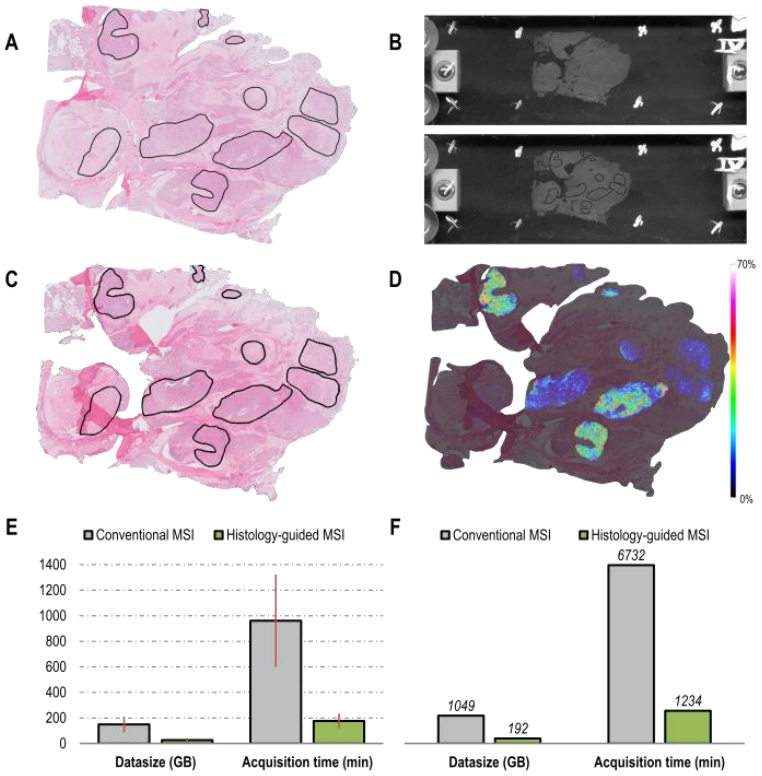
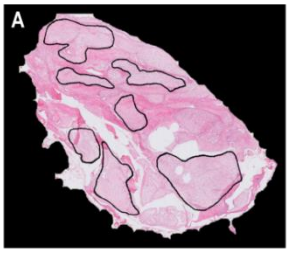
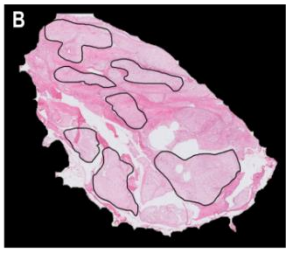


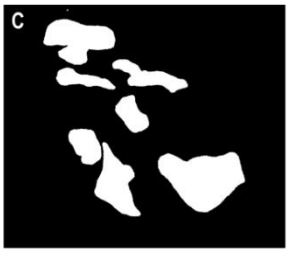
Figure 4-3: (A) *Histological-tissue-section* image of a myxoid liposarcoma tissue, including annotations marking the clinically relevant regions. **(B-top)** *Pre-MSI-optical-image* of the tissue prior to the MALDI-MSI analysis. The section was adjacent to the section shown in (A). **(B-bottom)** The output image of the Histology_guided_MS registration algorithm. The image shows the aligned, annotated regions from (A) propagated to the *pre-MSI-optical-image* (B-top). **(C)** *Post-MSI-optical-image*; the H&E stained tissue after MALDI-MSI data acquisition (B), including the aligned annotated regions from (A). **(D)** Example MSI image of *m/z* 944,51 Da corresponding to Histone H2A peptide AGLQFPVGR represented by the Jet intensity scale. **(E-F)** Comparison between conventional MSI and histology-guided MSI. Shown are the average (E) and total (F) data size and acquisition times for the MLS tissue series (n = 7).

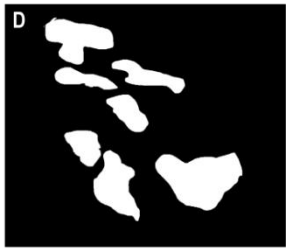
As the MSI measurement regions are determined by the histological annotations from an adjacent section, an investigation of the image registration accuracy is necessary. After MSI data acquisition excess MALDI matrix was removed and the tissue sections were H&E stained and, digitally scanned, this is the *post-MSI-optical-image*. The automatically registered annotations from the *histological-tissue-section* were mapped on to this image (Figure 4-3C & 4-4A), and then compared with the results of an expert manual histopathological annotation of the *post-MSI-optical-image* (Figure 4-4B). Visually it can be seen that there is a good agreement between the annotations.

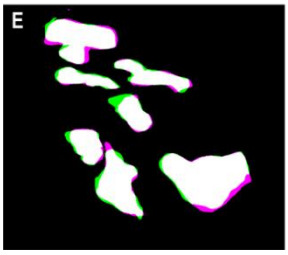
The HG-MSI software was tested by analyzing the proteolytic peptides from an on-tissue digestion investigation of a patient series of FFPE MLS tissues (n = 7) using a 9.4 T MALDI-FTICR instrument (Figure S4-1, Supplementary information). MLS was chosen as these tumors have a very heterogeneous histology (Willems et al., 2010b) that allowed the selection of ROIs with different tumour grades and/or additional histological features (e.g. fat differentiation, and mucine pools).











F

Tissue	Dice coefficient
L1357	0.8525
L2039	0.8440
L2235	0.9175
L2398	0.8375
L3242	0.8355
L3403	0.8268
L4175	0.8071
Mean	*0.8458 (± 0.0012)*

Figure 4-4: An evaluation of the image registration through the measure of annotation overlap using the Dice coefficient. **(A)** Automatically aligned annotations propagated to the post-MSI-optical-image. **(B)** Manually annotated post-MSI-optical-image. **(C)** + **(D)** Binary masks of the annotated regions in (A) and (B). **(E)** Fusion of the two binary mask showing the degree of overlap. In white, truly overlapping regions. In green and pink the non-overlapping regions. **(F)** An overview of the Dice coefficients for the analyzed tissues.

Analyzing the tissue series, utilizing the HG-MSI software, resulted in a total data load of 192 GB, which was just 18% of the 1,049 GB that would be generated by analyzing each tissue section of the series completely (Figure S4-2, Supplementary information). Similarly, the data acquisition time was also reduced by 82% from 112 hours to 20 hours (Figure 4-3E-F). Ultimately, throughput and data load are determined by the total area of the selected ROIs. For these tissues only regions with clear histologic grades (e.g. low, moderate and high grade) based on distinct histological features were selected, resulting in annotated areas varying between 12 - 29% of the total tissue surface.

The effectiveness of the HG-MSI software to focus on specific histological regions was then quantified using the Dice similarity coefficient (DSC) (Figure 4-4A-E), to provide a metric for the spatial overlap of the annotated regions in the automatically and manually annotated *post-MSI-optical image*. No overlap would result in a DSC of 0, perfect overlap in DSC of 1. For this tissue series the average DSC was 0.8458 (± 0.0012), confirming that the HG-MSI approach does indeed allow the analyst to target the specific regions (Figure 4-4F).

There are several causes that may contribute to the 15% reduction of the DSC. The first is that adjacent sections are not identical and so small differences are expected. Second, it was found that some of the histological regions were harder to define in the *post-MSI-optical-image*, and which may have led to slight differences in the annotation areas. Third, artifacts from the tissue sectioning and mounting such as stretches, tears and folds can cause large differences even between adjacent sections. For example, tissue L2235 and L4175 (Figure S4-1, Supplementary information) represent the experiments with the highest and lowest DSCs respectively (Figure 4-4F). Whereas the images of L2235 exhibit no clear sectioning/mounting artifacts and very distinct histological features, the *histological-tissue-section* of L4175 was folded and the histological features less distinct. To limit the impact of sectioning/mounting artifacts on the registration a manual, free-hand cropping function was added to the registration tool to exclude such regions from the registration process.

We have demonstrated this HG-MSI software for the analysis of proteolytic peptides from specific histopathological regions in tumor tissues using a Bruker Daltonics MALDI-FTICR-MSI. However, the tool is vendor neutral and is applicable to all MSI methodologies (e.g. DESI, SIMS, etc.), the only requirement is that an optical image of the tissue section is used to define the measurement regions. Similarly, the HG-MSI software may be used to better focus high spatial resolution as well as high mass resolution analysis, for all analyte classes and, all tissue types. The software is freely available for download (http://www.maldi-msi.org/download/SW_HG_MSI.rar). A detailed overview of the histology-guided MSI workflow and software manual is available as online Supplementary information. The only prerequisite to run the *Histology_guided_MS* executable is the installation of the Matlab Compiler Runtime (R2012a (7.17)) which is free to download from the Mathworks website (www.mathworks.com/product/compiler/mcr).

