Proteomics characterization of HPV-independent Vulvar Cancer Tumour Progression and Metastasis by Mass Spectrometry



A thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

By

Chao ZHANG

B.Sc. (Biotechnology), M.Sc. (Biotechnology)

Adelaide Proteomics Centre

School of Molecular and Life Sciences

The University of Adelaide, South Australia

October 2018

This page intentionally left blank

Tables of Contents

Declaration XII

Acknowledgments XIII

List of Publications XV

Abbreviations XVI

Abstract XVIII

Chapter 1: Introduction 1

1.1 Vulvar Cancer 2

1.1.1 Vulvar cancer histopathology 3

1.1.2 Vulvar cancer incidence 6

1.1.3 Vulvar cancer mortality and survival 7

1.1.4 Vulvar cancer risk factor 8

1.1.5 Vulvar cancer stages and grades 9

1.1.6 Non-neoplastic vulvar skin disorders 12

1.1.7 Vulvar intraepithelial neoplasm 13

1.1.8 Vulvar cancer carcinogenesis 15

1.1.8.1 HPV-dependent vulvar cancer 18

1.1.8.2 HPV-independent vulvar cancer 19

1.1.9 Vulvar cancer diagnosis 21

1.1.10 Vulvar cancer prognosis and metastasis 23

1.1.11 Vulvar cancer metastasis prediction 26

1.1.12 Treatment of vulvar cancer 28

1.1.13 Indigenous vulvar cancer 29

1.1.14 Previous molecular biomarkers of vulvar cancer 33

1.2 Mass spectrometry 35

1.2.1 MALDI-TOF-MS and MS/MS 37

1.2.2 MALDI-Mass Spectrometry Imaging 39

1.2.3 *nano*-LC-ESI-MS/MS 42

1.2.4 Label-free quantification in *nano*-LC-ESI-MS/MS 44

1.2.5 Tissue samples for mass spectrometry analysis 48

1.2.5.1 Fresh-frozen tissue samples for MS study 49

1.2.5.2 The FFPE samples for MS study 50
Perspective 51
Hypothesis 52
Aims: 52
Chapter 2: Materials and Methods 53
2.1 Basic materials, reagents and instruments 54
2.1.1 Solvents and chemicals 54
2.1.2 Consumable materials 56
2.1.3 Proteolytic enzymes 57
2.1.4 Peptide calibration standard 57
2.1.5 Matrices 58
2.1.6 Primary, secondary antibodies and IHC reagents 58
2.1.7 Instruments 60
2.1.8 Software 61
2.2 Methods 63
2.2.1 Sample preparation methods 63
2.2.1.1 Tissue samples for experiments 63
2.2.1.2 Tissue microarray construction (TMA) for chapter 4 experiments 64
2.2.1.3 Experimental tissue sample preparation and heat-induced antigen retrieval (HIAR) 64
2.2.1.4 TMA sample reduction and alkylation for chapter 4 experiments 65
2.2.1.5 *in-situ* tryptic digestion 65
2.2.1.6 Internal calibrants application 66
2.2.1.7 Matrix deposition 66
2.2.1.8 Laser capture microdissection (LCM) 66
2.2.1.9 In-solution tryptic digestion, antigen retrieval and purification for the peptide library construction 67
2.2.1.10 In-solution tryptic digestion, antigen retrieval and purification for peptide label-free quantitation (LFQ) 68
2.2.2 Data acquisition methods 69
2.2.2.1 MALDI-MSI for chapter 3 and 5 experiments 69
2.2.2.2 MALDI-MSI for chapter 4 experiments 69
2.2.2.3 Raw spectra processing and peak picking 70
2.2.2.4 *in-situ* tandem mass spectrometry (MALDI-TOF/TOF-MS) 70
2.2.2.5 *nano*-LC-ESI-Orbitrap-MS/MS for the peptide library construction 71

2.2.2.6 *nano*-LC-ESI-qTOF-MS/MS for the peptide library construction and data dependent acquisition (DDA) based label-free quantification 71
2.2.2.7 Data independent acquisition (DIA) based label-free quantification 72
2.2.2.8 Hematoxylin and eosin (H&E) staining 73
2.2.2.9 Immunohistochemistry (IHC) staining 73
2.2.3 Data analysis methods 75
2.2.3.1 MALDI-MSI peaks grouping and statistical analysis 76
2.2.3.2 Canonical correlation analysis and Linear discriminant analysis (CCA-LDA) patient classification 77
2.2.3.3 MALDI-MSI spectra processing, statistical analysis and visualization in the SCiLS Lab 77
2.2.3.4 Peptide identification in the MASCOT 78
2.2.3.5 Peptide identification and LFQ data processing in the MaxQuant 78
2.2.3.6 DDA LFQ data analysis in the Perseus 79
2.2.3.7 DIA LFQ analysis in the Skyline MS 80
2.2.3.8 IHC staining analysis in the IHC Profiler 80
2.2.4 Analytical R Scripts 81
2.2.5 Reactome Pathway Analysis 83
Chapter 3: Proteomic Characterization of HPV-independent Vulvar Cancer Development 85
3.1 Introduction 86
3.2 Material and Methods 88
3.3 Results 91
3.3.1 Results of the manuscript 91
3.3.1.1 Matrix-assisted Laser Desorption/Ionization Mass Spectrometry Imaging (MALDI-MSI) 92
3.3.1.2 Cytokeratin 5 (CK5) Identified as a Protein of Interest 94
3.3.1.3 Immunohistochemistry (IHC) Analysis of CK5 comparing Healthy Epithelium, Differentiated Vulvar Intraepithelial Neoplasia, and Vulvar Squamous Cell Carcinoma 97
3.3.2 Proteomic characterization of metastatic vulvar cancer tissues 100
3.3.2.1 MALDI-MSI 100
3.3.2.2 Identification of discriminative *m/z* values 103
3.3.2.3 Immunohistochemistry analysis of ACTB 107
3.3.2.4 Verification of ACTB by data-independent acquisition *nano*-LC-ESI-MS/MS 110
3.4 Discussion 112
3.5 Conclusion 115
Chapter 4: Identification of predictive protein markers of vulvar cancer lymph node metastasis 117

4.1 Introduction 118
4.2 Materials and Methods 121
4.3 Results 124
4.3.1 Patient cohort 124
4.3.2 MALDI-MSI 127
4.3.3 CCA-LDA classification 130
4.3.4 Identification of discriminative *m/z* values 133
4.3.5 Immunohistochemistry (IHC) validation of candidate proteins identified in MALDI-MSI 150
4.3.6 Label-Free Quantification of Data-dependent data acquisition 164
4.4 Discussion 169
4.5 Conclusion 173
Chapter 5: Proteomics analysis of Indigenous vulvar cancer using *nano*-LC-ESI-MS/MS 175
5.1 Introduction 176
5.2 Materials and Methods 179
5.3 Results 181
5.3.1 Patient Samples 181
5.3.2 Relative quantification of HPV-independent Indigenous and non-Indigenous VSCC tissues 184
5.3.3 LFQ comparison of Indigenous and non-Indigenous vulvar cancer patients 186
5.3.4 Biological implications of Indigenous vulvar cancer proteomic variations 191
5.3.5 Validation of spatial distribution of proteins of interests in MALDI-MSI 195
5.3.6 Immunohistochemistry (IHC) validation of proteins of interest 205
5.4 Discussion 212
5.5 Conclusion 218
Chapter 6: Discussion and Future directions 220
6.1 Main findings 221
6.2 Application of MALDI-MSI on vulvar cancer research 223
6.3 Cytoskeleton remodeling as an important biological process in vulvar cancer progression 224
6.4 Tumor hypoxia proteins in vulvar cancer metastasis 228
6.5 Neutrophil degranulation promotes Indigenous vulvar cancer progression 230
6.6 Clinical considerations and future directions 231
REFERENCE 236
Supplementary ..... A
Appendix C

List of Figures and Tables

Figure 1.1: Representative histological sections of the benign vulvar epithelium.
Figure 1.2: Vulvar intraepithelial neoplasia of usual type.
Figure 1.3: Models of two distinctive etiological pathways of vulvar cancer carcinogenesis.
Figure 1.4: Model of HPV-independent vulvar cancer tissue types presenting in different cancer development stages.
Figure 1.5: Workflow of peptide MALDI mass spectrometry imaging (MALDI-MSI) of vulvar cancer tissue sections.
Figure 1.6: Workflow of vulvar cancer peptide sample analysis via label-free quantification approach in *nano*-LC-ESI-MS/MS.

Figure 2.1: The flow chat of analytical software used in this vulvar cancer proteomics characterization study.

Figure 3.1: A MALDI-MSI annotated ion intensity map from a representative vulvar tissue section for the peak group *m/z* 1410.67.
Figure 3.2: IHC analysis of CK5 on vulvar tissues containing regions of healthy epithelium, dVIN, and VSCC.
Figure 3.3: A comparison of the expression pattern observed for the CK5 peptides.
Figure 3.4: Hematoxylin and eosin (H&E) staining of tissue sections from four patients.
Figure 3.5: Representative healthy (green), dVIN (blue), VSCC (red) and LNM (grey) mass spectrums obtained by MALDI-MSI detection from a cohort of single patient tissue samples.
Figure 3.6: Representative MALDI-MSI images for ACTB peptide AVFPSIVGRPR which matched the peptide 1198.687Da.
Figure 3.7: IHC analysis of ACTB on vulvar tissue samples containing regions of healthy epithelium, dVIN, VSCC and LNM.
Figure 3.8: Comparison of expression pattern observed for the ACTB peptide.
Figure 3.9: Quantification of ACTB by data-independent *nano*-LC-ESI-MS/MS.

Figure 4.1: Spatial distribution of annotated VSCC tumour cores in two TMAs.
Figure 4.2: Receiver operating characteristic (ROC) curves for prediction of LNM condition by two *m/z* bins.
Figure 4.3: Representative MALDI-MSI images for *m/z* 1335.721Da.
Figure 4.4: Representative MALDI-MSI images for peptide GEQGAPGPR and GEQGAPGPR which matched the peptide bin of 868.583Da, AWM at 868.444.
Figure 4.5: Representative MALDI-MSI images for peptide SVSSSSYRR *m/z* 1028.555Da.
Figure 4.6: Representative MALDI-MSI images for keratin 17 peptide TIVEEVQDGK *m/z* of AWM 1117.584Da.
Figure 4.7: Representative MALDI-MSI images for keratin 1 peptide QLDSIVGER *m/z* of AWM 999.481Da.
Figure 4.8: Representative MALDI-MSI images for Ig gamma-3 chain C region peptide SCDTPPPCPR *m/z* of AWM 1186.552Da.
Figure 4.9: Peakgroup AWM 1391.741 from bin 1391.833 was identified as of LALDVEIATYRK of cytokeratin 6 with overall AUC= 0.481.
Figure 4.10: Representative MALDI-MSI images for NDRG1 peptide SHTSEGTR *m/z* 874.470Da.
Figure 4.11: IHC analysis of K2C5/K2C6 on TMA sections containing samples of LNM positive and negative VSCC samples.
Figure 4.12: IHC analysis of K2C1/K1C10 on TMA sections containing samples of LNM positive and negative VSCC samples.
Figure 4.13: IHC analysis of VIME on TMA sections containing samples of LNM positive and negative VSCC samples.
Figure 4.14: IHC analysis of COL-1 on TMA sections containing samples of LNM positive and negative VSCC samples.
Figure 4.15: IHC analysis of NDRG1 on TMA sections containing samples of LNM positive and negative VSCC samples.
Figure 4.16: IHC analysis of ANXA8 on TMA sections containing samples of LNM positive and negative VSCC samples.
Figure 4.17: Morphological patterns revealed in H&E staining, IHC staining and MALDI-MSI.
Figure 4.18: Total ion count chromatograms of *nano*-LC-MS/MS DDA analysis between six LNM positive and six LNM negative vulvar cancer patients.
Figure 4.19: Volcano plot of student t-test results of protein expression difference in LFQ analysis between six LNM positive and six LNM negative vulvar cancer patients.

Figure 4.20: Clustering heat map of LFQ candidate protein intensities among six LNM positive patients and six LNM negative patients.

Figure 5.1: Peptide abundance distribution in four Indigenous patient and four non-Indigenous patient VSCC tissue samples.

Figure 5.2: Volcano plot of student t-test results of protein expression differences in LFQ analysis between Indigenous (n=4) and non-Indigenous (n=4) vulvar cancer patients.

Figure 5.3: Clustering heatmap of LFQ candidate protein intensities for all Indigenous patients and all non-Indigenous patients.

Figure 5.4: The log2 transformed LFQ intensity analysis revealed the relative abundance difference of POSTN in VSCC sections of Indigenous (n=4) and non-Indigenous (n= 4) vulvar cancer patients.

Figure 5.5: Volcano plot of student t-test results of protein expression difference in LFQ analysis between two LNM positive and two LNM negative Indigenous vulvar cancer patients.

Figure 5.6: Clustering heatmap of LFQ candidate protein intensities for two LNM positive Indigenous and two LNM negative vulvar cancer patients.

Figure 5.7: The innate immune system signalling pathways that involved with proteins of interests showed differential abundance between Indigenous and non-Indigenous cancers in the Reactome Pathway Analysis.

Figure 5.8: The metabolism signalling pathways that involved with proteins of interests showed differential abundance between Indigenous and non-Indigenous cancers in the Reactome Pathway Analysis.

Figure 5.9: MALDI-MSI annotated ion intensity map from a representative Indigenous (n=4) and non-Indigenous (n=4) patient VSCC tissue section for POSTN peptide *m*/*z* 928.48.

Figure 5.10: MALDI-MSI annotated ion intensity map from a representative Indigenous (n=4) and non-Indigenous (n=4) patient VSCC tissue section for FGA peptide *m*/*z* 848.52.

Figure 5.11: MALDI-MSI annotated ion intensity map from a representative Indigenous (n=4) and non-Indigenous (n=4) patient VSCC tissue section for FN1 peptide *m*/*z* 872.44.

Figure 5.12: MALDI-MSI annotated ion intensity map from a representative Indigenous (n=4) and non-Indigenous (n=4) patient VSCC tissue section for FN1 peptide *m*/*z* 980.47.

Figure 5.13: MALDI-MSI annotated ion intensity map from a representative Indigenous vulvar tissue section for NDRG1 peptide *m*/*z* 2376.11.

Figure 5.14: MALDI-MSI annotated ion intensity map from a representative Indigenous vulvar tissue section for RAP1A peptide *m*/*z* 985.60.

Figure 5.15: IHC analysis of POSTN on VSCC tissue sections containing samples of Indigenous (n=4) and non-Indigenous samples (n=4), at 40× magnification.

Figure 5.16: IHC analysis of NDRG1 for Indigenous VSCC tissue sections containing samples of LNM positive and negative VSCC samples, at 40× magnification.

Figure 5.17: IHC analysis of RAP1A on Indigenous VSCC tissue sections containing samples of LNM positive and negative VSCC samples, at 40× magnification.

Figure 6.1: Interactions among proteins found to be potentially involved in vulvar cancer progression.

Figure 6.2: The study design of future vulvar cancer biomarker validation and expected patient numbers in each group of the validation cohort.

Table 1.1: The American Joint Committee on Cancer (AJCC) TNM cancer staging classification and the International Federation of Gynecology and Obstetrics (FIGO) cancer staging system for vulvar cancer.

Table 1.2: Comparing different analysis approaches in the current vulvar cancer study.

Table 3.1: The vulvar cancer patient sample cohort (n=12) of the published manuscript.

Table 3.2: Matrix-assisted laser desorption/ionization mass spectrometry imaging peak groups ranked heuristically

Table 3.3: CK5 matrix-assisted laser desorption/ionization mass spectrometry imaging peak groups of interest with matched nanoflow liquid chromatography-tandem mass spectrometry results.

Table 3.4: Top 10 MALDI-MSI peak groups ranked heuristically by KW test results among healthy, dVIN, VSCC and LNM tissue samples.

Table 3.5: Abundant peaks groups identifiable by *in-situ* MALDI MS/MS found to have significantly different signal to noise ratios between the four tissue types (p<0.01).

Table 4.1: Demographic and clinical information of study cohort patients with LNM positive and LNM negative vulvar squamous cell carcinoma.

Table 4.2: MALDI-MSI peakgroups ranked heuristically by the S/N or log intensity discriminative power between VSCC with positive LNM and negative LNM.

Table 4.3: Top-ranked *m/z* bins from selective CCA-LDA classification list of VSCC primary tumour TMAs.
Table 4.4: Identification of discriminative *m/z* values from MALDI-MSI results by matching a peptide library created from sample LC-MS/MS and *in-situ* MS/MS proteomic profiling.
Table 4.5: Summary of significantly different (p<0.01) proteins in t-test results of LFQ analysis between six LNM positive and six LNM negative patients.

Table 5.1: Clinical information of Indigenous and non-Indigenous vulvar squamous cell carcinoma and VIN patients in the cohort.
Table 5.2: Identification of discriminative *m/z* values from MALDI-MSI results by matching a peptide library created from sample *nano*-LC-ESI-MS/MS proteomic profiling and label-free quantification.

DECLARATION

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint award of this degree.

I give consent to this copy of my thesis when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968. I acknowledge that copyright of published works contained within this thesis resides with the copyright holder(s) of those works.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

Chao ZHANG

June 2018

ACKNOWLEDGMENTS

Studying my Ph.D. at the Adelaide Proteomics Centre is an important life experience but also a great challenge for international students like me. Luckily, I received countless help from those kindness people from our lab and external organizations. Their valuable suggestions were not only assisted me to overcome all technical difficulties in my experiments but also inspired my future scientific life. All my works achieved in my Ph.D. period are closely related with their generous support.

Important of all, I would like to give my first sincere appreciation to my supervisor, leader of the lab, Prof. Peter Hoffmann, who granted this vital opportunity to me to start my academic life and Ph.D. study in his lab. His advice is the key for me to accomplish my postgraduate research. Here, I also want to express my gratitude to my co-supervisor, Prof. Martin K Oehler, who provided me with all patient samples and histopathological knowledge in this project. Besides, I am extremely grateful to postdoctoral fellows in the Adelaide Proteomics Centre who guided my daily experiments and academic writings step by step in the past four years. Dr. Georgia Arentz, who always being patient and seeking possibilities to solve my research and writing problems, while Dr. Yinying Ho, who are willing to share her knowledge and expertise in the last year of my study. Their hard working guaranteed the success of my research projects and my following research life will benefit a lot from their experience.

I would like to thank the co-authors of my article, Dr. Lyron Winderbaum, for his critical mathematics scripts and comments that subsequently improved the quality of my paper and thesis. I would like to thank Dr.Gurjeet Kaur and Dr. Christopher Carter, for their assistance annotating the tumour tissue sections and provide professional clinical analysis. Our lab technician, Chris Cursaro, deserves my sincere acknowledgment because of his kindly experimental support. I would also like to thank Dr. Manuela Klingler-Hoffmann, Dr. Noor Lokman, Parul Mittal, Matthew Briggs, Mitchell Acland and other members of the Adelaide Proteomic Centre, without their help, my Ph.D. degree would never have been possible.

Finally, I should remember warmest supports given by my family, my parents and brother who do not only support all my tuition fee but also my life in Australia. They encourage me to finish my scientific dream and never give up hope even in my hardest moment. Thank you for all the listening to my complaints and most importantly allowing me to chase my own dream.

Thank you again for all your help.

Chao ZHANG

LIST OF PUBLICATIONS

This thesis is based on the following publications, which are referred to in the text and attached as appendices:

I. **Chao Zhang**, Georgia Arentz, Lyron Winderbaum, Noor A Lokman, Manuela Klingler-Hoffmann, Parul Mittal, Christopher Carter, Martin K. Oehler, Peter Hoffmann, MALDI Mass Spectrometry Imaging reveals decreased CK5. Levels in vulvar squamous cell carcinomas compared to the precursor lesion differentiated vulvar intraepithelial neoplasia, IJMS, 2016, 17; 1088-1099.

II. Parul Mittal, Manuela Klingler-Hoffmann, Georgia Arentz, **Chao Zhang**, Gurjeet Kaur, Martin K. Oehler, Peter Hoffmann, Proteomics of endometrial cancer diagnosis, treatment, and prognosis, Proteomics Clinical Applications, 2016; 217-229.

III. Georgia Arentz, Parul Mittal, **Chao Zhang**, Ying-Ying Ho, Matthew Briggs, Lyron Winderbaum, Manuela Klingler-Hoffmann, Peter Hoffmann, Chapter 2 - Application of Mass Spectrometry Imaging to cancer, in: Richard, R. D., Liam, A. M. (Eds.), Advances in Cancer Research, Academic Press 2017, Volume 134, 1st Edition, 27-66.

IV. Parul Mittal, Manuela Klingler-Hoffmann, Georgia Arentz, Lyron Winderbaum, Noor Alia Lokman, **Chao Zhang**, Lyndal Anderson, James Scurry, Yee Leung, Colin JR Stewart, Jonathan Carter, Gurjeet Kaur, Martin K. Oehler, Peter Hoffmann, Lymph node metastasis of primary endometrial cancers: Associated proteins revealed by MALDI imaging, Proteomics, 2016, 16, 1793-1801.

ABBREVIATIONS

ACTB	β-actin
ACTN4	Alpha-actinin-4
ANXA1	Annexin A1
ANXA2	Annexin A2
ANXA8	Annexin A8
CHCA	α-cyano-4-hydroxycinnamic acid
CCA	Canonical Correlation Analysis
COL1	Collagen Type 1
CO1A1	Collagen alpha-1(I) chain
CO1A2	Collagen alpha-1(II) chain
Da	Dalton
DDA	Data-dependent analysis
DIA	Data-independent analysis
DHB	2,5-Dihydroxybenzoic acid
dVIN	Differentiated type, vulvar intraepithelial neoplasia
ECM	Extracellular matrix
ESI	Electro-spray ionization
EMT	Epithelial-mesenchymal transition
FFPE	Formalin-fixed paraffin-embedded
FIGO	International Federation of Gynecology and Obstetrics
H&E	Haematoxylin and eosin
HIAR	Heat-induced citric acid antigen retrieval
IHC	Immunohistochemistry
ITO	Indium tin oxide
K1C10 or CK10	Keratin, type I cytoskeletal 10

K1C17 or CK17	Keratin, type I cytoskeletal 17
K2C1 or CK1	Keratin, type II cytoskeletal 1
K2C5 or CK5	Keratin, type II cytoskeletal 5
K2C6 or CK6	Keratin, type II cytoskeletal 6
LC	Liquid chromatography
LCM	Laser capture micro-dissection
LC-MS/MS	Liquid chromatography-tandem mass spectrometry
LDA	Linear discriminant analysis
LFQ	Label-free quantification
LNM	Lymph Node Metastasis
MALDI	Matrix assisted laser desorption/ionization
MALDI-MSI	Matrix-assisted laser desorption/ionization mass spectrometry imaging
MS	Mass spectrometry
m/z	Mass to charge ratio
S/N	Signal to Noise ratio
PEN	Polyethylene naphthalate
POSTN	Periostin
ppm	Parts per million
qTOF	Quadrupole time of flight
TMA	Tissue microarray
VIN	Vulvar intraepithelial neoplasia
VIME	Vimentin
VSCC	Vulvar squamous cell carcinoma
uVIN	Usual type, vulvar intraepithelial neoplasia

ABSTRACT

Vulvar cancer is the fourth most common gynecological cancer around the world, constituting 5% of all gynecological cancers and 0.6% of total cancer cases in 2016. As a relatively rare cancer type, the average vulvar cancer incidence rate is three in 100 000 women each year. But, Australian Indigenous showed significantly higher incidence rate than the global level. Previous etiological studies have identified two distinctive oncogenetic origins of vulvar cancer based on the presence of HPV infection or not. In general, over 60% of patients, especially those diagnosed at the post-menopause age or with a history of precancerous lesions of differentiated VIN (dVIN), are classified as the high metastatic risk HPV-independent type vulvar cancer. Currently, most scientific efforts of vulvar cancer research are given to the low-risk HPV-dependent type, whereas protein studies focused on the HPV-independent type are quite limited.

The protein biomarker is an important diagnostic tool in cancer early detections and metastasis predictions. However, there is no protein biomarker targeting on HPV-independent vulvar cancer available in the clinical practice yet. Tumour biomarkers, such as TP53 and CA125, do not have enough sensitivity and specificity in vulvar cancer detection. Therefore, new protein biomarker candidates are needed for current vulvar cancer diagnosis and prognosis. Mass spectrometry (MS) is a high-throughput protein analysis and biomarker discovery technique in modern cancer studies. Recent developments of cut-edge MALDI mass spectrometry imaging (MALDI-MSI) and *nano*-LC-ESI-MS label-free quantification (LFQ) provide new opportunities to reveal differential abundances and spatial distributions of proteins from vulvar cancer tissue samples.

In this thesis, protein markers for HPV-independent vulvar cancer carcinogenesis and metastasis were identified from the MS based tissue characterization and validated by Immunohistochemistry (IHC). Firstly, the intermedia filament protein, Cytokeratin 5 (CK5), was identified as a tissue-specific biomarker for HPV-independent vulvar cancer in the MALDI-MSI analysis of 12 representative patient samples, which has been published in the manuscripts. Other potential tissue markers, β-Actin (ACTB) as an example, were also detected with differential abundances among benign epithelium, differentiated vulvar intraepithelial neoplasia (dVIN), vulvar squamous cell

carcinoma (VSCC) and metastatic lymph nodes (LN). These proteins could be used as tissue markers to distinguish VSCC from dVIN and other vulvar tissues.

The following results of this thesis focused on proteins with potential discriminative abilities between lymph node metastasis (LNM) positive and LNM negative HPV-independent vulvar cancer patients. Two vulvar cancer VSCC tissue microarrays (TMAs) built from 28 patient samples (16 LNM positive and 12 LNM negative) were analyzed using the MALDI-MSI data. Peptide peaks of N-myc downstream regulated 1 (NDRG1) and Annexin 8 (ANXA8) showed higher discriminative powers in the Receiver Operating Characteristic Curve analysis with Area Under Curve scores of 0.641 and 0.625 respectively. Further Canonical Corresponding Analysis and Linear Discrimination Analysis (CCA-LDA) classification based on log-transformed peptide peaks achieved the best 28 out of 28 corrections in determining LNM status of TMA patients. Proteins with differential abundance were further validated by the LFQ and IHC staining, which could be used as molecular references complement to the histopathological classification system.

Lastly, four Indigenous HPV-independent vulvar cancer patients were studied in Chapter 5 to compare differential protein abundances between Indigenous and non-Indigenous patients in LFQ experiment. As the result, the Periostin (POSTN) demonstrated significantly lower abundance in Indigenous vulvar cancers as compared to other non-Indigenous cancers. Besides, NDRG1 and Ras-related protein 1A (RAP1A) were identified able to determine Indigenous VSCCs with and without LNM. Spatial distributions of those proteins of interest were revealed by the MALDI-MSI and IHC staining on consecutive tissue sections. In addition, the Reactome Pathway Analysis suggested neutrophil degranulation and metabolism alternations could be key biological processes involved with the prevalence of vulvar cancer in Indigenous communities.

Discovery proteins associated with vulvar cancer progression is important for vulvar cancer diagnosis and prognosis while understanding differential protein abundance in Indigenous patients will help to improve vulvar cancer treatments. This study proved MALDI-MSI and *nano*-LC-ESI-MS/MS LFQ are ideal methods for vulvar cancer protein identification and quantification. Protein identified in this study will be promising vulvar cancer tissue markers in the future practice.

This page intentionally left blank

CHAPTER 1: INTRODUCTION

1.1 VULVAR CANCER

Vulvar cancer is a rare gynecological cancer affecting the outermost part of the female reproductive system. This cancer generally initiates at the skin of the labia majora or labia minora, perineum or clitoris, or in unusual cases at the bottom of the urethra or vagina glands (Lancaster, 2010). Vulvar cancer is a slowly progressing disease and it takes several years for obvious symptoms to develop. In the early stage of vulvar cancer malignant tumour lesions on the vulva present as thickened skin patches or small moles with abnormal shape and colour, with the subsequent wart-like growth of vulvar cancer potentially causing itching, soreness, bleeding and other symptoms. However, some symptoms described above are very similar to the benign vulvar neoplasm which results in a high misdiagnosis rate of early-stage vulvar cancer (Canavan and Cohen, 2002).

1.1.1 Vulvar cancer histopathology

According to tumour origins, vulvar cancer can be classified into the following cancer subtypes. The dominant vulvar cancer type is vulvar squamous cell carcinoma (VSCC), accounting for over 90% of diagnosed vulvar cancer cases (Ansink, 1996). Normally, vulvar skin squamous cells are multilayered, flat cells which tightly overlap to form skin barriers. In contrast, VSCC tumour cells are malignant neoplastic cell growths, with deregulated shape, that invade deeper parts of the skin (LEIBOWTTCH et al., 1990). The histological difference between healthy vulvar epithelium and VSCC is shown in ***Figure 1.1***. In clinical practice, VSCC is classified into several subtypes. The common keratinizing subtype appears as exophytic or ulcerative like tumour lesions and this subtype is frequently detected in post-menopausal patients, independent from human papillomavirus (HPV) infection, and has much higher lymph node metastatic risk than other subtypes (Monk et al., 1995). VSCC patients with HPV infection are frequently classified as basaloid or warty subtypes. These two VSCC subtypes mainly affect young women (younger than 40) and display better prognostic outcomes following treatment (Parkin and Bray, 2006). The last subtype of VSCC is verrucous carcinoma which is extremely rare (less than 1% of VSCC cases) (Schwartz, 1995).

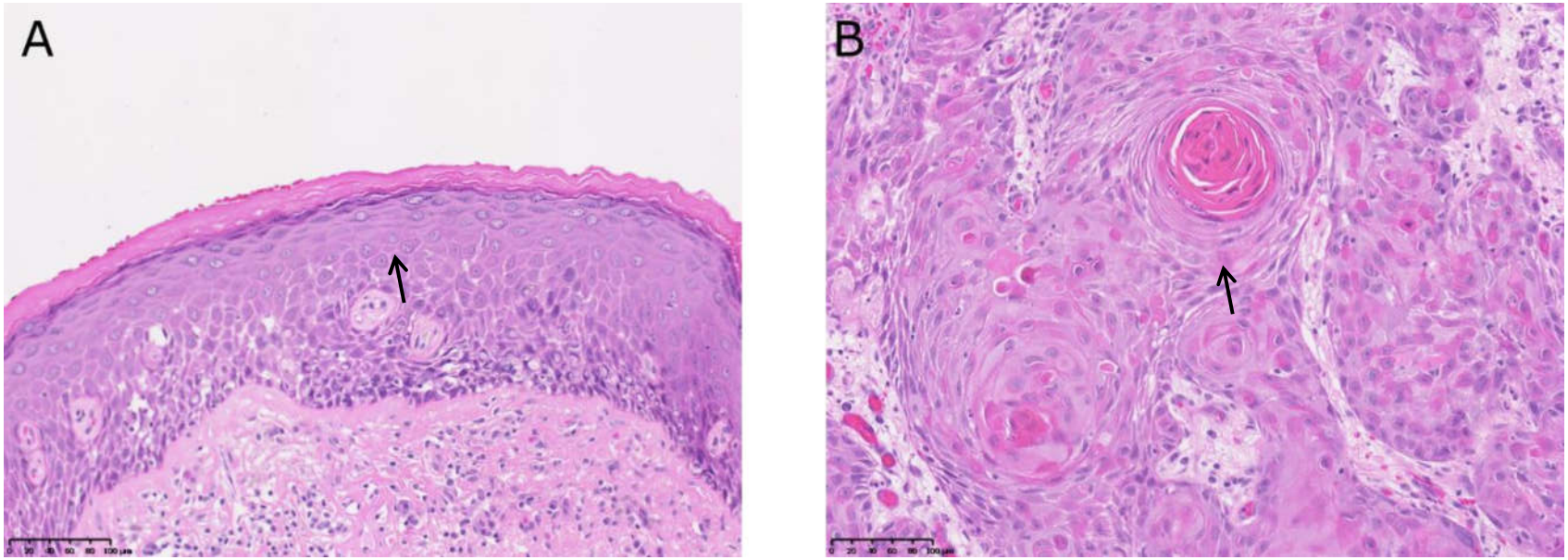


Figure 1.1: Representative histological sections of benign vulvar epithelium (A) and the vulvar cancer lesion (VSCC type) (B) in H & E stains. A: Benign skin tissue shows well-differentiated and multilayered regular shape squamous cells (black arrow). B: VSCC shows abnormal large and atypical keratinocytes with pearl form structure (black arrow) in a malignant tumour. Bar = 100µm.

Other than the most common VSCC type, vulvar cancer may initiate from Bartholin gland cells, classified as vulvar adenocarcinoma (Leuchter et al., 1982). This type of vulvar cancer accounts for 8% of total diagnosed cases and recent clinical investigation suggests that vulvar Paget disease may be a precursor lesion of vulvar adenocarcinoma, but the mechanism of vulvar cancer carcinogenesis remains unclear due to limited patient numbers (Shaco-Levy et al., 2010). The rest of diagnosed vulvar cancer cases include vulvar melanomas (about 6% of cases) and sarcoma and basal cell carcinoma (2% of cases) (Canavan and Cohen, 2002). This study mainly focuses on the dominant VSCC type, rather than rare vulvar cancer types.

1.1.2 Vulvar cancer incidence

As a relatively rare gynecological cancer, the average global vulvar cancer incidence rate is three in 100 000 women each year (Board, 2016). In the US, vulvar cancer was the 28th most common cancer type in 2016 and it is estimated that 5 950 women will be diagnosed with vulvar cancer in 2017, accounting for 0.4% of newly diagnosed cancer cases (Board, 2016). Over the past three decades, the Australian national vulvar cancer incidence rate remained stable in the range of two to five per 100 000 women per year, with most patients diagnosed with vulvar cancer at postmenstrual age (over 50) (Board, 2016). However, the vulvar cancer incidence rate among young women (younger than 50) has dramatically increased over the past few years to 2.93 per 100 000 in 2013, which is two times higher than incidence rate in 1990 (Rottmann et al., 2016). The increasing number of young vulvar cancer patients is shown to be related to HPV prevalence and other virulent infections which can be prevented by vaccination and healthy lifestyle (Kenter et al., 2009). Although vulvar cancer is a rare disease in Australia, it is more common among Indigenous women compared to women in the broader population. In 2005, the incidence rate for Indigenous women was two-fold more than for Caucasian women. But, in Indigenous communities of Arnhem Land, the Indigenous vulvar cancer incidence rate was as high as 31.1 per 100 000, ten times higher than the national level (Condon et al., 2005).

1.1.3 Vulvar cancer mortality and survival

Generally, the vulvar cancer mortality rate is very low compared to other gynecological cancer types, such as breast and ovarian cancer. In 2017, 1769 Australian women died from vulvar cancer, accounting for 8.6 % of all female cancer deaths (AIHW, 2016). Statistical analysis suggests that median death age of vulvar cancer patients is 78 years old, and 90% of young patients (younger than 55) fully recover after receiving appropriate treatment (Stroup et al., 2008). Between 2008 and 2012, 71.9% per cent of vulvar cancer patients survived the first 5 years postdiagnosis. The key factors in vulvar cancer death are lymph node metastasis (LNM), followed by local recurrence and surgery related infections (Van der Zee et al., 2008). The 5 year relative survival rate of vulvar cancer patients with carcinoma *in-situ* was over 80%, whereas only 57.1% per cent of LNM positive patients survive past the first 5 year period after treatment, including some patients with distant metastasis who suffer the worse prognostic outcomes and rarely recover even following radiotherapy treatment (Board, 2016). Unfortunately, nearly half of vulvar cancer patients have positive node status when clinically diagnosed. In summary, vulvar cancer is low-risk cancer in terms of death, resulting in only 0.6% of patients deaths each year (Hampl et al., 2008, Board, 2016). With the development of tissue specific markers and lymph node metastasis predictors, vulvar cancer patients will receive better treatments with less invasive surgery and lower local recurrences in the future practice.

1.1.4 Vulvar cancer risk factor

The potential risk of developing vulvar cancer, especially the HPV-independent vulvar cancer, is lower than other gynecological cancers and only 0.3% of women may be diagnosed with vulvar cancer in their lifetime (AIHW, 2016). Currently, there is limited evidence of a causative link between the development of vulvar cancer and specific hereditary genetic mutations. There are well-known risk factors that increase the chance of vulvar cancer development. The first risk factor is exposure to human papillomavirus (HPV) or another oncovirus infection. The HPV family contains more than 150 different virus genotypes and some genotypes are frequently identified in warty growth lesions of vulvar cancer. Not all HPV infections result in carcinogenesis, the only HPV genotypes 16, 18 and 31 have strong cancerous effects on vulva skin, while other genotypes have not been reported as potential cancer risks (Kenter et al., 2009). The high-risk HPV infection genotypes are mainly identified in young vulvar cancer patients. Therefore, positively avoiding direct skin contact with HPV-positive patients and HPV vaccination helps to minimize HPV-dependent vulvar cancer risk (Lee et al., 1994). Similar to HPV, human immunodeficiency virus (HIV) may trigger vulvar cancer development. The second risk factor is the previous history of vulvar diseases, which is frequently reported in HPV-independent vulvar cancer cases. Patients with chronic vulvar skin disorders, such as lichen sclerosis, are more vulnerable to vulvar cancer, while the vulvar intraepithelial neoplasia (VIN) is confirmed as the direct pre-cancerous lesion of vulvar cancer (Modesitt et al., 1998, Carlson et al., 2013). The third risk factor for vulvar cancer is the aging of women. Over 50% of HPV-independent vulvar cancer patients are diagnosed at the post-menopausal age of over 60, whereas HPV-dependent vulvar cancer is more common among females younger than 50 (Hampl et al., 2008). Other risk factors important in vulvar cancer are smoking, obesity and other unhealthy lifestyle factors which increase the risk of vulvar cancer in women (Daling et al., 1992).

1.1.5 Vulvar cancer stages and grades

The cancer staging system is the global standard to determine the progressing condition of vulvar cancer in clinical practice. Currently, the International Federation of Gynaecology and Obstetrics (FIGO) and American Joint Committee (TNM) cancer staging systems are applied to indicate the level of vulvar cancer progression (***Table 1.1***) (Pecorelli, 2009, Canavan and Cohen, 2002). The TNM system describes cancer based on histological characteristics, including tumour size (T), metastasis in regional lymph nodes (N) and distant lymph nodes (M), while FIGO guidelines classify vulvar cancer progression into five clinical stages (from 0 to IV) (Canavan and Cohen, 2002). Based on the FIGO cancer staging system (Pecorelli, 2009), *in-situ* carcinoma or intraepithelial carcinoma is stage 0 cancer, with tumours located on the vulva skin surface. At stage I, the size of *in-situ* tumour is less than 2 cm and the affected area is limited to vulva and perineum. Stage I can be further divided into IA lesions (<= 2cm diameter and <=1mm stromal invasion) and IB unilateral lesion (<=2cm diameter, >1mm stromal invasion and >=1cm from tumor centre) depending on tumor size and invasion depth. At stage II the larger size tumour is still localized to the vulva and perineum regions without metastasis spread to nearby lymph nodes. However, at stage III, the tumour has extended from the vulva and perineum to the lower urethra, vagina or anus with unilateral regional positive lymph node metastasis of either single positive lymph node metastasis of more than 5mm or multiple node metastasis under 5mm (Makar et al., 2001). Stage IV is considered the terminal stage of vulvar cancer, as cancer has widely spread to other organs, such as bladder mucosa, pelvic bone and bowel and the metastasis tumour may be found in distant pelvic lymph nodes.

Currently, both FIGO and TNM cancer stage systems are widely used for vulvar cancer classification according to morphological tissue patterns, rather than microscopic reporting of tumour appearance (Origoni et al., 1992). Tumour grades of vulvar cancer describing tumour cell appearance are employed as complementary microscopic detail for cancer pathological analysis, rather than to determine the cancer stage. Cancer stages and grades are frequently used as references for vulvar cancer classification. However, highly descriptive parameters of vulvar cancer staging and grading systems may be confused by pathologists in clinical practice,

particularly as some pre-cancerous vulvar lesions share similar morphological descriptions with stage 0 and stage I vulvar cancer. Moreover, stage III and IV vulvar cancer can be only confirmed by lymph node biopsy results, rather than primary tumour histopathology. Such deficiencies of current FIGO and TNM cancer staging systems indicate that a more sensitive and specific standard for vulvar cancer stage classification is required for application in the clinical setting (Pecorelli, 2009).

Table 1.1: The American Joint Committee on Cancer (AJCC) TNM cancer staging classification and the International Federation of Gynecology and Obstetrics (FIGO) cancer staging system for vulvar cancer.

Vulvar cancer	Description			
FIGO stage	Description	TNM stage		
		T	N	M
0	Not evaluated	TX	NX	M0
	No primary tumour detected	T0	N0	
I	*in-situ* Carcinoma	Tis		
IA	lesions ≤ 2 cm, vulva or perineum and with stromal invasion ≤ 1.0 mm	T1a		
IB	Lesions > 2 cm size or any size with stromal invasion > 1.0 mm	T1b		
II	Tumour of any size with extension to adjacent perineal structures (lower/distal 1/3 urethra, lower/distal 1/3 vagina, anal involvement)	T2		
IIIA	One or two positive regional lymph nodes		N1	
	One or two node metastases, < 5 mm		N1a	
	One lymph node metastasis ≥ 5 mm		N1b	
IIIB	Over two regional lymph node metastasis		N2	
	Three or more positive lymph node < 5 mm		N2a	
	Two or more positive lymph node ≥ 5 mm		N2b	
IIIC	Lymph node metastasis with extracapsular spread		N2c	
IVA	Tumour of any size with extension (upper/ proximal 2/3 urethra, upper/proximal 2/3 vagina, bladder mucosa, rectal mucosa or fixed to pelvic bone)	T3	N3	
	Fixed or ulcerated regional lymph node metastasis			
IVB	Distant metastasis (including pelvic lymph node metastasis)			M1

1.1.6 Non-neoplastic vulvar skin disorders

More than half of vulvar cancer cases develop from HPV-free chronic inflammation in aging. Recently, it has been shown that non-neoplastic vulvar skin disorders, including lichen sclerosis (LS), lichen planus (LP) and lichen simplex chronicus, are involved in vulvar squamous cell carcinoma carcinogenesis and recurrence (Sclerosus, 2008, O'Connell et al., 2008). For example, LS lesions may be identified in some VSCC patients from different age groups, but LS etiology remains unclear in current vulvar disease studies (Powell and Wojnarowska, 1999). A potential cause of LS is autoimmune disorder resulting from women producing a lower amount of estrogen post-menopause. Intensive vulvar itching, soreness or burning and fragile skin are common symptoms of most non-neoplastic vulvar skin disorders and these symptoms are very similar to early stage vulvar cancer. A recent study of vulvar diseases showed that 3-5% of LS cases progressed to vulvar cancer and 25% of VSCC patients had a clinical record of LS lesions (Bleeker et al., 2016). Further, immunochemistry tests showed the aberrant expression and accumulation of TP53 protein in non-neoplastic vulvar skin dysplasia, like LS, may link LS with vulvar cancer development (Liegl and Regauer, 2006, Vanin et al., 2002). Another report highlighted tumorigeneses, like structural changes in non-neoplastic vulvar lesions, and increased inflammatory infiltration in LS regions (Jones et al., 2005). Although many LS and other non-neoplastic disorders patients are diagnosed without tumour, surgery and other treatments are required for affected women to reduce or eliminate the risk of vulvar cancer in the latter part of life.

1.1.7 Vulvar intraepithelial neoplasm

The strong carcinogenetic effect and etiological classification of vulvar intraepithelial neoplasia (VIN) have been revealed in recent decades. Patients with VIN rarely experience pain, itching and ulceration and some patients do not show obvious symptoms on the vulva (Jones and Rowan, 1994). Previously, VIN was thought to be low malignant lesions, but recently VIN has been confirmed as a direct pre-cancerous lesion of vulvar cancer. VIN lesions generally present with unique tissue histology of abnormal nuclei, disturbed structure and mitotic figures with different levels of cytoplasmic and nuclear maturation (Modesitt et al., 1998). Nearly 90% of VIN cases are caused by HPV infection and, as a result, pathogenic principles are frequently misunderstood. VIN lesions were briefly classified into three classes: VIN I (mild atypia and loss of 1/3 epidermis differentiation), VIN II (moderate atypia and 2/3 epidermis differentiation loss) and VIN III (severe dysplasia) (Edwards et al., 1996, Canavan and Cohen, 2002). This VIN classification system was based on tissue morphological patterns, rather than distinctive histopathological differences between HPV dependent and independent origins. It is now understood that VINs with and without HPV infection have different etiological causes and associated treatments (Van Den Einden et al., 2013). According to molecular evidence, VINs have been re-classified into types that include HPV infected flat condyloma, warty, basaloid and mixed lesion, and a differentiated type present in aged patients without HPV (***Figure 1.2***) (Kokka et al., 2011a). The usual type of VIN has been well studied due to its identical morphology and obvious symptoms of HPV infected tissues, it is suggested that one in four usual type VIN cases may develop into HPV-dependent VSCC if untreated (van de Nieuwenhof et al., 2009a). However, another type of VIN, differentiated VIN (dVIN), has much stronger carcinogenic effects inducing HPV-independent vulvar cancer, despite treatment. Further, dVIN tissue morphology is difficult to discriminate from benign tissues and VSCC, even by experienced pathologists (McAlpine et al., 2017). dVIN tissue frequently coexists with adjacent LS and keratinizing type VSCC in aged patients, while atypical squamous cell hyperplasia of dVIN lesions may be misdiagnosed as benign squamous cell hyperplasia with healthy hyperkeratosis (Chafe et al., 1988). Incomplete removal of dVIN is a key reason for HPV-independent vulvar cancer recurrence, but a significant number of dVIN patients have been misclassified in clinical practice (Kokka et al., 2011b). Currently, the use of genetic

profiles to distinguish dVINs from LS and early stage VSCC have not yet been established, as some similar genetic mutations are detected in dVIN and other lesions. For example, over 80% of dVIN is identified with TP53 mutation, which is also a key biomarker for HPV-independent VSCC diagnosis (van de Nieuwenhof et al., 2011, Pinto et al., 2010). Therefore, the discovery of novel biomarkers that only respond to dVIN affected tissue is required to prevent misdiagnosis of vulvar cancer.

1.1.8 Vulvar cancer carcinogenesis

Previous etiological studies have identified two distinctive oncogenetic origins of VSCCs based on vulvar cancer patient samples and different HPV infection status. Each VSCC etiological phenotype has unique tumour morphological patterns and development is associated with either by HPV infection or mutation accumulation. Therefore, vulvar cancer can be classified as having an HPV-dependent phenotype or an HPV-independent phenotype. The distinctive tumorigenesis pathways of these two phenotypes are shown in ***Figure 1.3*** (Hørding et al., 1994, Regauer and Reich, 2013).

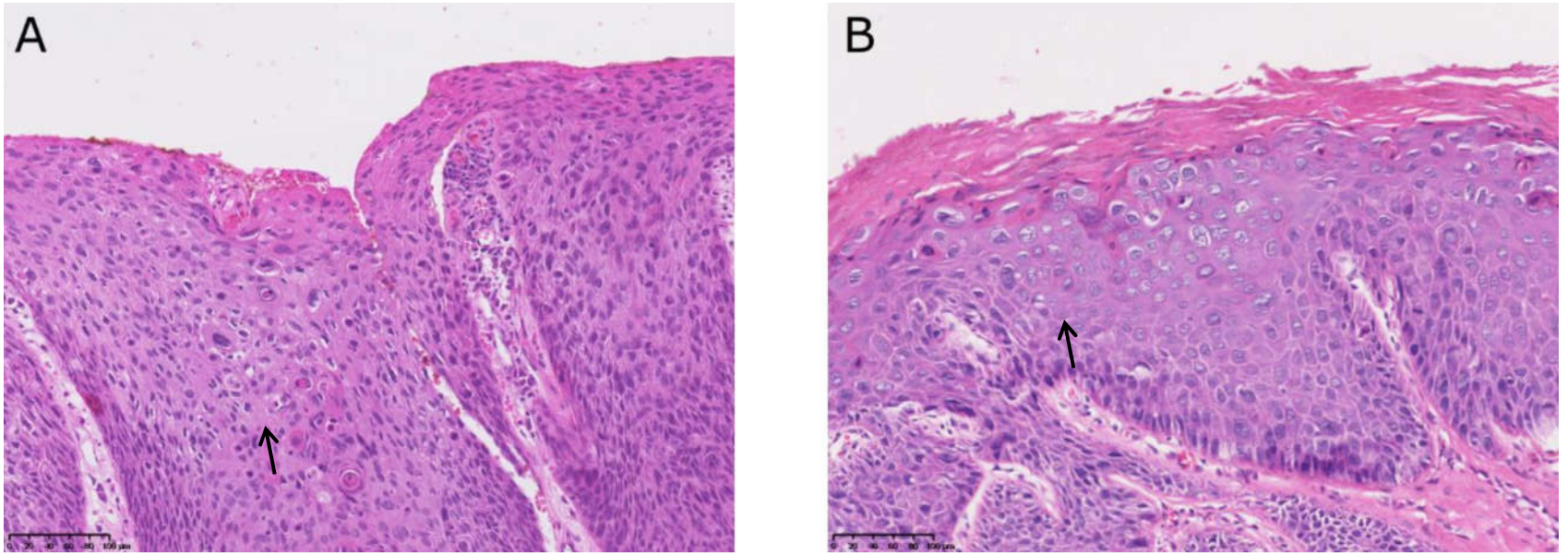


Figure 1.2: Vulvar intraepithelial neoplasia of usual type (A) and differentiated type (B) with H & E staining. A: High-grade usual type VIN displays significant warty growth and deeply invades the skin. B: Differentiated VIN showing acanthosis, variable parakeratosis and irregular elongation and anastomoses of rete ridges. Tissue morphology is difficult to discriminate from other vulvar tissues. Bar = 100µm.

.

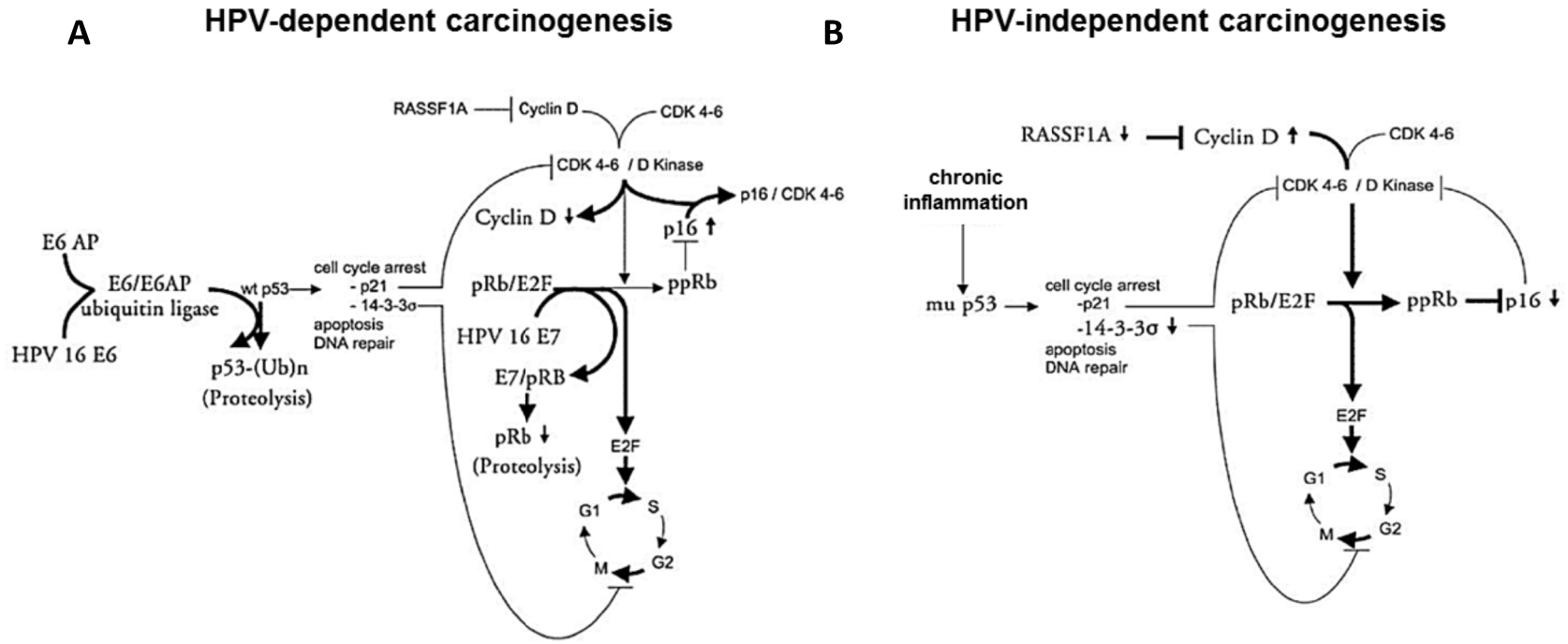


Figure 1.3: Models of two distinctive etiological pathways of vulvar cancer carcinogenesis (Leemans et al., 2011). A: HPV-dependent vulvar cancer carcinogenesis induced by HPV infection with wild-type p53 and high expression of p16 proteins. B: HPV-independent vulvar cancer carcinogenesis related to chronic inflammation with high expression of mutated p53 expression.

.

1.1.8.1 HPV-dependent vulvar cancer

HPV-dependent vulvar cancer accounts for one-third of all vulvar cancer cases. HPV-dependent vulvar squamous cell carcinoma displays a warty, basaloid or mixed cancer form and more commonly arise in young women (Parkin and Bray, 2006). This cancer directly developed from high-grade usual type VIN within one to three years. The most frequently identified virus genotype in vulvar cancer, HPV16, and other high-risk HPV genotypes including types 18, 31, 33 and 45, may all induce vulvar squamous cell carcinoma following HPV infection (Hildesheim et al., 1997, Hørding et al., 1994). Similar to HPV-dependent tumorigenesis in HPV positive cervical, vagina and anus cancer, the progression of VSCC is triggered by expressions of virulent genes, such as HPV genes E6 and E7, in host vulvar keratinocytes which lead to infected skin cell immortalization and transformation. Immortalized keratinocytes develop into malignant cancer when other oncogenes are activated (Read and Douglas, 2014). Expression of E6 and E7 are widely detected in basal epithelial cells of precancerous lesions in PCR analysis (Park et al., 2000). Interestingly, usual type VIN and HPV-dependent vulvar cancer have negative immunohistochemistry staining results for mutated p53 protein, while genetic studies indicate that most HPV E6/E7 tumorigenesis mainly regulates p53 independent signalling pathways and are not directly involved in TP53 regulation. However, E6 protein complex can interact with another cellular factor to suppress transcription of TP53 genes or lead to p53 inactivation (Munoz et al., 2006, Lee et al., 1994, Sano et al., 1998), while virus E7 protein binds to another tumour suppressor, retinoblastoma protein (pRB), to decrease antitumor activities. As a consequence, inactivation of pRB leads to significant increase in downstream p16 protein expression which is encoded by the CDKN2A gene in chromosome 9p21 and regular cellular cycle via inhibiting CDK4 or CDK6 complex (Ishii et al., 1999). These changes facilitate cells from HPV infected lesions to enter S phase. Strong immunohistochemical staining of p16 on vulvar tissues has become a hallmark of usual type VIN or HPV-dependent vulvar cancer (Sano et al., 1998). