High-resolution MALDI-FT-ICR MS Imaging for the analysis of metabolites from formalin-fixed paraffin-embedded clinical tissue samples

Achim Buck[1‡], Alice Ly[1‡], Benjamin Balluff[2], Na Sun[1], Karin Gorzolka[1], Annette Feuchtinger[1], Klaus-Peter Janssen[3], Peter J. K. Kuppen[4], Cornelis J. H. van de Velde[4], Gregor Weirich[5], Franziska Erlmeier[5], Rupert Langer[5], Michaela Aubele[6], Horst Zitzelsberger[7], Michaela Aichler[1], and Axel Walch[1]*

Authors` affiliations:

[1]Research Unit Analytical Pathology, Helmholtz Zentrum München, Neuherberg, Germany

[2]Centre for Proteomics and Metabolomics, Leiden University Medical Center, Leiden, The Netherlands

[3]Department of Surgery, Klinikum Rechts der Isar, Technische Universität München, Munich, Germany

[4]Department of Surgery, Leiden University Medical Center, Leiden, The Netherlands

[5]Institute of Pathology, Technische Universität München, Munich, Germany

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/path.04560

[6]Institute of Pathology, Helmholtz Zentrum München, Neuherberg, Germany

[7]Research Unit Radiation Cytogenics, Helmholtz Zentrum München, Neuherberg, Germany

[‡]These authors contributed equally to this work.

*Correspondence to: A Walch, Research Unit Analytical Pathology, Helmholtz Zentrum München, Ingolstädter Landstr. 1, 85764 Neuherberg, Germany; E-mail: axel.walch@helmholtz-muenchen.de; Phone: +49 89 3187 2739; Fax: +49 89 3187 3349

Conflict of interest:

The authors disclose no potential conflicts of interest.

Keywords: formalin-fixed paraffin-embedded tissue; MALDI; Imaging Mass Spectrometry; metabolite; cancer

ABSTRACT

We present the first analytical approach to demonstrate the *in situ* imaging of metabolites from formalin-fixed paraffin-embedded (FFPE) human tissue samples. Using high-resolution Matrix-Assisted Laser Desorption/Ionization Fourier-Transform Ion Cyclotron Resonance Mass Spectrometry Imaging (MALDI-FT-ICR MSI), we conducted a proof of principle experiment comparing metabolite measurements from FFPE and fresh frozen tissue sections, and found an overlap of 72% amongst 1700 m/z species. In particular, we observed conservation of biomedically relevant information at the metabolite level in FFPE tissues. In biomedical applications, we analysed tissues from 350 different cancer patients, and were able to discriminate between normal and tumour tissues, and different tumours from the same organ, and found an independent prognostic factor for patient survival. This study demonstrates the ability to measure metabolites in FFPE tissues using MALDI-FT-ICR MSI, which can then be assigned to histology and clinical parameters. Our approach is a major technical, histochemical and clinicopathological advance that highlights the potential for investigating diseases in archived FFPE tissues.

INTRODUCTION

A broad spectrum of diseases are diagnosed based on histology - the microscopic anatomy of samples resected from patients. The gold standards for histopathological analysis, valid diagnosis and patient treatment are based on formalin-fixed paraffin-embedded (FFPE) patient samples. Formalin fixation was first reported in 1893 and has been shown to maintain tissue integrity for many years [1]. Formalin fixation preserves tissues by forming methylene bridge cross-linkages between proteins, and the sample is then dehydrated with graded ethanol, immersed in an organic solvent (e.g. xylene) to clear the alcohol, and then infiltrated

with molten paraffin. Much effort has been spent in the last decades on standardizing FFPE protocols. Compared to fresh frozen (FF) samples, FFPE tissues can be easily archived for many years with corresponding patient data for retrospective reassessment and clinicopathological studies. While histology is a reliable tool, advanced techniques have expanded knowledge of the molecular and genetic content of tissues, which is not readily apparent when assessing patient samples based on histology [2,3].

Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry Imaging (MALDI-MSI) is a powerful technique combining mass spectrometry with histology, allowing for the spatially-resolved and label-free detection of hundreds to thousands of compounds within a single tissue section. Depending on the sample preparation and instrumentation, MALDI-MSI can detect a wide variety of molecular classes, e.g. proteins, peptides, lipids, drugs and others [4,5]. Recently, the development of devices that have a high resolving power for small molecular weight compounds, e.g. AP-SMALDI-Orbitrap or MALDI-FT-ICR MS, have allowed for the *in situ* detection and visualization of metabolites and different metabolic states within tissues [6,7]. FF samples are currently preferred over FFPE in tissue-based MALDI-MSI studies. This is primarily due to concerns regarding the effect of chemical processing on molecular content, molecular delocalization from washing steps and the belief that paraffin would cause ion suppression [4,8]. At present, MALDI-MSI of FFPE tissue has primarily concentrated on protein and peptide content, the protocol for which utilizes deparaffination, antigen-retrieval and enzyme digestion adapted from immunohistochemical procedures [4].

Measurements of small molecule metabolites (<1500 Da) are emerging in importance as these levels reflect gene expression, protein function, environmental factors and metabolic pathways [9]. In addition, perturbed metabolism is a hallmark of many diseases, e.g. cancer

and degenerative conditions [10]. While metabolite measurements from homogenized FFPE tissue using liquid based mass spectrometry have been conducted [11,12], liquid samples cannot provide information on molecular spatial distribution. This is particularly important as tissues are highly complex and heterogeneous systems. Given the ability to measure metabolites from FFPE samples using liquid mass spectrometry, we hypothesized that it should also be feasible in MSI.

We present the first study of *in situ* MALDI-FT-ICR MSI of metabolites from FFPE tissue samples. Using a simple experimental workflow and different comparison paradigms, we demonstrate the localization of different metabolites, which was used to answer tumour biology and clinical questions.

MATERIALS AND METHODS

Collection of tissue samples

Primary surgical resection specimens were obtained from patients diagnosed with colon adenocarcinoma (n = 102; all pT2 according to UICC classification [13]), breast cancer (n = 33), gastric adenocarcinoma (n = 102), renal oncocytoma (n = 21), chromophobe renal cell carcinoma (n = 11), and oesophageal adenocarcinoma (n = 53). Samples were collected between 1993 and 2010 at the Dept. of Surgery, Klinikum Rechts der Isar, Munich, Germany, and at Leiden University Medical Center, Leiden, the Netherlands. This study was approved by the local Ethics Committees; all patients provided informed, signed consent (Germany); Code of conduct for responsible use of human tissue for research (the Netherlands). Follow-up data were available for patients and survival rates calculated from the date of surgical resection to the date of last follow-up or death. All FFPE samples were collected in Munich.

Table S1 is a summary of patient numbers and sample type (FF or FFPE).

FFPE samples were fixed for 12-24 hours in 10% Neutral Buffered Formalin then embedded in paraffin using standardised and automated procedures. The samples were dehydrated in graded ethanol, cleared with xylene, and embedded in molten paraffin [14]. Fresh-frozen samples were stored post-surgery in liquid nitrogen until measurement. Frozen (12µm) and FFPE sections (4µm) were mounted onto indium-tin-oxide (ITO) coated glass slides (Bruker Daltonik, Bremen, Germany) pre-treated with 1:1 poly-L-lysine (Sigma Aldrich, Munich, Germany), and 0.1% Nonidet P-40 (Sigma).

MALDI imaging mass spectrometry

Prior to matrix application, FF sections were dehydrated in a vacuum for 1h at room temperature, FFPE sections were incubated for 1 hr at 70°C, de-paraffinized in xylene (2x 8 minutes), and allowed to air-dry. Matrix application and MALDI-MSI were conducted as previously described.[7] Briefly, frozen and FFPE samples were coated in 10mg/ml 9-aminoacridine matrix in 70% methanol using a SunCollect sprayer (Sunchrom, Friedrichsdorf, Germany) [7]. MSI was performed in negative ion mode on a Bruker Solarix 7T FT-ICR MS (Bruker), over a mass range of *m/z* 50 - 1000, and 60µm lateral resolution. All measurements included a non-tissue measurement region as a background control.

Following measurement, the matrix was removed with 70% ethanol, stained with HE, coverslipped, scanned with a Mirax Desk scanner (Zeiss, Göttingen, Germany) with a 20x magnification objective, and co-registered to respective MSI data using FlexImaging v. 4.0 (Bruker). Stained slides were examined by a pathologist and samples excluded if they contained no tumour or if the section was lost during HE staining. Using FlexImaging, acquired data were normalised against the root mean square of all data points, and histology-guided regions of interest (ROIs) were annotated (e.g. tumour or mucosa) and the spectral

data exported [7]. On-tissue MS/MS was conducted on FFPE tissue samples using continuous accumulation of selected ions mode and collision induced dissociation (CID) in collision cell [7]. To test the reproducibility we measured consecutive serial sections of breast cancer tissue microarrays (TMAs, n = 3; tissue cores = 40) on separate days.

Bioinformatics and Statistical Analysis

Subsequent data processing was performed with a MATLAB script using the bioinformatics and image processing toolboxes (v.7.10.0, MathWorks, Natick, MA, USA). All exported spectra underwent baseline subtraction to remove noise, resampling to lower data dimensionality, and smoothing to remove tissue and measurement artifacts [15]. Peak picking was conducted using a modified version of the LIMPIC algorithm [16], using *m/z* 0.0005 minimal peak width, signal to noise threshold of 2, and intensity threshold of 0.01%. Peaks were picked if found in 80% of the samples in each group (e.g. tumour or mucosa). Isotopes were automatically identified and excluded. To identify statistically significant differences in *m/z* values, the peaklists were analysed with the Mann-Whitney *U* test and Benjamini-Hochberg *post hoc* test (alpha = 0.05). If not otherwise stated, *m/z* values that were found also to occur in the non-tissue measurement region (acting as a background control) were excluded from these peak lists on the basis that *m/z* values that occurred in the non-tissue region could also be molecules from the slide surface.

PCA and heatmaps were created with MetaboAnalyst 3.0 (http://www.metaboanalyst.ca) [17]. Peak lists with respective intensities for ROIs were uploaded to MetaboAnalyst, processed with a mass tolerance of *m/z* 0.0001, no data filtering, and log transformed. The resultant PCA and heatmaps of metabolite data were exported as figures.

Significance Analysis of Microarrays (SAM) was used to identify disease-free survival-associated *m/z* values (q - value < 0.05). The average peak intensity was used to split patients

into good and poor survivor groups and significant differences in their survival were determined with the log-rank test. Multivariate survival analysis was performed using Cox proportional hazards regression models. All calculations were performed within R ('samr' and 'survival' packages). P-values <0.05 were considered significant.

Metabolite Identification

Metabolites were identified either by comparison to observed MS/MS spectra with standard compounds or by matching accurate mass with Metlin (http://metlin.scripps.edu/index.php) and the Human Metabolome Database (http://www.hmdb.ca) [18,19]. Search parameters were 3 ppm mass tolerance and $[M-H]^-$ charge. *In silico* analysis using Bruker DataAnalysis and SmartFormula was used to determine the molecular formula of *m/z* 256.9975. Raw spectra from the Barrett's TMA where m/z 256.9975 was highly intense were extracted and examined in Bruker DataAnalysis. The spectra were recalibrated using 9-AA as a reference peak ([M-H]- 193.07712), and then analysed using a FTMS peak finder. Using SmartFormula, H/C ratio was set from 0 to 3; the number of rings and double bonds were restricted to 0.5 to 40; and the option to automatically detect isotopic mass was activated. Molecular formulas were predicted allowing structures containing sulphate and phosphate; and M-H, M+H2O-H, M+Na-H2, M+K-H2 as negative adducts. Mass error tolerance was set at < 1 ppm; mass charge as -1. mSigma values – the match factors between the measured isotopic pattern and theoretical pattern of a given formula – less than 50 were considered a good match. A database search using Compound Crawler, including ChEBI, Chemspider, KEGG, and Metlin, was performed to produce compound structures.

RESULTS

In a first proof of principle study, we examined whether it was possible to conduct metabolite measurements from FFPE samples and if the detected compounds were comparable to those found in FF tissues. As stated, MALDI-MSI for FFPE protein and peptide content require many preparation steps [1]. In contrast, our workflow for FFPE metabolite imaging consisted only of deparaffination and matrix application (Figure 1).

Comparison of FF and FFPE tissue samples revealed a high overlap of metabolite content

Measurements from 21 FF and 81 FFPE colon cancer samples were compared. Automated peak picking found 1470 peaks in FFPE and 1465 peaks in FF colon cancer samples (Figure 2A), indicating that the number of detected peaks is not overly affected by sample preparation. 1226 peaks overlapped between the FFPE and FF datasets indicating that detected metabolite content of the samples is similar. FF and FFPE samples were measured over a mass range of *m/z* 50 - 1000. Examples of average overall spectra are presented in Figure 2B. Peaks in the *m/z* 50 - 400 range are more intense in the FFPE sample while peaks in the *m/z* 600 - 1000 range are more prominent in the FF sample. The mass range between *m/z* 50 – 400 is composed of non-lipid low molecular weight metabolites, while *m/z* 400 - 1000 consists mostly of lipids [20].

Reproducibility assessment

To examine reproducibility of our FFPE metabolite imaging approach, serial sections of breast cancer tissue microarrays (TMAs, $n = 3$; tissue cores = 40) were measured on separate days. TMAs are paraffin blocks containing samples from many different patients; an advantage of which is the simultaneous measurement of many samples in a single

experiment. Overall peaks of *m/z* values from each replicate, scatterplots of the intensities from individual cores and visualizations are presented in Figure S1A. Hierarchical clustering of the measurements was used to group five individual patients into separate clustering groups showing that our approach is reproducible (Figure S1B).

Spatial and chemical conservation of metabolites in tissue sections

Localizations for different *m/z* values in FF and FFPE tissues from colon cancer patients are presented in Figure 3. *m/z* 259.0140 (galactose-1-phosphate) localized to mucus. In contrast, hexose-6-phosphate (H6P; *m/z* 259.0230) is present across both samples but most intensely in tumour regions. Carnitine (*m/z* 160.8420) and *N*-Acetylglucosamine sulfate (*m/z* 300.0400) are localized to blood and tumour stroma, respectively. Figure S2 demonstrates the high spatial coherence between the histological features and *m/z* localization. Higher magnification images of single patient cores from the FFPE TMA show that the *m/z* distributions are spatially conserved even in small tissue sections (for further information see also Figure S3). This is a particular breakthrough as it has been thought that FFPE tissue processing would cause delocalization of molecules leading to inaccurate distributions [8]. Sample ion maps of other *m/z* values that occur in both FF and FFPE tissue are presented in Figure S4. It can be seen that most metabolites occur with similar intensities and localizations in both FF and FFPE. Interestingly, despite the embedding process, we could annotate some *m/z* values to lipids in the *m/z* 600 - 1000 range. Representative ion images of lipids are shown in Figure S4, for example *m/z* 599.3216 (phosphatidylinositol, PI(18:0/0:0)) and *m/z* 701.5150 (phosphatidic acid, PA(18:0/18:1)).

Identification and validation of metabolites

On-tissue MS/MS was conducted on FFPE sections to confirm that the detected peaks corresponded to metabolites. Focusing on nucleotides and metabolites from the central

carbon metabolism pathway, AMP, GMP, and GDP were confirmed by comparison of on-tissue MS/MS spectra to standard compounds (Figure S5). Hexose-6-phosphate, glycerol monophosphate and UDP were confirmed by accurate mass matching to databases.

Distinction between tumour and normal tissue specimens

To test whether metabolite profiles could distinguish between normal and diseased tissue, a TMA (n = 28 patient tissue samples) with matched normal and colon cancer tissue was measured and stained afterwards with haematoxylin and eosin (HE). Regions of interest (ROIs) for normal mucosa and tumour were annotated using histology as a guide. For the layout of the TMA, please refer to Figure S6A. Principal component analysis (PCA) scatterplots of the individual ROIs showed that tumour and mucosa sample clusters are distinct from each other (Figure S7). Mass spectra from these ROIs were then compared and statistical analysis found 182 significantly different *m/*z values (p <0.05). A heatmap of the top 25 significant *m/z* values demonstrates clear differences in the intensity patterns between tumour and normal mucosa (Figure 4A). Figure 4B shows peak intensity differences and visualizations for two of these *m/z* values (*m/z* 256.9975 and *m/z* 599.3200). These *m/*z values were also found to correspond to mucosa and tumour in FF tissue (Figure S4). A higher magnification of a normal mucosa sample indicates that *m/z* 256.9975 directly localizes to epithelium within the tissue. In another example, a gastric cancer TMA and a whole resected cancer sample, *m/z* 282.0292 coincided with desmoplastic stroma and *m/z* 444.0808 revealed molecular heterogeneity within the tumour (Figure S8).

Distinction between chromophobe renal cell carcinoma and renal oncocytoma

Chromophobe renal cell carcinoma (ChRCC) and renal oncocytoma have different clinical outcomes. The separation of these tumour types using light microscopy of conventional HE staining remains difficult in some cases. Therefore our metabolite imaging approach was

applied to discriminate between these tumour types that arise from renal tissue.

MALDI-FT-ICR MSI was conducted on TMAs consisting of 11 ChRCC and 21 oncocytoma cases. For the layout of the ChRCC and oncocytoma TMAs, please refer to Figure S6B and S6C, respectively. 123 peaks displayed significantly different intensities between the tumour types (p <0.05). As an example, phosphoethanolamine (*m/z* 862.6105) was detected with increased intensity in ChRCC samples compared to oncocytoma (Figure 5A). In contrast, 2-amino AMP (*m/z* 345.0720) was elevated in oncocytoma compared to ChRCC (Figure 5A). PCA (Figure 5B) was able to distinguish between ChRCC and oncocytoma samples, and a heatmap (Figure 5C) of the 50 most significant *m/z* values demonstrates that hierarchical clustering can differentiate between oncocytoma and ChRCC based on *m/z* expression patterns. Metabolite imaging can therefore differentiate between tumours of the same origin.

Survival analysis reveals an independent prognostic marker in oesophageal adenocarcinomas

We next investigated whether metabolite imaging could be used to address patient survival outcome. Tumour-specific spectral data was obtained from a TMA containing 53 oesophageal adenocarcinomas and survival analysis was conducted (Figure S9). The mean age at the time of surgery was 61 years, and the maximum follow up time was 162 months. Median disease-free survival time was 15 months, and median overall survival time was 18 months. The clinicopathological parameters of the tumour set are summarized in Table S2. Uni- and multivariate statistical analysis (with correction of cumulative alpha-errors) found *m/z* 256.9975 correlated significantly with disease-free survival, and constitutes an independent prognostic parameter in multivariate analysis, in comparison to the clinical standard for tumour staging, the TNM system according to UICC classification (Figure 6) [13]. As no database annotation was available for *m/z* 256.9975, *in silico* molecular formula

analysis utilising isotope fine structure was used to provide an identity. Analysis of the isotopic pattern predicts that *m/z* 256.9975 has a molecular formula $C_6H_{10}O_9S$ (mSigma 28.6). An integrated structure search found eight different compounds with the molecular formula $C_6H_{10}O_9S$ (Figure S10), which were deoxy sugar acids with sulphate esters.

DISCUSSION

This is the first study describing metabolite measurements with MALDI-FT-ICR MSI using FFPE tissues, thereby demonstrating a high-degree of chemical metabolite conservation in FFPE tissue samples. The comparison between FF and FFPE datasets indicate similar metabolite spatial localization in tissue sections. For example, hexose-phosphate was more concentrated in tumour than in normal tissue in both FF and FFPE samples (Figure 3), confirming the energy competition also known as Warburg effect – the increased rate of glycolysis in tumours [10]. In FFPE samples, peaks of metabolites were detected in low mass range (*m/z* 50 - 400) having similar intensity compared to FF. In contrast, peak intensities were significantly reduced above the mass range of *m/z* 600. This result corresponds with the expectation that tissue embedding process and the removal of paraffin wax by solvents (e.g. xylene) remove lipids, too. Despite this, we found a conservation of solvent-resistant lipids which remain present in formalin-fixed tissue (Figure S4). It is known that cellular components such as nucleic acids, carbohydrates, and lipids are not fixed directly by formaldehyde, but can be trapped in the network of cross-links between proteins [21]. In a previous study, it was suggested that methylene chains of free, unbound tissue lipids are removed from the tissue, while solvent-resistant lipids remain present in formalin-fixed tissue due to being locked into protein–lipid complex matrices, predominantly in the membranes [22].

Currently, the outcome of metabolic research based on FFPE tissue is very restricted although analysis of FFPE samples in a clinical and biological context is desirable. Commonly used techniques in metabolomics are nuclear magnetic resonance spectroscopy and liquid based MS (e.g. liquid chromatography MS and gas chromatography MS). Efforts were made to establish extraction and purification protocols for metabolite detection by liquid based MS [11,12]. But these protocols face challenges as liquid based MS techniques require large tissue amounts in order to achieve satisfactory metabolite yields for analyses. Additionally, tissue homogenizing leads to loss of specificity reducing significant results. An advantage of MSI is that spatial and chemical integrity are maintained and therefore enables detailed readouts from images by performing virtual micro-dissection (defined as region of interest) of spatially-resolved mass spectra. In a proof of principle this information was used to successful discriminate the metabolite content of normal and diseased tissue (Figure 4). Moreover, we determined the metabolic 'fingerprints' of chromophobe renal cell carcinoma (ChRCC) and renal oncocytoma using TMAs (Figure 5). The distinction between ChRCC and oncocytoma is challenging in some cases because both tumours share morphological and immunophenotypic features. ChRCC and renal oncocytoma are thought to arise from the distal tubuli of the kidney but have differing prognostic and clinical outcomes that influence treatment, making accurate diagnosis of utmost importance [23]. The separation of ChRCC and renal oncocytoma by metabolite signatures, as well as the possible high-throughput of large patient cohorts and the easy sample handling demonstrate advantages of MALDI FT-ICR MSI as a potential diagnostic tool. Furthermore, our approach was used to analyse clinical endpoints such as patient prognosis. Linking molecular data from 53 oesophageal adenocarcinomas by MSI to clinical data revealed an independent prognostic factor for overall survival (Figure 6). *In silico* analysis of the prognostic *m/z* value indicates that this molecule is most likely a deoxy sugar acid with ester sulphate. Previous studies have shown

that duodenal and gastric mucus contains acid polysaccharides with ester sulphate [24,25]. It has been reported that sulphated polysaccharides are associated with Barrett's oesophagus and Barrett's carcinoma [26]. Given the functions of sulphated polysaccharide structures as ligands, alterations in the sulphation of metaplastic cells, as present in Barrett's oesophagus, could have important effects on differentiation and malignant progression [27].

MALDI-MSI of FFPE tissue has so far concentrated only on protein and peptide analyses. Protocols for peptide measurements from TMAs are more complex and incorporate an antigen retrieval step, deparaffinization, *in situ* trypsin digestion and matrix application which can challenge the reproducibility. In this study we acquired reliable and reproducible data from TMAs and detected much more *m/z* species compared to the achieved number of molecules obtained in peptide measurements on FFPE tissues [28].

The ability to image metabolites is the starting point to investigating metabolic pathways in FFPE tissues, and should enable distinguishing the exact location of metabolic disturbances. In addition, the continual improvement in the mass resolution of MSI instrumentation may increase the likelihood of detecting low abundance metabolites that may be masked by the peaks of more abundant molecules. This would in turn lead to improved annotation of detected peaks. Prospectively, it is likely possible to measure hundreds of molecules and generate pathways from thousands of patients when using TMAs. Given that FFPE samples have been archived for decades in hospitals and universities around the world, this presents an unparalleled source of information on diseases. In summary, our study is therefore not only a major technical advance but may have direct impact on predicting patient outcome and diagnosis, and tissue-based research in general.

Acknowledgements

This project was funded by the Ministry of Education and Research of the Federal Republic of Germany (BMBF) (Grant No. 01ZX1310B) and the Deutsche Forschungsgemeinschaft (Grant Nos/ HO 1254/3-1, SFB 824 TP Z02 and WA 1656/3-1) to AW. Ulrike Buchholz, Claudia-Mareike Pflüger, Gabriele Mettenleiter, and Andreas Voss provided technical assistance. Daniela Borgmann provided bioinformatics assistance.

Author contributions:

AB and AL were responsible for the MALDI imaging mass spectrometry data acquisition, data analyses and wrote the manuscript. BB was involved in data analyses and provided bioinformatics assistance. NS and KG carried out validation experiments. MAi, AF and KPJ assisted in interpretation of the results. MAi, BB and HZ assisted in interpretation of the results and writing of the manuscript. KPJ, PJKK, CJHvdV, GW, FE, RL and MAu provided samples and assisted in histological interpretation of the samples. AW performed histopathological interpretation, conceived the study design, as well as wrote the manuscript. All authors approved the final version of this manuscript.

REFERENCES

1. Casadonte R, Caprioli RM. Proteomic analysis of formalin-fixed paraffin-embedded tissue by MALDI imaging mass spectrometry. *Nature protocols* 2011; **6**: 1695-1709.
2. Cancer Genome Atlas Research N. Comprehensive molecular characterization of gastric adenocarcinoma. *Nature* 2014; **513**: 202-209.
3. Wagle N, Berger MF, Davis MJ, *et al.* High-throughput detection of actionable genomic alterations in clinical tumor samples by targeted, massively parallel sequencing. *Cancer discovery* 2012; **2**: 82-93.
4. Norris JL, Caprioli RM. Analysis of tissue specimens by matrix-assisted laser desorption/ionization imaging mass spectrometry in biological and clinical research. *Chemical reviews* 2013; **113**: 2309-2342.
5. Rompp A, Spengler B. Mass spectrometry imaging with high resolution in mass and space. *Histochemistry and cell biology* 2013; **139**: 759-783.
6. Rompp A, Guenther S, Schober Y, *et al.* Histology by mass spectrometry: label-free tissue characterization obtained from high-accuracy bioanalytical imaging. *Angewandte Chemie* 2010; **49**: 3834-3838.
7. Sun N, Ly A, Meding S, *et al.* High-resolution metabolite imaging of light and dark treated retina using MALDI-FTICR mass spectrometry. *Proteomics* 2014; **14**: 913-923.
8. Nilsson A, Goodwin RJ, Shariatgorji M, *et al.* Mass Spectrometry Imaging in Drug Development. *Analytical chemistry* 2015.
9. Wishart DS. Proteomics and the human metabolome project. *Expert review of proteomics* 2007; **4**: 333-335.
10. Warburg O. On the origin of cancer cells. *Science* 1956; **123**: 309-314.
11. Kelly AD, Breitkopf SB, Yuan M, *et al.* Metabolomic profiling from formalin-fixed, paraffin-embedded tumor tissue using targeted LC/MS/MS: application in sarcoma. *PloS one* 2011; **6**: e25357.
12. Yuan M, Breitkopf SB, Yang X, *et al.* A positive/negative ion-switching, targeted mass spectrometry-based metabolomics platform for bodily fluids, cells, and fresh and fixed tissue. *Nature protocols* 2012; **7**: 872-881.
13. UICC. TMN Classification of Malignant Tumours. (Seventh ed). Wiley-Blackwell, 2009.
14. Ergin B, Meding S, Langer R, *et al.* Proteomic analysis of PAXgene-fixed tissues. *Journal of proteome research* 2010; **9**: 5188-5196.
15. McDonnell LA, van Remoortere A, de Velde N, *et al.* Imaging mass spectrometry data reduction: automated feature identification and extraction. *Journal of the American Society for Mass Spectrometry* 2010; **21**: 1969-1978.
16. Mantini D, Petrucci F, Pieragostino D, *et al.* LIMPIC: a computational method for the separation of protein MALDI-TOF-MS signals from noise. *BMC bioinformatics* 2007; **8**: 101.
17. Xia J, Psychogios N, Young N, *et al.* MetaboAnalyst: a web server for metabolomic data analysis and interpretation. *Nucleic acids research* 2009; **37**: W652-660.
18. Smith CA, O'Maille G, Want EJ, *et al.* METLIN: a metabolite mass spectral database. *Therapeutic drug monitoring* 2005; **27**: 747-751.
19. Wishart DS, Jewison T, Guo AC, *et al.* HMDB 3.0--The Human Metabolome Database in 2013. *Nucleic acids research* 2013; **41**: D801-807.
20. Wang J, Qiu S, Chen S, *et al.* MALDI-TOF MS Imaging of Metabolites with a N-(1-Naphthyl) Ethylenediamine Dihydrochloride Matrix and Its Application to Colorectal Cancer Liver Metastasis. *Analytical chemistry* 2015; **87**: 422-430.

21. Renshaw S. Immunochemical staining techniques. ed). Scion Publishing Ltd: Bloxham, UK, 2007; ch. 4.
22. Hughes C, Gaunt L, Brown M, *et al.* Assessment of paraffin removal from prostate FFPE sections using transmission mode FTIR-FPA imaging. *Anal Methods* 2014; **6**: 1028-1035.
23. Carvalho JC, Wasco MJ, Kunju LP, *et al.* Cluster analysis of immunohistochemical profiles delineates CK7, vimentin, S100A1 and C-kit (CD117) as an optimal panel in the differential diagnosis of renal oncocytoma from its mimics. *Histopathology* 2011; **58**: 169-179.
24. Andre F, Bouhours D, Andre C, *et al.* The carbohydrate and ester sulfate content of mucosal biopsies in health and in patients with duodenal ulcer. *Clinica chimica acta; international journal of clinical chemistry* 1974; **56**: 255-259.
25. Schrager J. Sulphated Mucopolysaccharides of the Gastric Secretion. *Nature* 1964; **201**: 702-704.
26. Endo T, Tamaki K, Arimura Y, *et al.* Expression of sulfated carbohydrate chain and core peptides of mucin detected by monoclonal antibodies in Barrett's esophagus and esophageal adenocarcinoma. *Journal of gastroenterology* 1998; **33**: 811-815.
27. Bodger K, Campbell F, Rhodes JM. Detection of sulfated glycoproteins in intestinal metaplasia: a comparison of traditional mucin staining with immunohistochemistry for the sulfo-Lewis(a) carbohydrate epitope. *Journal of clinical pathology* 2003; **56**: 703-708.
28. De Sio G, Smith AJ, Galli M, *et al.* A MALDI-Mass Spectrometry Imaging method applicable to different formalin-fixed paraffin-embedded human tissues. *Molecular bioSystems* 2015.

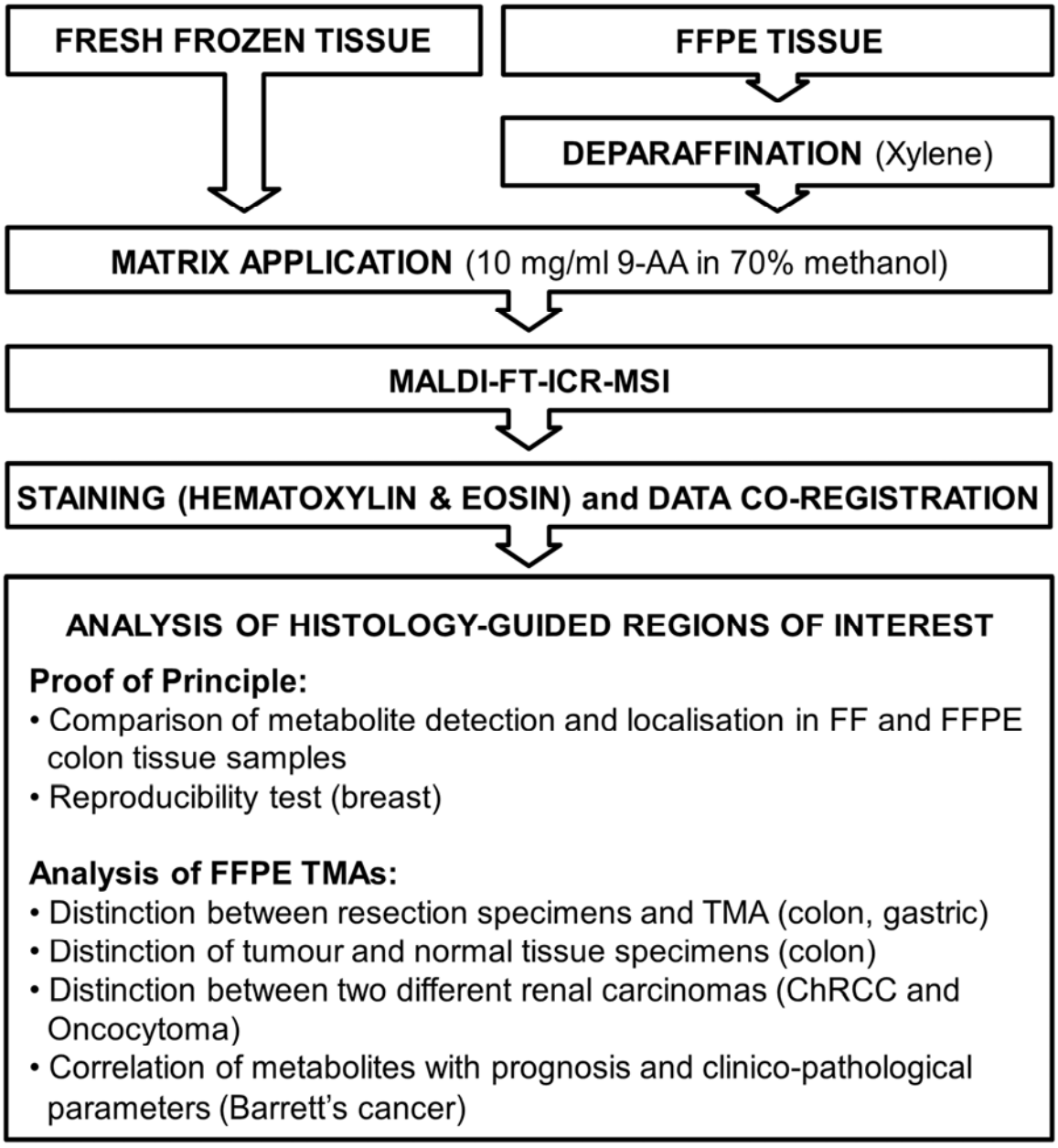

Figure 1. Scheme of the method and study design, and application of the analytical approach to important biomedical questions. (FF, fresh frozen; FFPE, formalin-fixed paraffin-embedded).

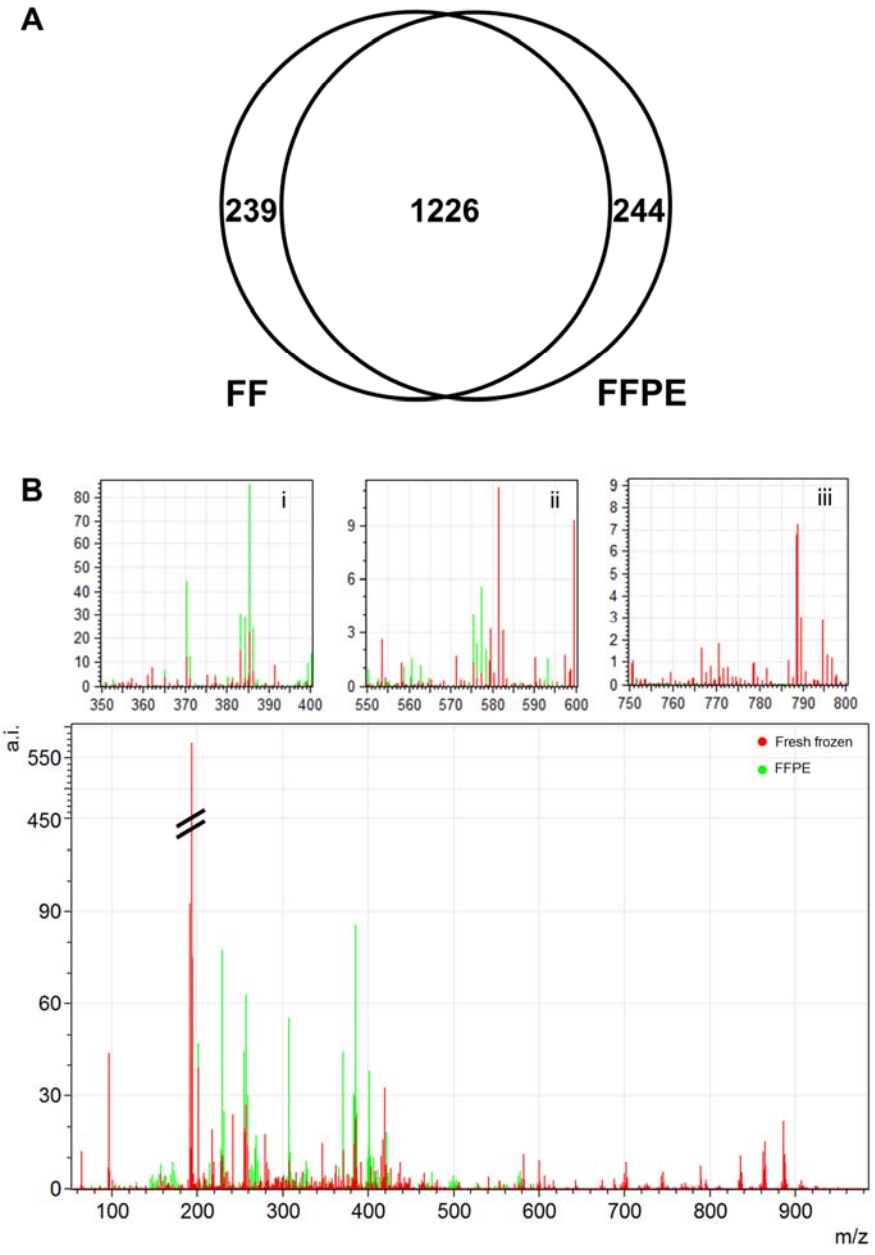

Figure 2. (A) Venn diagram showing the number of peaks found in FF tissue and FFPE samples. A total of 1465 peaks were detected in FF tissue and 1470 from FFPE, with an overlap of 1226 peaks found in both datasets (grey). **(B)** Examples of spectra from FFPE (green) and FF (red) samples. Inset: peaks from low (i), middle (ii) and higher (iii) mass ranges.

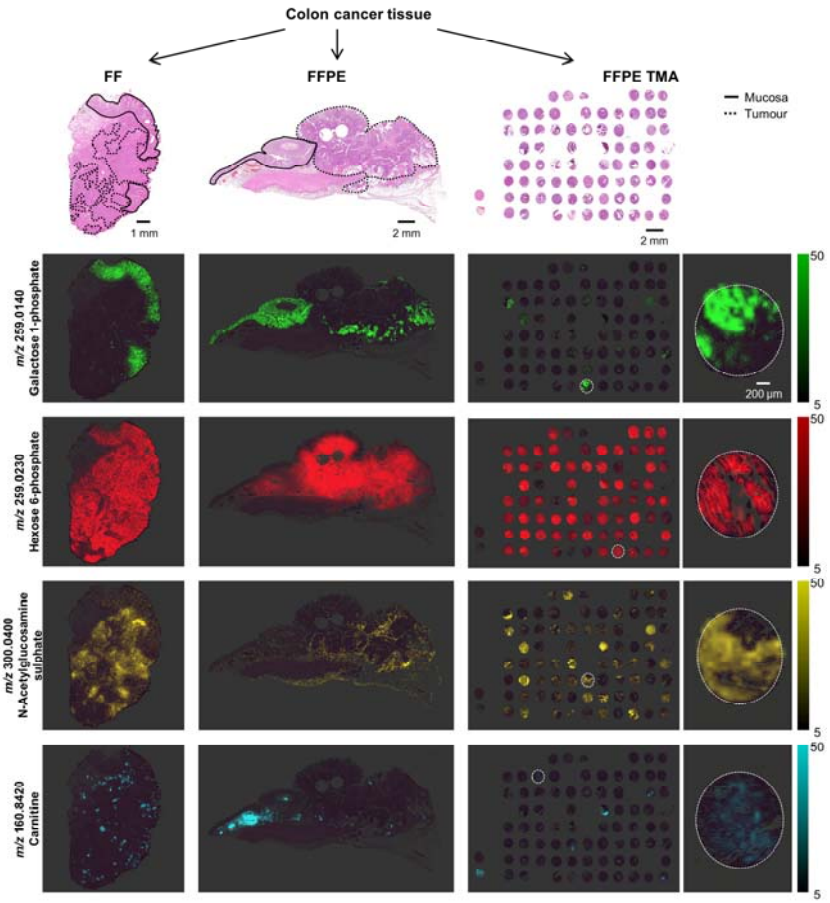

Figure 3. Comparison of *m/z* localisation in colon cancer samples from whole resected FF, FFPE, and FFPE TMA. Top panels are HE-stained FF, FFPE, and TMA colon cancer tissue sections showing regions corresponding to tumour (black dotted line), and mucosa (black solid line). *m/z* 259.0140 (galactose-1-phosphate; green) localises to mucus. *m/z* 259.0224 (H6P; red) is most intense in the tumour regions in FF and FFPE samples. *m/z* 300.0400 (*N*-Acetylglucosamine sulphate; yellow) corresponds to tumour stroma – tissue which supports tumour growth. *m/z* 160.8420 (Carnitine; blue) is found in the vicinity of blood. Higher magnification images of single patients samples from the TMA (right) show the *m/z* distributions are specific even in small tissue sections. White dotted lines indicate the magnified cores.

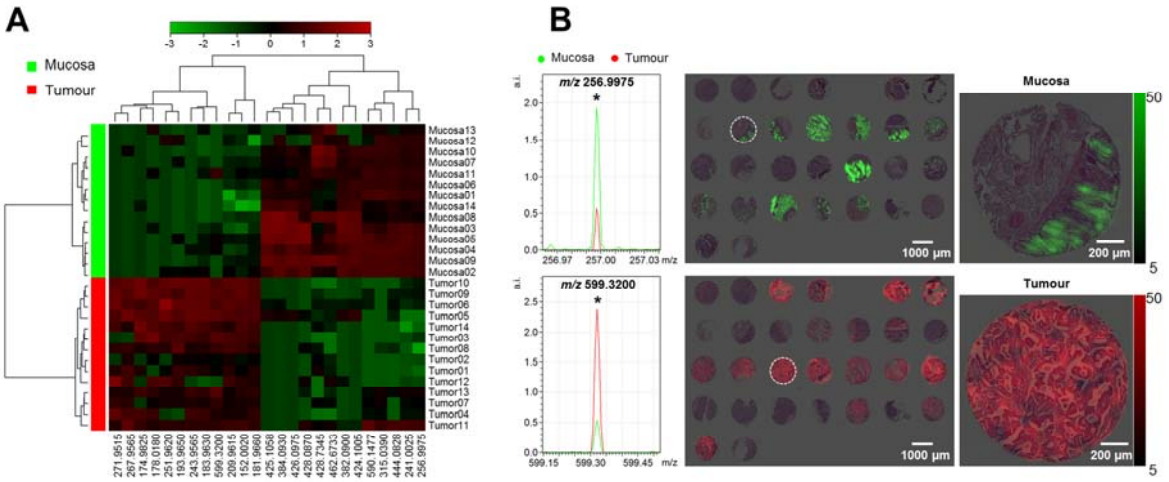

Figure 4. Discrimination of tumour tissue from normal colonic mucosa in FFPE TMA. **(A)** Heatmap of top 25 significant *m/z* values demonstrate different *m/z* expression patterns in mucosa vs. tumour. **(B)** Average spectra of *m/z* 256.9975 and *m/z* 599.3200 and localisation overlaid on corresponding HE-stained samples. *m/z* 256.9975 was more intense in mucosa (green) than tumour (red); a close up shows it to be specific for mucus-producing epithelium. *m/z* 599.3200 is more intense in tumour regions.

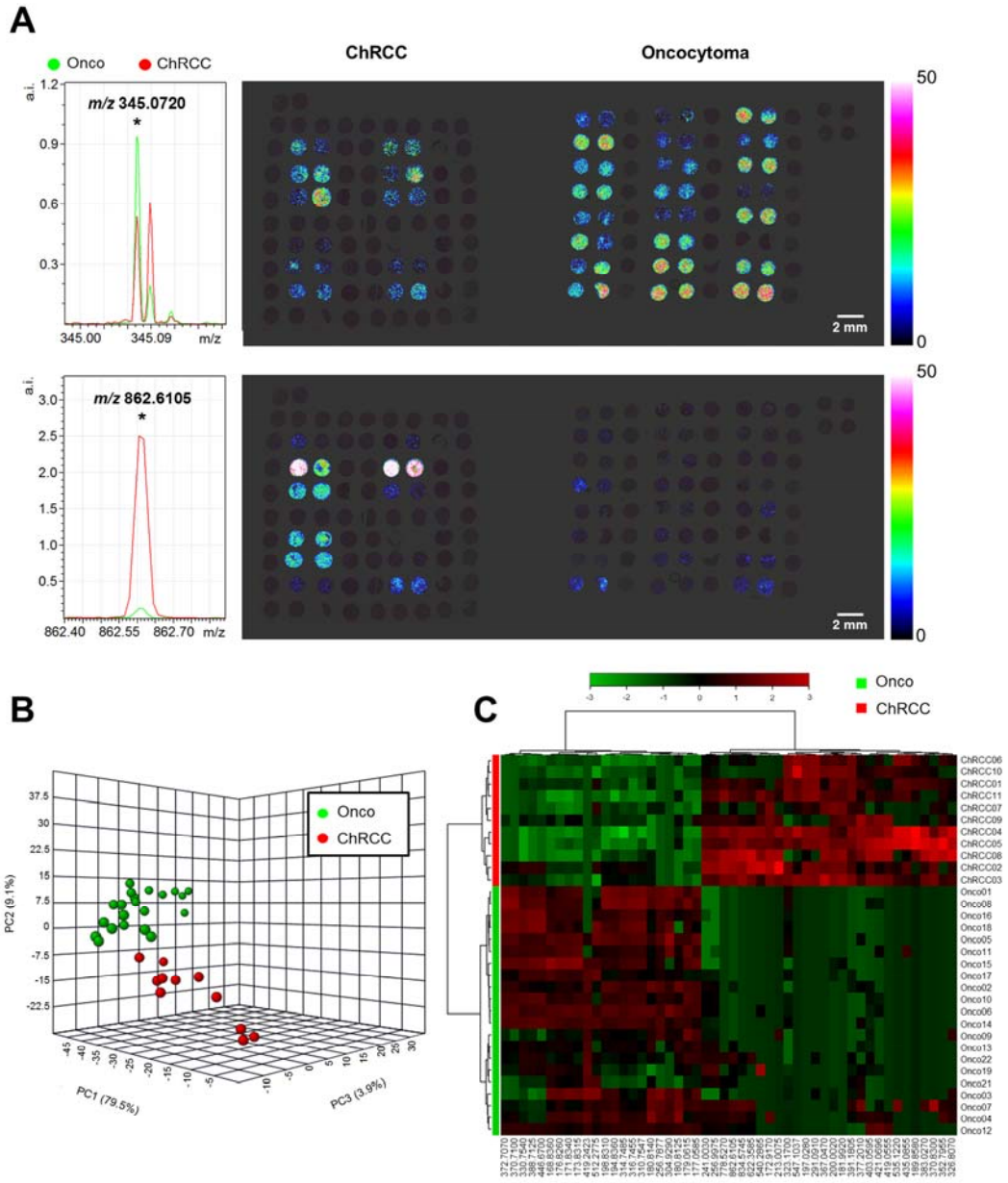

Figure 5. Metabolite MALDI-FT-ICR-MSI of renal oncocytoma vs. ChRCC. **(A)** Spectra and ion maps for *m/z* 862.6105 (phosphoethanolamine) and 345.0720 (2-amino AMP) show differences in distribution and intensities in oncocytoma and ChRCC. **(B)** PCA accurately distinguishes ChRCC (red) from oncocytoma (green). **(C)** Heatmap of top 50 differentially intense *m/z* values.

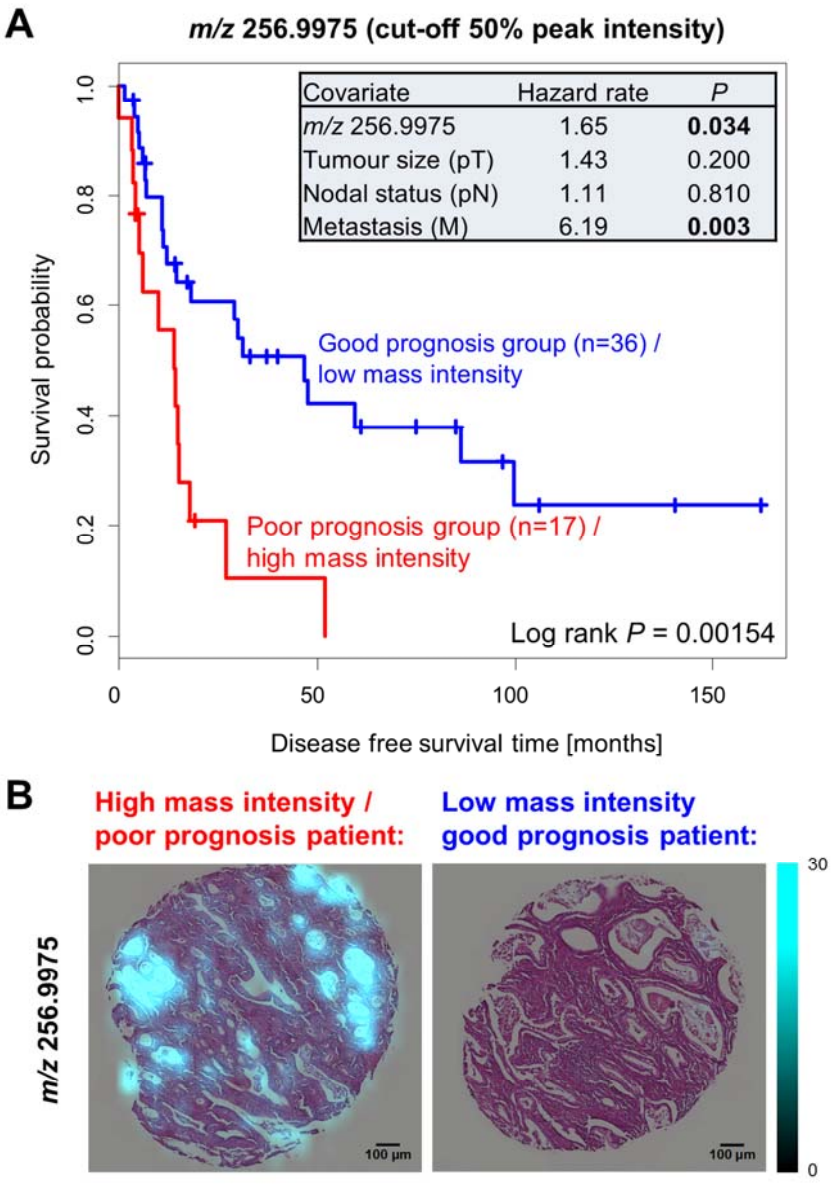

Figure 6. Metabolite imaging for addressing patient survival outcome. **(A)** Tumour-specific metabolomic data from a TMA containing 53 oesophageal adenocarcinomas was submitted to uni- and multivariate statistical analysis. *m/z* 256.9975 was found a significant prognostic factor for disease-free survival of the patients (P=0.00154), independently of other survival determinants given by the clinical TNM classification (A, inset; P=0.034). **(B)** Ion distribution maps showing localisation of deoxy sugar acid with ester sulphate (*m/z* 256.9975) in regions of mucus (visualization in blue).