The compromise associated with specifying data acquisition from distinct histological regions is that retrospective data analysis possibilities are more limited than if each tissue section was analyzed in its entirety. For example, for the myxofibrosarcoma sample shown in Figure 4-1 if only the well-differentiated tumor regions (3.6% of the example tissue section) were selected for MSI data acquisition then it would no longer be possible to compare its mass spectral signature with those from the moderately differentiated tumor (9.3%) and undifferentiated tumor regions

(4.2%), or the tumor stroma and healthy tissue. It is thus crucial to specify, at the outset, which histological comparisons will/may be made. Accordingly, the tool is better suited to those analyses in which the slow scan speed (high mass resolution) or large number of pixels (high spatial resolution) makes it impractical to apply to a large patient series of complete tissue sections.

4.4 Conclusions

In this work we report a new freely available tool for histology-guided MALDI-MSI and provide the software required to perform automated image registration so that the annotations from an adjacent histological image can be transferred to the tissue section that will be analyzed by MSI. The results demonstrate the strength of the approach, particularly for high resolution MSI, in which the high resolution leads to longer acquisition times. The ease of use and general applicability of the tool in any MSI workflow will ultimately result in a higher throughput, a smaller data load, and more comparable data.

4.5 Acknowledgements

This work was supported by the ZonMW Zenith project "Imaging Mass Spectrometry-Based Molecular Histology: Differentiation and Characterization of Clinically Challenging Soft Tissue Sarcomas" (No. 93512002; B.H.), the ICT consortium COMMIT project "e-biobanking with Imaging" and the Cyttron II project "Imaging Mass Spectrometry".

4.6 Supplementary information

The full supplementary information, files and tables are available on the internet via http://pubs.acs.org/doi/abs/10.1021/acs.analchem.5b03610.

4.6.1 Supplementary figures

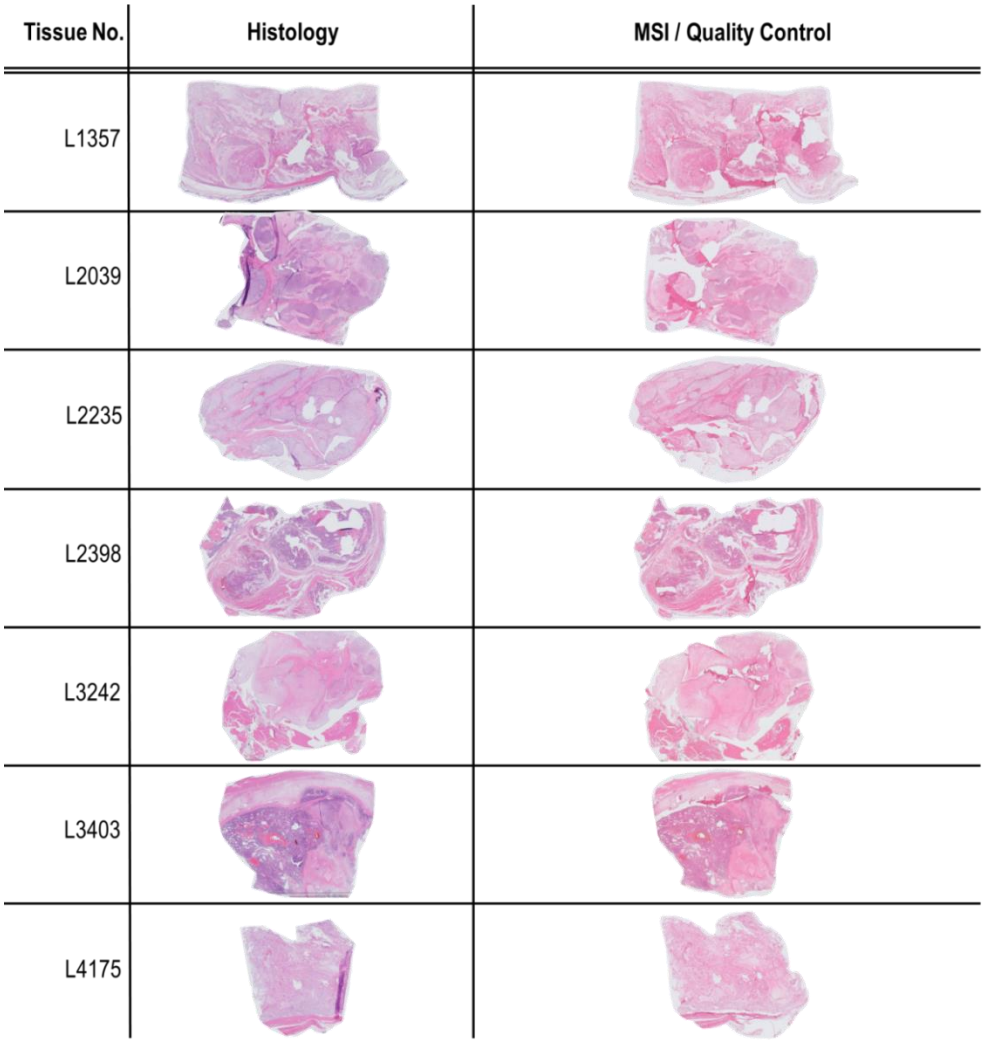


Figure S4-1: An overview of the H&E stained MLS tissues. **(left)** Tissue numbers; **(middle)** histology image prior to annotation; **(right)** MSI tissue after quality control staining.

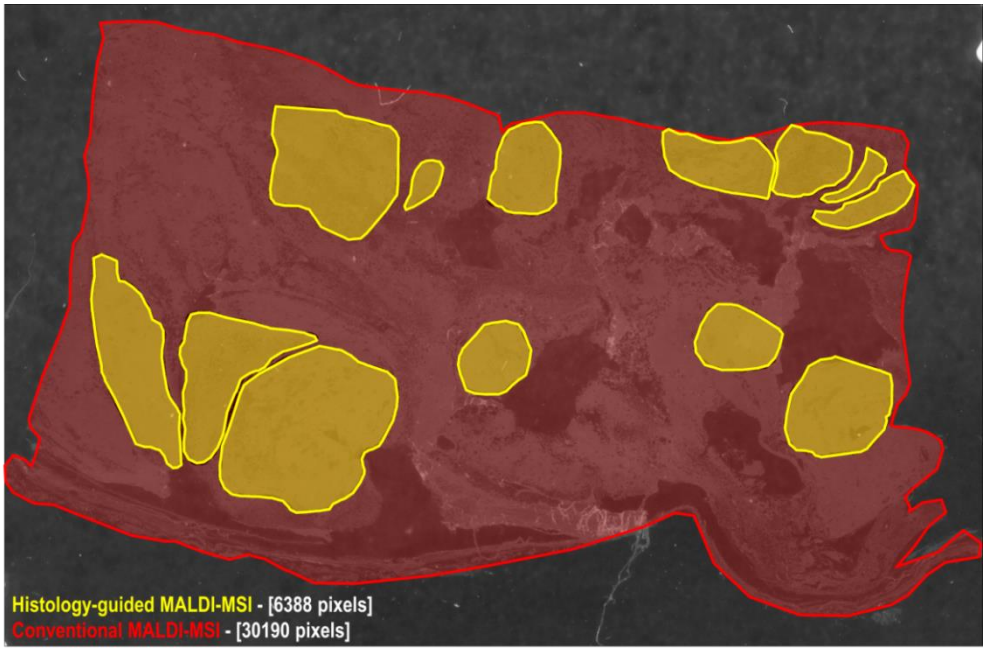


Figure S4-2: An example showing the differences in analysis surface between the HG-MSI platform and conventional MSI analysis. In yellow, the histology-guided analysis regions, and in red, the full surface area of the tissue analyzed in a conventional MSI experiment. The number of pixels is based on a spatial resolution of 100 µm. The datasize of the conventional MALDI-MSI dataset was calculated using (1).

$$Size_{Normal} = Pixels_{Normal}\left(\frac{Size_{HG}}{Pixels_{HG}}\right) \quad (1)$$