LYMPH NODE METASTASIS OF PRIMARY ENDOMETRIAL CANCERS: ASSOCIATED PROTEINS REVEALED BY MALDI IMAGING.

Parul Mittal[1,2*], Manuela Klingler-Hoffmann[1,2*], Georgia Arentz[1,2], Lyron Winderbaum[1,2], Noor A Lokman[1,8], Chao Zhang[1,2], Lyndal Anderson[3], James Scurry[4], Yee Leung[5], Colin JR Stewart[5], Jonathan Carter[6], Gurjeet Kaur[7], Martin K. Oehler[8,9], Peter Hoffmann[1,2]#

[1]Adelaide Proteomics Centre, School of Biological Sciences, University of Adelaide, Adelaide, South Australia, 5005

[2]Institute for Photonics and Advanced Sensing (IPAS), University of Adelaide, Adelaide, South Australia, 5005

[3]Department of Tissue Pathology and Diagnostic Oncology, Royal Prince Alfred Hospital, Camperdown, New South Wales, 2050

[4]University of New South Wales, Kensington, New South Wales, 2033

[5]School of Women's and Infants' Health, University of Western Australia, Crawley, Western Australia, 6008

[6]Department of Gynaecological Oncology, Chris O'Brien Lifehouse, Camperdown, New South Wales 2050

[7]Institute for Research in Molecular Medicine, Universiti Sains Malaysia, 11800 Minden, Pulau Pinang, Malaysia

[8]Discipline of Obstetrics and Gynaecology, Research Centre for Reproductive Health, School of Medicine, Robinson Research Institute, University of Adelaide, Adelaide, South Australia, 5005

[9]Department of Gynaecological Oncology, Royal Adelaide Hospital, Adelaide, South Australia, 5000

[*]these authors contributed equally to the work

[#]To whom all correspondence and requests for reprints should be addressed:

Prof. Peter Hoffmann

Adelaide Proteomics Centre, University of Adelaide

Gate 8 Victoria Drive, Adelaide, South Australia, 5005

Phone: +61 (08) 8313 5507; Fax: +61 (08) 0 8313 4362

Email: peter.hoffmann@adelaide.edu.au

Keywords: Biomarker, Endometrial cancer, Lymph node metastasis, Proteomics

Abbreviations: EC (Endometrial cancer), LNM (Lymph Node Metastasis), FFPE (Formalin Fixed Paraffin Embedded), MALDI-MSI (Matrix Assisted Laser Desorption/Ionisation Mass Spectrometry Imaging), LC-MS/MS (Liquid Chromatography Mass Spectrometry), CCA (Canonical Correlation Analysis), LDA (Linear Discriminant Analysis), LOO (Leave One Out), α-Actin-2 (Alpha-actin-2)

Abstract

Metastasis is a crucial step of malignant progression and is the primary cause of death from endometrial cancer. However, clinicians presently face the challenge that conventional surgical-pathological variables, such as tumour size, depth of myometrial invasion, histological grade, lymphovascular space invasion or radiological imaging are unable to predict with accuracy if the primary tumour has metastasized. In the current retrospective study, we have used primary tumour samples of endometrial cancer patients diagnosed with (n=16) and without (n=27) lymph node metastasis to identify potential discriminators. Using peptide matrix assisted laser desorption/ionisation mass spectrometry imaging (MALDI-MSI), we have identified *m/z* values which can classify 88% of all tumours correctly. The top discriminative *m/z* values were identified using a combination of *in situ* sequencing and LC-MS/MS from digested tumour samples. Two of the proteins identified, plectin and α-Actin-2, were used for validation studies using LC-MS/MS data independent analysis (DIA) and immunohistochemistry. In summary, MALDI-MSI has the potential to identify discriminators of metastasis using primary tumour samples.

Statement of Significance

Endometrial cancer is the most common gynaecological malignancy in Australia with 2,256 diagnosed cases in 2010 and 381 associated deaths in 2011. The presence or absence of lymph node metastasis is the most important prognostic factor in early stage I endometrial cancer. Of the patients diagnosed with stage I disease, around 10% will have pelvic lymph node metastases (LNM). Despite the small percentage of patients who suffer from metastasis the majority undergo radical treatment including the removal of lymph nodes; a precautionary measure carried out due to our current inability to accurately stage the disease. Lymph node removal is associated with significant complications including lower extremity lymphoedema, occurring in up to 38% of patients. A classification system based around predictive tissue markers of metastasis is therefore essential to determine the optimal treatment strategy for endometrial cancer patients and to reduce disease morbidity. In this study we show that data acquired from the MALDI imaging of primary endometrial carcinomas can be used to successfully predict the presence or absence of LNM with an overall accuracy of 88.4%. The development of such a classification method shows the diagnostic potential of MALDI imaging in determining the metastatic potential of primary carcinomas.

1. Introduction

Endometrial cancer (EC) is the most frequent malignant tumour occurring in the female reproductive system. Based on the histology and clinical practice EC is divided into two subgroups: low-grade endometroid adenocarcinomas (Type I) and high-grade endometrioid and non-endometroid (Type II) carcinomas [1]. Type I EC accounts for 65% of all EC cases. These cancers are usually low grade, associated with oestrogen excess, obesity and atypical endometrial hyperplasia. Due to presence of early symptoms, the vast majority of type I carcinomas are diagnosed when the tumour is still confined to the uterus and these are cured by surgery in most cases resulting in a 5-year survival rate above 70% [2]. However, patients with recurrent EC have a 5 year survival rate below 40% despite intervention with chemotherapy and/or radiotherapy [2].

Currently accepted prognostic [prognostic = survival] factors for EC include the histological subtype, grade and International Federation of Gynecology and Obstetrics (FIGO) stage of the disease [3]. Deep (>50%) myometrial invasion and high histological grade are associated with the presence of lymph node metastasis (LNM) and adverse prognosis in EC [4]. Based on pathology findings, *Milam et al*. categorized EC patients as low risk for LNM if the tumour size was ≤ 2cm, well or moderately differentiated, and depth of myometrial invasion was ≤ 50%. Tumours that did not meet all three criteria were considered at high risk for LNM [5]. A study by *Jacques et al*. showed that a large percentage of EC will be misclassified before surgery [6]. Although only 15% of EC patients have or develop metastasis, the majority undergo radical treatment including removal of lymph nodes. This procedure is associated with significant complications including lower extremity lymphoedema, deep vein thrombosis and vascular or nerve injury [7].

Metastasis is a complex process in which tumour cells from the primary neoplasm acquire the ability to survive detachment, intravascular circulation, and implant and proliferate at a secondary site [8]. Recent studies in a variety of cancers have shown that metastatic potential of a primary tumour can be determined by genomic and proteomic analysis [9-12]. Moreover, in recent years a number of proteomic approaches have identified primary tumour signatures that accurately predict the presence of LNM, overall survival or disease recurrence [13-18]. Therefore, we hypothesised that the metastatic potential of a primary EC could be reflected at the proteome level and could be determined through the identification of molecular discriminators using MALDI mass spectrometry imaging (MALDI-MSI).

MALDI-MSI allows for the *in situ* characterisation of tissue sections, enabling the relative quantification and spatial expression profiling of thousands of peptides within and between tissues [19]. This technique allows a comparison of tissue histology with corresponding spatially resolved

mass spectrometric information. The identification of differentially expressed *m/z* values cannot be achieved from a standard MALDI-MSI experiment alone [20] and moreover, single *m/z* values can match several peptide masses from the corresponding LC-MS/MS data of the tissue extract, even when using relative high mass accuracy in the MALDI-MSI experiment using internal calibrants. Although, *in situ* MS/MS data are often poor and don't reveal direct identification; combining the two methods and matching *m/z* values of the intact peptides and y and b ions from both fragmentation patterns provides identification in most cases.

Recently, we have shown the capacity of MALDI-MSI to discriminate regions of healthy endometrial tissue from tumour [21]. Here we present an analysis of primary EC specimens with (n=16) and without LNM (n=27) by MALDI-MSI. Upon data acquisition, a Canonical Correlation Analysis (CCA) based method was applied to rank the intensities of the acquired MALDI *m/z* values based upon their power to discriminate the primary carcinomas with metastasis from those without. This ranking was used to reduce the dimension of the data to the top ranked *m/z* values prior to classification by linear discriminant analysis (LDA), and the performance of this classification was judged by leave one out (LOO) cross validation (for details see *Winderbaum et al*.[22]).

The top *m/z* values were targeted for identification using the complementary techniques of *in situ* MALDI MS/MS and matching to peptide sequences obtained from traditional nano-flow liquid chromatography electrospray ionization tandem mass spectrometry (nanoLC-ESI-MS/MS). The differential expression of plectin and α-Actin-2 between the primary carcinomas with and without LNM was further validated using label-free quantitative LC-MS/MS and immunohistochemistry. In summary we have identified α-Actin-2 as a potential discriminator of increased risk of LNM in EC.

2. Material and Methods

2.1. Sample collection and Tissue specimens

Formalin-fixed paraffin-embedded (FFPE) tissue samples were retrieved from the archives of the Institute of Medical and Veterinary Science, Adelaide, South Australia, Royal Prince Alfred Hospital, Sydney, New South Wales, John Hunter Hospital, Newcastle, New South Wales and King Edward Memorial Hospital, Perth, Western Australia. The study was approved by the ethics committees of the different institutions. The histo-morphological and clinical information for the patients is provided in Supplementary Table 1.

2.2. Tissue microarray construction

TMAs were constructed as previously described [23]. Primary tumour sections were annotated by a pathologist and two cores representing tumour centre were used to construct two TMA's. However, after TMA construction, all medical records were carefully reviewed by a clinician again and samples with mixed carcinoma were excluded from the study. This led to overall exclusion of 14 samples from 57 patient samples and the final cohort was comprised of a total of 43 patients (n=16 with LNM and n=27 without LNM), which were assembled in two TMAs (named TMA1 and TMA2). Serial 6µm sections were mounted onto indium titanium oxide (ITO) conductive glass slides for MALDI-MSI analysis (Bruker Daltonics, Bremen, Germany). Sections were also placed on plain glass slides and Polyethylene Naphthalate (PEN) membrane slides (MicroDissect, Herborn, Germany), and haematoxylin and eosin (H&E) stained for pathology annotation and laser microdissection (LMD) respectively.

2.3. Sample preparation for FFPE MALDI-MSI

MALDI-MSI was carried out as previously described [24] in duplicate on consecutively cut TMA sections. Briefly, the tissues were deparaffinised, subjected to heat induced citric acid antigen retrieval (HIAR) (10 mM citric acid, pH=6), followed by digestion with trypsin gold (Promega, Madison, WI) at 37°C for 2 hours. Internal calibrants and α-cyano-4-hydroxycinnamic acid (CHCA) matrix were overlayed using an ImagePrep station [25].

2.4. TMA analysis by MALDI-MSI

MALDI imaging of the TMAs was carried out using an ultrafleXtreme MALDI-TOF/TOF MS (Bruker Daltonics, Bremen, Germany) with flexControl v3.0.1 and flexImaging v4.0.1 software (Bruker Daltonics) in positive reflectron mode over a detection range of *m/z* 800-4000 Da. MS spectra was acquired in a raster based grid with a centre to centre resolution of 60 µm. Technical replicates of each TMA were measured. After data acquisition, the matrix was removed with 70% ethanol, the TMA cores were H&E stained, and digitally scanned using a Nanozoomer (Hamamatsu Photonics, Shimadzu, Japan) and images were obtained using imaging software (NDP scan software v2.2, Hamamatsu Photonics). To align the MS data with the tissue histology the H&E scanned cores were co-registered with the MALDI-MSI results and annotated using the flexImaging software. Tissue regions containing only areas of primary tumour were selected and the spectra lists for these regions were exported as .XML files.

2.5. MALDI-MSI Canonical Correlation Analysis (CCA)

A detailed description of the CCA method can be found in *Winderbaum et al*.[22]. Briefly, arbitrarily located bins with a width of 0.25 Da were used to discretise *m/z* domains in order to group peaks. The intensity values of the peaks in each of these defined peak groups (*m/z* bin) were log-transformed and

averaged across the annotated tumour areas for each patient. These averages were assembled into a data matrix with columns representing each analysed patient and rows corresponding to the *m/z* bins. The rows of this matrix were then ranked using the developed CCA based method for their ability to distinguish between primary carcinomas with and without LNM. A dimension reduced submatrix consisting of the top ranked rows was then analysed using linear discriminant analysis (LDA) in order to predict the LNM status of the 43 patients. All analyses were replicated in parallel using two alternate (shifted) bin locations, resulting in three analyses in total. A majority rule was used to combine the data.

2.6. Data analysis using SCiLS lab

For data dependent visualization of tissue morphological regions, raw data was uploaded into the SCiLS lab software (v2015a, GmbH, Bremen, Germany). Here the data was pre-processed including top hat baseline removal and total ion count (TIC) normalization, and peak alignment and picking was performed [26]. The spatial expression profiles of the *m/z* values found to have discriminative power in the CCA were visualised in the form of ion intensity maps, and receiver operating characteristic (ROC) curves of these *m/z* values comparing the intensities of the tumours with and without LNM was generated.

2.7. Identification of m/z values by in-situ MALDI MS/MS and nanoLC-ESI-MS/MS

In order to gain peptide identifications for the *m/z* values of interest, *in situ* MS/MS was performed directly from the tissue used in the MALDI-MSI analysis and searched using Mascot (Version 2.3.02) as previously described [24]. For matching back to peptide sequences obtained by data dependent acquisition, nanoLC-ESI-MS/MS, primary tumour and normal tissue regions of interest were collected from the TMA cores using LMD, subjected to HIAR, and digested with trypsin [23]. NanoLC-ESI-MS/MS was performed using an Ultimate 3000 RSLC system (Thermo-Fisher Scientific) coupled to an Impact II™ QTOF mass spectrometer (Bruker Daltonics) via an Advance CaptiveSpray source (Bruker Daltonics). Peptide samples were pre-concentrated onto a C18 trapping column (Acclaim PepMap100 C18 75 µm × 20 mm, Thermo-Fisher Scientific) at a flow rate of 5 µL/ min in 2% ACN 0.1% TFA for 10 minutes. Peptide separation was performed using a 75 µm ID C18 column (Acclaim PepMap100 C18 75 µm × 50 cm, Thermo-Fisher Scientific) at a flow rate of 0.2 µL/ min using a linear gradient from 5 to 45% B (A: 5% ACN 0.1% FA, B: 80% ACN 0.1% FA) over 130 minutes, followed by a 20 minute wash with 90% B, and a 20 minute equilibration with 5% A. MS scans were acquired in the mass range of 300 to 2200 *m/z* in a data-dependent fashion using Bruker's Shotgun Instant Expertise™ method.

Singly charged precursor ions were excluded from acquisition. Collision energy ranged from 23% to 65% as determined by the *m/z* of the precursor ion.

Acquired spectra were subjected to peak detection, de-convolution, and re-calibration according to a lock mass using Compass DataAnalysis for OTOF (Version 1.7, Bruker Daltonics). Processed spectra were then exported to Mascot generic format and submitted to Mascot (Version 2.3.02) for identification. Search parameters were as follows; SwissProt Homo sapiens database search, trypsin digestion with up to 2 missed cleavages, variable modification of oxidation of methionine, MS mass tolerance of 40 ppm and a MS/MS mass tolerance of 0.2 Da. In Mascot an ion score cut off of 20 with a peptide significance threshold of ≤0.05 was used, which corresponds to a false discovery rate (peptide level) of <2%. Matching between the MALDI-MSI and nanoLC-ESI-MS/MS was done by comparing the experimental *m/z* values of the nanoLC-ESI-MS/MS sequenced peptides to the *m/z* values from the compiled peak bins.

2.8 Quantification of peptides by data independent acquisition (DIA) nano-LC-ESI-MS/MS results

DIA nano-LC-ESI-MS/MS was performed on LMD normal and cancer tissues from 4 patients with, and 4 patients without LNM. Nano-LC was performed as described above using an Ultimate 3000 RSLC system coupled to an Impact II™ QTOF mass spectrometer. The Impact II™ QTOF acquired data using Bruker's Middle Band CID™ method where a mass range of *m/z* 375 to 1206 is scanned in 26 Da increments with increasing collision energies of 20 to 36. Data were analysed in the Skyline software against a spectral library generated from the previous nano-LC-ESI-MS/MS experiments [27]. The peptide and transition settings during analysis were as follows; trypsin was specified as the cleavage enzyme with a maximum of 1 missed cleavage, precursor charge states 2 and 3, ion charges 1 and 2, ion types y and b from ion 3 to 6, ion match tolerance 0.1 *m/z*, a MS/MS filtering DIA isolation scheme from *m/z* 400 - 1206 (26 Da windows), a resolution 10 000, and only scans within 5 minutes retention time window of spectral library MS/MS identification used. Summed area intensities for the analysed peptides were calculated from y and b ions 3 to 6 (starting from ion 3). For each peptide analysed, the relative intensity in the tumour tissue was normalised to the relative intensity of the normal tissue from each patient.

2.9. Immunohistochemistry (IHC)

For the analysis of plectin and α-Actin-2 by IHC, 6 µm TMA sections were analysed as previously described [28]. Briefly, the tissue sections were dewaxed, rehydrated with xylene and ethanol and subjected to microwave antigen retrieval for 10 minutes at 100^{o}C (Sixth Sense, Whirlpool, VIC, Australia) in 10mM citric acid buffer pH=6. TMA sections were incubated overnight at 4^{o}C with either

α-Actin-2 (1/500, rabbit polyclonal, ProteinTech, Chicago, USA) or plectin (1/250, rabbit monoclonal, Abcam, MA, USA) in blocking buffer (5% goat serum), followed by incubation with biotinylated anti-rabbit immunoglobulin (1/400, Dako, NSW, Australia) and streptavidin-HRP (1/500, Dako). Immunoreactivity was detected using diaminobenzidine (DAB)/H_2O_2 (Sigma Aldrich) substrate and counterstaining with haematoxylin (Sigma Aldrich). TMA slides were digitally scanned using a Nanozoomer and images were obtained using NDP view imaging software. Analysis was carried out in IHC Profiler-Image J [29]. For each tissue core, three representative photo-micrographic images at 40x magnification were analysed.

3. Results and Discussion

3.1. MALDI MSI

MALDI-MSI was carried out on two TMA (TMA1 and TMA2), two replicates per patient were used resulting in a total of 86 primary endometrial carcinoma tissue cores from 43 patients with (n=16) and without (n=27) LNM. Peak groups were generated from the MALDI-MSI data and then ranked using a CCA based method for their ability to distinguish between the primary carcinomas with and without LNM. A list of the top *m/z* bins (peak groups) with the capacity to differentiate the primary cancer types is shown in Supplementary Table 2. Reducing the data to these *m/z* values, and using LDA to discriminate between primary carcinomas with and without LNM, a classification accuracy of 38 out of 43 patients (88.4%) was achieved by LOO cross validation (for details see *Winderbaum et al*.[22]).

3.2. Identification of discriminative m/z values

The top discriminating *m/z* bins of 0.25 Da were centred at: 802.42, 857.42, 915.42, 941.42, 944.42, 967.42, 975.42, 976.42, 1027.67, 1032.67, 1115.42, 1138.67, 1157.67, 1161.67, 1167.67, 1198.67, 1242.67, 1406.67and 1612.92 (in order of size). Potential peptide identifications for the top *m/z* bins, as ranked by the CCA, were compiled by matching back to sequences obtained by data dependent acquisition nanoLC-ESI-MS/MS (Supplementary Table 2). Of the 20 *m/z* bins, 3 had no matches back to the tandem MS data, with the remaining 17 having 2 or more sequence matches. In order to confirm peptide identifications, the *m/z* values were targeted for *in situ* MS/MS directly off the tissue, from which 2 peptides could be verified; *m/z* 1198 AVFPSIVGRPR (α-Actin-2), and *m/z* 976 AGFAGDDAPR (α-Actin-2).

α-Actin-2 was targeted for further analysis given 3 of the top *m/z* bins matched to α-Actin-2 peptides. Moreover, the *m/z* value of 1501.42 was proteotypic for α-Actin-2 (data not shown). The *m/z* 967.42 matching to plectin was selected given only two possible nano-LC-ESI-MS/MS sequence matches were obtained for this *m/z* value, and of these two matches the relative abundance of the plectin related

peptide was significantly greater in the nano-LC-ESI-MS/MS data than the alternative candidate and matched the high abundance in the MALDI-MSI data (data not shown).

Figure 1 and 2 shows the ion intensity images for *m/z* 967.42 and 976.42 across the TMA. Tumour cores with LNM are circled red; tumours without LNM are circled green, yellow circles indicate controls and blue circles indicate cores which have been excluded from the analysis. All tumour regions within the tumour cores are annotated in the corresponding colour. Magnification of two replicate cores from one representative patient with LNM (Figure 1B, 2B) and without LNM (Figure 1C, 2C) reveals the expression of the *m/z* value within the tissue cores (H&E stain top panel, ion intensity images bottom panel). Although *m/z* 967.42 was included in the list of the top 20 list, SCiLS analysis only revealed a slight difference between the ion intensity between tumours with and without LNM (Figure 1). However, the SCiLS analysis of *m/z* 976.42 revealed a difference between the ion intensity map between tumours with and without LNM (Figure 2).

3.3. Validation by DIA

The MALDI-MSI results for the α-Actin-2 and plectin peptides were verified by DIA nano-LC-ESI-MS/MS. Analysis was carried out on LMD on normal and primary tumour tissue from 4 patients with LNM and 4 without LNM, who had not been included in the TMA analysis. DIA allows the differential quantification of the isobaric peaks by matching the retention time, as the unique fragment ions of the two species generate two different MS/MS chromatograms [30]. Therefore, retention time was used when matching back the data to the spectral library for DIA. The area intensities of the peptides matching back to the α-Actin-2 and plectin were summed for both tumour and normal tissues. The summed area intensity of the tumour was then normalised to that of the paired normal tissues. The normalised tumours were then compared with and without LNM using an unpaired T-Test in GrahPad Prism. A trend of increased expression in the primary tumours without LNM was observed for both plectin and α-Actin-2 peptides (Figure 3).

3.4. Validation by immunohistochemistry

The spatial expression profile of plectin was verified across the patient cohort by immunohistochemistry (Figure 4A). Quantitative analysis of immunostaining was performed using IHC Profiler-Image J [29] and as expected this indicated no difference between tumours with and without LNM; shown is the average staining intensity across all tumour cores (Figure 4B). Plectin scarcely stained normal tissue including stroma (Figure 4C); the staining of tumours cells was strong (Figure 4C-E). In summary plectin IHC can highlight tumour cells, but staining intensity does not distinguish cases with and without LNM.

This is in contrast to data presented in the human proteome atlas (Version 14, updated 2015-10-16), where a medium intensity stain of plectin was detected in healthy endometrium and absent or low staining was usually detected in EC with one of three antibodies. However, the other antibodies showed different staining patterns, making a precise interpretation of the data difficult.

It is known that tumour cell motility is required for invasion and metastasis [31, 32]. Plectin, has been found to be important in cytoskeletal network organization [33]. A number of studies have shown that increased levels of plectin correlate with migration and invasion [33-36]. However, our data show an increase in plectin staining in EC tumour cells unrelated of their metastatic potential. Furthermore, we have identified a number of peptides from the protein plectin which showed potential to discriminate between tumours with and without LNM. One representative *m/z* 967.4 (amino acid 1045-1052) is shown in Figure 1 and has been confirmed by DIA (Figure 3). However, the immunohistochemistry analysis failed to identify a difference in staining intensity of tumours with LNM when compared to tumours without LNM. A recent study has identified S1047 in plectin as a potential phosphorylation site offering one possible explanation of the discrepancy of results [37].

The spatial expression profile and differential expression of α-Actin-2 was verified across the patient cohort by immunohistochemistry (Figure 5A). Quantitative analysis of staining was performed using IHC Profiler-ImageJ and a significant difference in the negative staining (1.8 fold, $p<0.05$) between tumours without and with LNM was observed; shown is the average staining intensity across all tumour cores (Figure 5B). α-Actin-2 stained normal tissue (Figure 5C),while the staining of tumours with LNM was reduced (Figure 5D) when compared to tumour without LNM (Figure 5E). In summary α-Actin-2 IHC staining intensity has the potential to distinguish between tumours with and without LNM.

This is in agreement with the data presented in the human proteome atlas (Version 14, updated 2015-10-16), where a medium intensity stain of α-Actin-2 was detected in healthy endometrium and absent or low staining was detected in EC with four different antibodies.

Cytoskeletal proteins facilitate the biological modes of cells: migration, cell division, differentiation and cell death. It is therefore not surprising that these proteins are frequently identified in comparative proteomic studies. α-Actin-2, the human aortic smooth muscle actin gene, is one of six different actin isoforms which have been identified and has been described to facilitate migration of cells. A number of studies have shown that increased expression of α-Actin-2 leads prevents cellular motility [38, 39]. Accordingly, decreased expression of α-Actin-2 has been shown to contribute to the

metastatic potential of basal cell carcinoma [40]. Our data indicate a down-regulation of α-Actin-2 in tumours with LNM and therefore α-Actin-2 may have potential as a biomarker for EC metastasis.

4. Concluding remarks

In summary, MALDI-MSI has the potential to identify the markers of tumour metastasis by providing spatial intensity of proteins/peptides that might be associated with different tissue types and facilitate developing disease. Using MALDI-MSI data, we found a number of *m/z* values that could predict the status of LNM with an overall accuracy of 88.4%. Additionally, the *m/z* values were identified as α-Actin-2 and plectin via *in situ* MS/MS (Supplementary Figure 1) and label free quantification (Supplementary Table 2). Furthermore, DIA (Supplementary Figure 2-6) and immunohistochemistry was used for relative quantification and validation. The role of α-Actin-2 and plectin in metastasis has already been described previously and could be useful as potential biomarkers for distinguishing EC with and without LNM.

Figure 1: Representative MALDI-MSI images for *m/z* 967.42± 0.125 Da. (**A**) Overview of one TMA slide, MALDI-MSI image of (intensity range from blue (lowest) to red (highest)). The samples belonging to different groups are indicated by different coloured circles: control (yellow), with LNM (red), without LNM (green) and mixed carcinoma/not included in the study (blue). The tumour regions (red) within the samples have been annotated by a pathologist. (**B**) Magnification of two cancer tissue spots with LNM showing H&E stain and the ion intensity images of *m/z* 967.42. (**C**) Magnification of two cancer tissue spots without LNM showing H&E stain and the ion intensity images of *m/z* 967.42. (**D**) MALDI-MSI spectra displaying the mean spectrum from regions with LNM (red) and without LNM (green). (**E**) ROC curve with AUC of 0.396

Figure 2: Representative MALDI-MSI images for *m/z* 976.42± 0.125 Da. (**A**) Overview of one TMA slide, MALDI-MSI image of (intensity range from blue (lowest) to red (highest)). The samples belonging to different groups are indicated by different coloured circles: control (yellow), with LNM (red), without LNM (green) and mixed carcinoma/not included in the study (blue). The tumour regions (red) within the samples have been annotated by a pathologist. (**B**) Magnification of two cancer tissue spots with LNM showing H&E stain and the ion intensity images of *m/z* 976.42. (**C**) Magnification of two cancer tissue spots without LNM showing H&E stain and the ion intensity images *m/z* 976.42. (**D**) MALDI-MSI spectra displaying the mean spectrum from regions with LNM (red) and without LNM (green). (**E**) ROC curve with AUC of 0.313.

Figure 3: DIA analysis of tumour sections with (n=4) and without (n=4) LNM. (**A**) The relative abundance of plectin was analysed in comparison to normal tissue set to 1.0. (**B**) The relative abundance of α-Actin-2 was analysed in comparison to normal tissue set to 1.0. The error bar indicates the standard deviation.

Figure 4: Immunohistochemical staining of plectin (**A**) 6µm serial section of TMA 1 was used for immunohistochemistry (IHC). The different tissue types are encircled control (yellow), with LNM (red), without LNM (green) and mixed carcinoma/not included in the study (blue). (**B**) Quantitative analysis was performed using IHC profiler-Image J. For each tissue section, three representative photo-micrographic images at 40x magnification were used and each image was assigned a score of high positive, positive, low positive and negative staining. Shown is the average staining intensity distribution of all analysed images Representative image of plectin immunostaining of normal tissue (**C**), tumour with LNM (**D**) without LNM (**E**).

Figure 5: Immunohistochemical staining of α-Actin-2 (**A**) 6µm serial section of TMA 1 was used for immunohistochemistry (IHC). The different tissue types are encircled control (yellow), with LNM (red),

without LNM (green) and mixed carcinoma/not included in the study (blue). (**B**) Quantitative analysis was performed using IHC profiler-Image J. For each tissue section, three representative photo-micrographic images at 40x magnification were used and each image was assigned a score of high positive, positive, low positive and negative staining. Shown is the average staining intensity distribution of all analysed images. Representative image of α-Actin-2 immunostaining of normal tissue (**C**), tumour with LNM (**D**) and without LNM (**E**).

5 References

[1] Bokhman, J. V., Two pathogenetic types of endometrial carcinoma. *Gynecologic oncology* 1983, *15*, 10-17.

[2] American Cancer Society 2015.

[3] Larson, D. M., Connor, G. P., Broste, S. K., Krawisz, B. R., Johnson, K. K., Prognostic significance of gross myometrial invasion with endometrial cancer. *Obstetrics and gynecology* 1996, *88*, 394-398.

[4] Boronow, R. C., Morrow, C. P., Creasman, W. T., Disaia, P. J. *, et al.*, Surgical staging in endometrial cancer: clinical-pathologic findings of a prospective study. *Obstetrics and gynecology* 1984, *63*, 825-832.

[5] Milam, M. R., Java, J., Walker, J. L., Metzinger, D. S. *, et al.*, Nodal Metastasis Risk in Endometrioid Endometrial Cancer. *Obstetrics and gynecology* 2012, *119*, 286-292.

[6] Jacques, S. M., Qureshi, F., Munkarah, A., Lawrence, W. D., Interinstitutional surgical pathology review in gynecologic oncology: I. Cancer in endometrial curettings and biopsies. *International journal of gynecological pathology : official journal of the International Society of Gynecological Pathologists* 1998, *17*, 36-41.

[7] Creasman, W. T., Odicino, F., Maisonneuve, P., Quinn, M. A. *, et al.*, Carcinoma of the Corpus Uteri. *International Journal of Gynecology & Obstetrics* 2006, *95, Supplement 1*, S105-S143.

[8] Lodish H, B. A., Zipursky SL, et al., *Molecular Cell Biology*, W. H. Freeman, New York 2000.

[9] Meding, S., Nitsche, U., Balluff, B., Elsner, M. *, et al.*, Tumor classification of six common cancer types based on proteomic profiling by MALDI imaging. *Journal of proteome research* 2012, *11*, 1996-2003.

[10] Meding, S., Balluff, B., Elsner, M., Schöne, C. *, et al.*, Tissue-based proteomics reveals FXYD3, S100A11 and GSTM3 as novel markers for regional lymph node metastasis in colon cancer. *The Journal of pathology* 2012, *228*, 459-470.

[11] Hristov, A. C., Cope, L., Reyes, M. D., Singh, M. *, et al.*, HMGA2 protein expression correlates with lymph node metastasis and increased tumor grade in pancreatic ductal adenocarcinoma. *Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc* 2009, *22*, 43-49.

[12] Ji, F., Jin, X., Jiao, C. H., Xu, Q. W. *, et al.*, FAT10 level in human gastric cancer and its relation with mutant p53 level, lymph node metastasis and TNM staging. *World journal of gastroenterology : WJG* 2009, *15*, 2228-2233.

[13] Liu, N. Q., Stingl, C., Look, M. P., Smid, M. *, et al.*, Comparative proteome analysis revealing an 11-protein signature for aggressive triple-negative breast cancer. *Journal of the National Cancer Institute* 2014, *106*, djt376.

[14] Zanaruddin, S. N., Saleh, A., Yang, Y. H., Hamid, S. *, et al.*, Four-protein signature accurately predicts lymph node metastasis and survival in oral squamous cell carcinoma. *Human pathology* 2013, *44*, 417-426.

[15] Pavlou, M. P., Dimitromanolakis, A., Martinez-Morillo, E., Smid, M. *, et al.*, Integrating meta-analysis of microarray data and targeted proteomics for biomarker identification: application in breast cancer. *Journal of proteome research* 2014, *13*, 2897-2909.

[16] Cao, H. H., Zhang, S. Y., Shen, J. H., Wu, Z. Y. *, et al.*, A three-protein signature and clinical outcome in esophageal squamous cell carcinoma. *Oncotarget* 2014, *6*, 5435-5448.

[17] Shipitsin, M., Small, C., Giladi, E., Siddiqui, S. *, et al.*, Automated quantitative multiplex immunofluorescence in situ imaging identifies phospho-S6 and phospho-PRAS40 as predictive protein biomarkers for prostate cancer lethality. *Proteome science* 2014, *12*, 40.

[18] Mascini, N. E., Eijkel, G. B., Ter Brugge, P., Jonkers, J. *, et al.*, The use of mass spectrometry imaging to predict treatment response of patient-derived xenograft models of triple-negative breast cancer. *Journal of proteome research* 2015, *14*, 1069-1075.

[19] Cole, L. M., Bluff, J. E., Carolan, V. A., Paley, M. N.*, et al.*, MALDI-MSI and label-free LC-ESI-MS/MS shotgun proteomics to investigate protein induction in a murine fibrosarcoma model following treatment with a vascular disrupting agent. *PROTEOMICS* 2014, *14*, 890-903.
[20] Diehl, H. C., Beine, B., Elm, J., Trede, D.*, et al.*, The challenge of on-tissue digestion for MALDI MSI-a comparison of different protocols to improve imaging experiments. *Analytical and bioanalytical chemistry* 2015, *407*, 2223-2243.
[21] Mittal, P., Klingler-Hoffmann, M., Arentz, G., Zhang, C.*, et al.*, Proteomics of endometrial cancer diagnosis, treatment and prognosis. *Proteomics. Clinical applications* 2016, *10*, 217-229.
[22] Lyron Winderbaum, I. K., Parul Mittal, Peter Hoffmann, Classification of MALDI-MS imaging data of tissue microarrays using Canonical Correlation Analysis based variable selection. *Proteomics* 2016, *Epub ahead of print*.
[23] Meding, S., Martin, K., Gustafsson, O. J. R., Eddes, J. S.*, et al.*, Tryptic Peptide Reference Data Sets for MALDI Imaging Mass Spectrometry on Formalin-fixed Ovarian Cancer Tissues. *Journal of proteome research* 2012, *12*, 308-315.
[24] Gustafsson, O. J., Eddes, J. S., Meding, S., McColl, S. R.*, et al.*, Matrix-assisted laser desorption/ionization imaging protocol for in situ characterization of tryptic peptide identity and distribution in formalin-fixed tissue. *Rapid communications in mass spectrometry : RCM* 2013, *27*, 655-670.
[25] Gustafsson, J. O. R., Eddes, J. S., Meding, S., Koudelka, T.*, et al.*, Internal calibrants allow high accuracy peptide matching between MALDI imaging MS and LC-MS/MS. *Journal of Proteomics* 2012, *75*, 5093-5105.
[26] Alexandrov, T., Kobarg, J. H., Efficient spatial segmentation of large imaging mass spectrometry datasets with spatially aware clustering. *Bioinformatics* 2011, *27*, i230-i238.
[27] Egertson, J. D., MacLean, B., Johnson, R., Xuan, Y., MacCoss, M. J., Multiplexed peptide analysis using data-independent acquisition and Skyline. *Nat. Protocols* 2015, *10*, 887-903.
[28] Ween, M. P., Lokman, N. A., Hoffmann, P., Rodgers, R. J.*, et al.*, Transforming growth factor-beta-induced protein secreted by peritoneal cells increases the metastatic potential of ovarian cancer cells. *International Journal of Cancer* 2011, *128*, 1570-1584.
[29] Varghese, F., Bukhari, A. B., Malhotra, R., De, A., IHC Profiler: an open source plugin for the quantitative evaluation and automated scoring of immunohistochemistry images of human tissue samples. *PloS one* 2014, *9*, e96801.
[30] Sidoli, S., Simithy, J., Karch, K. R., Kulej, K., Garcia, B. A., Low Resolution Data-Independent Acquisition in an LTQ-Orbitrap Allows for Simplified and Fully Untargeted Analysis of Histone Modifications. *Analytical chemistry* 2015, *87*, 11448-11454.
[31] Liotta, L. A., Stracke, M. L., Aznavoorian, S. A., Beckner, M. E., Schiffmann, E., Tumor cell motility. *Seminars in cancer biology* 1991, *2*, 111-114.
[32] Stracke, M. L., Aznavoorian, S. A., Beckner, M. E., Liotta, L. A., Schiffmann, E., Cell motility, a principal requirement for metastasis. *Exs* 1991, *59*, 147-162.
[33] Rikardsen, O. G., Magnussen, S. N., Svineng, G., Hadler-Olsen, E.*, et al.*, Plectin as a prognostic marker in non-metastatic oral squamous cell carcinoma. *BMC Oral Health* 2015, *15*, 1-8.
[34] Katada, K., Tomonaga, T., Satoh, M., Matsushita, K.*, et al.*, Plectin promotes migration and invasion of cancer cells and is a novel prognostic marker for head and neck squamous cell carcinoma. *J Proteomics* 2012, *75*, 1803-1815.
[35] Bausch, D., Thomas, S., Mino-Kenudson, M., Fernández-del, C. C.*, et al.*, Plectin-1 as a Novel Biomarker for Pancreatic Cancer. *Clinical Cancer Research* 2011, *17*, 302-309.
[36] Bausch, D., Mino-Kenudson, M., Fernández-del Castillo, C., Warshaw, A.*, et al.*, Plectin-1 is a Biomarker of Malignant Pancreatic Intraductal Papillary Mucinous Neoplasms. *J Gastrointest Surg* 2009, *13*, 1948-1954.
[37] Bian, Y., Song, C., Cheng, K., Dong, M.*, et al.*, An enzyme assisted RP-RPLC approach for in-depth analysis of human liver phosphoproteome. *J Proteomics* 2014, *96*, 253-262.

[38] Beha, G., Sarli, G., Brunetti, B., Sassi, F.*, et al.*, Morphology of the Myoepithelial Cell: Immunohistochemical Characterization from Resting to Motile Phase. *The Scientific World Journal* 2012, *2012*, 8.
[39] Rønnov-Jessen, L., Petersen, O. W., A function for filamentous alpha-smooth muscle actin: retardation of motility in fibroblasts. *The Journal of Cell Biology* 1996, *134*, 67-80.
[40] Uzquiano, M. C., Prieto, V. G., Nash, J. W., Ivan, D. S.*, et al.*, Metastatic basal cell carcinoma exhibits reduced actin expression. *Mod Pathol* 2008, *21*, 540-543.

Figure 1

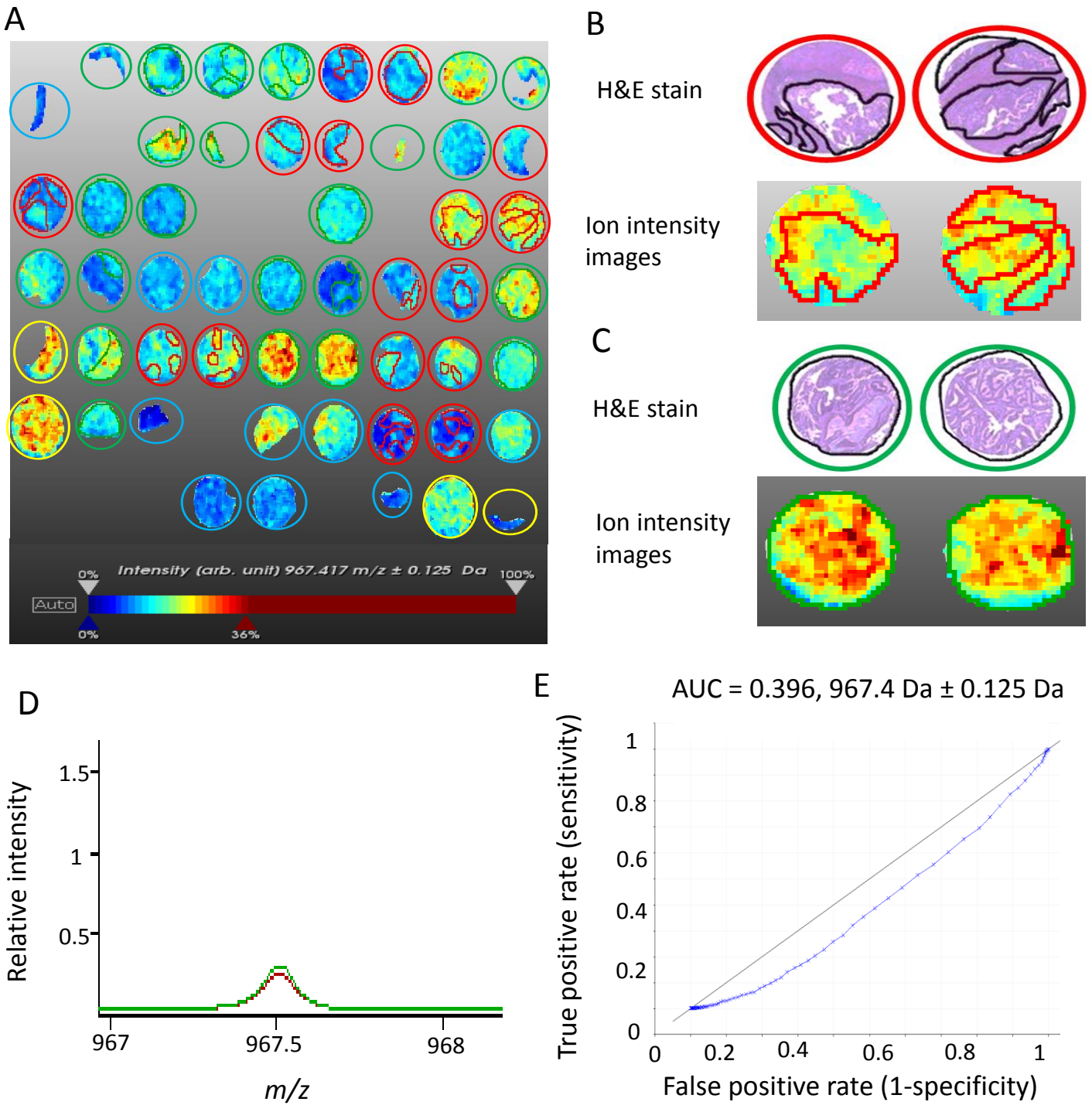

A
Intensity (arb. unit) 967.417 m/z ± 0.125 Da
0%
100%
Auto
0%
36%
B
H&E stain
Ion intensity images
C
H&E stain
Ion intensity images
D
Relative intensity
1.5
1
0.5
967
967.5
968
m/z
E
AUC = 0.396, 967.4 Da ± 0.125 Da
True positive rate (sensitivity)
1
0.8
0.6
0.4
0.2
0
0
0.2
0.4
0.6
0.8
1
False positive rate (1-specificity)

Figure 2

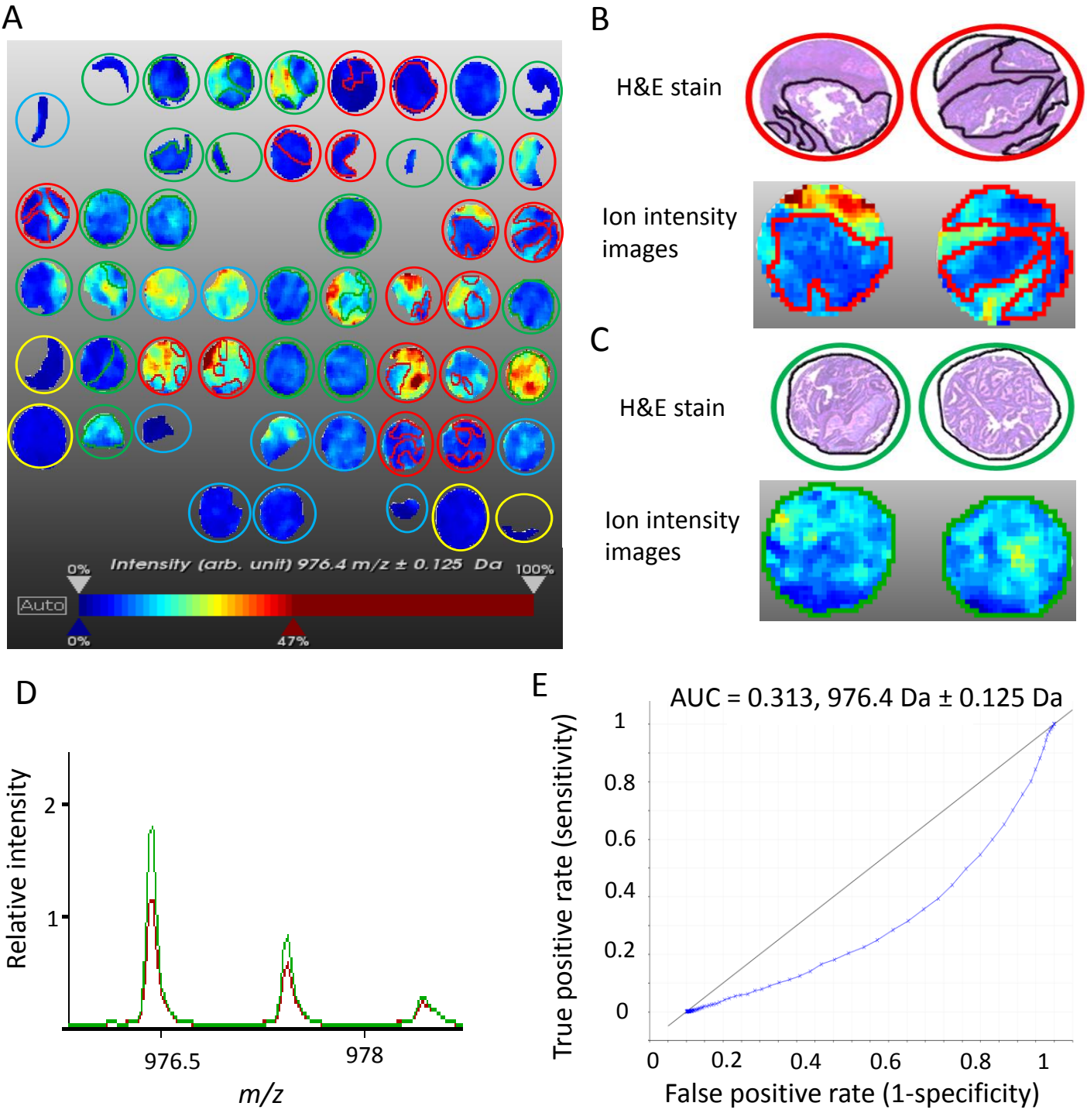

A
B
H&E stain
Ion intensity images
C
H&E stain
Ion intensity images
0%
Intensity (arb. unit) 976.4 m/z ± 0.125 Da
100%
Auto
0%
47%
D
Relative intensity
2
1
976.5
978
m/z
E
AUC = 0.313, 976.4 Da ± 0.125 Da
True positive rate (sensitivity)
1
0.8
0.6
0.4
0.2
0
0
0.2
0.4
0.6
0.8
1
False positive rate (1-specificity)

Figure 3

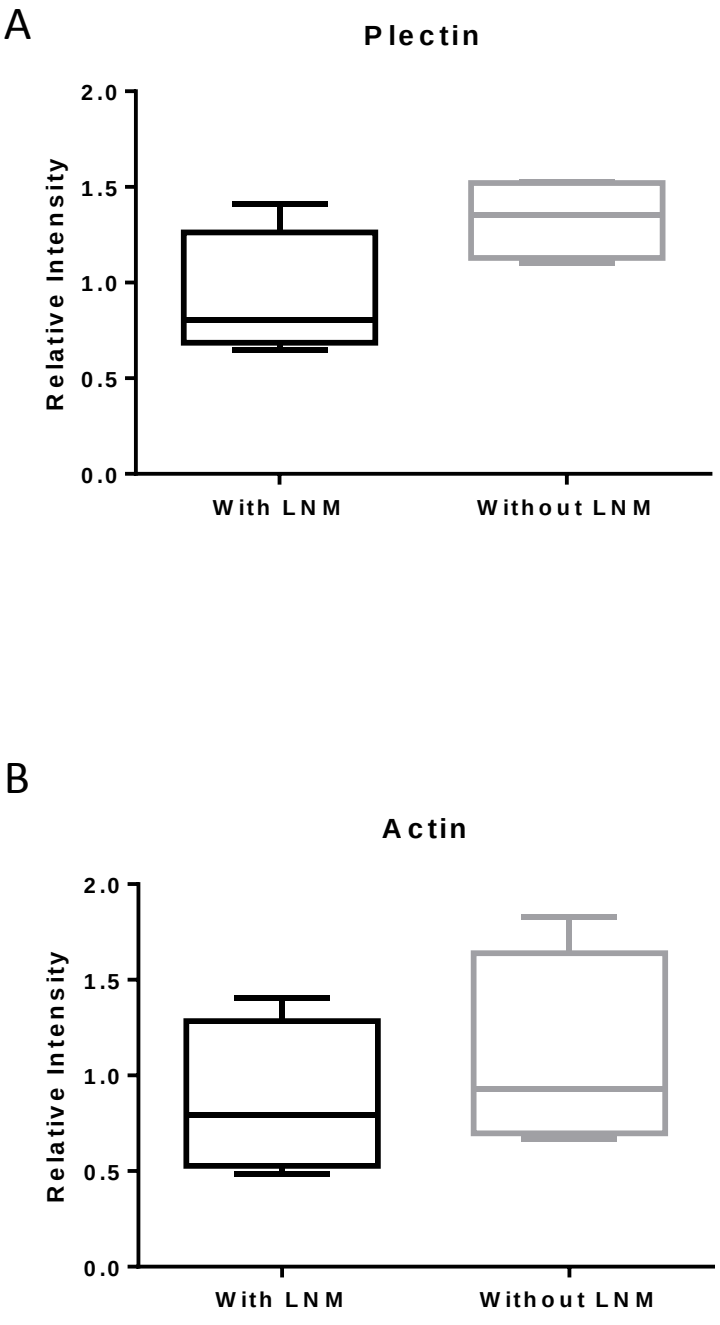
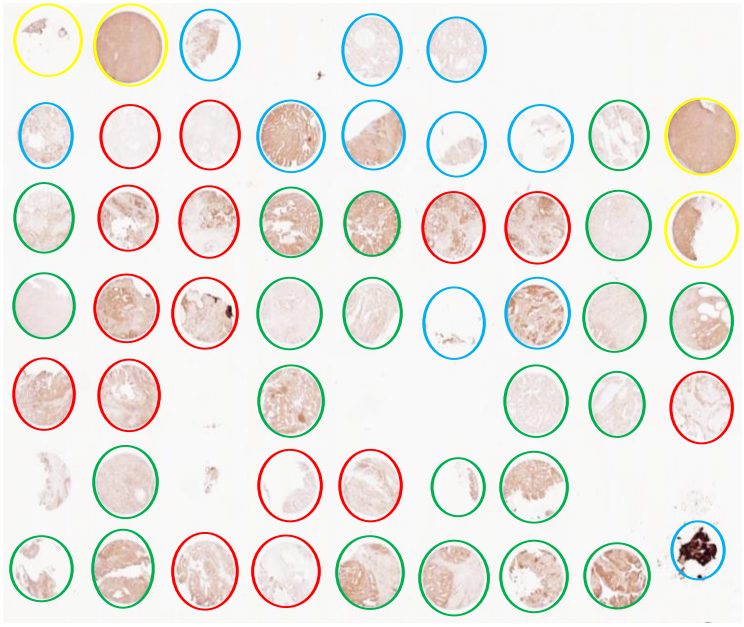
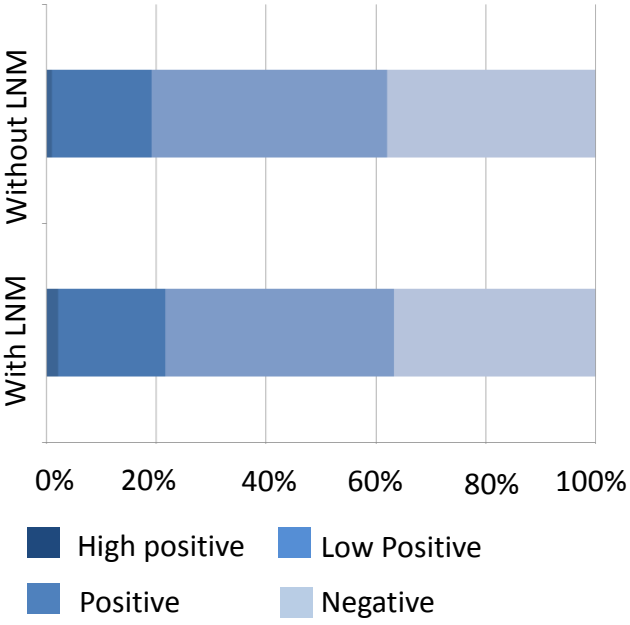

A
Plectin
Relative Intensity
2.0
1.5
1.0
0.5
0.0
With LNM
Without LNM
B
Actin
Relative Intensity
2.0
1.5
1.0
0.5
0.0
With LNM
Without LNM

Figure 4

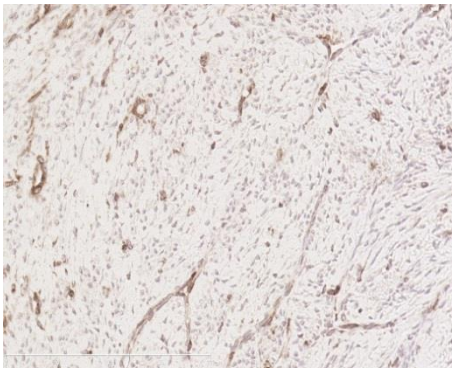
A



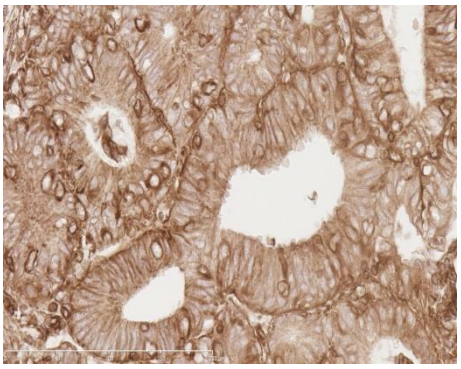
B



C Normal



D With LNM



E Without LNM

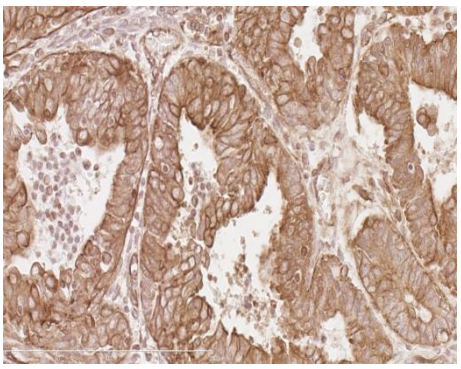
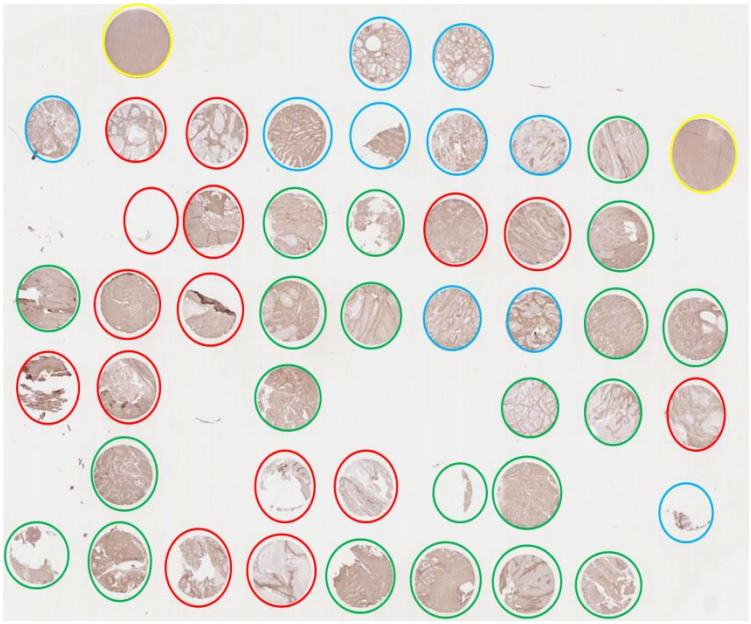
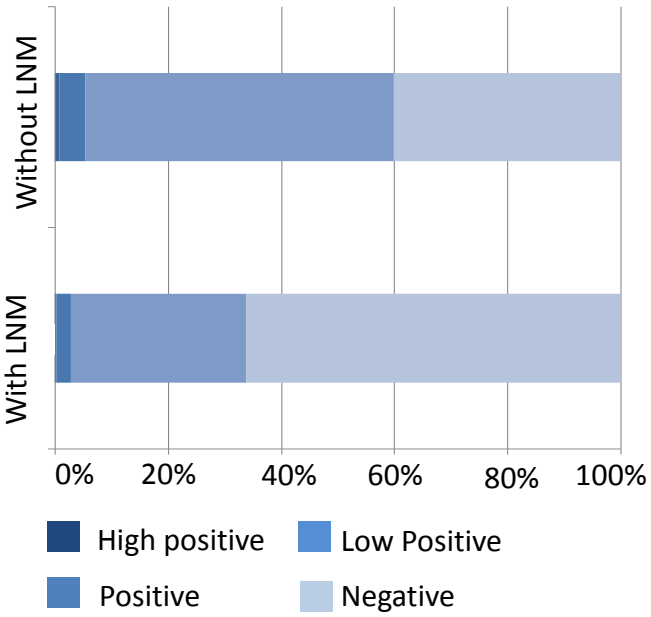


Figure 5

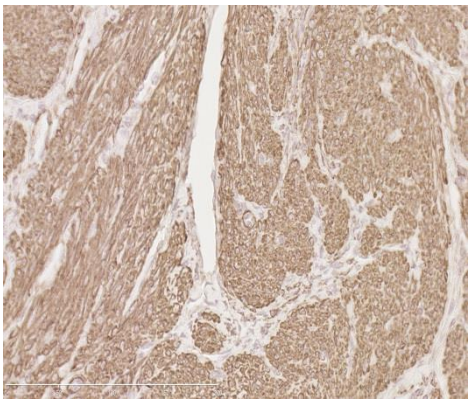
A



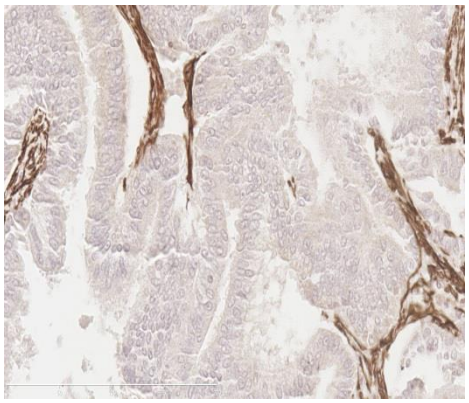
B



C Normal



D With LNM



E Without LNM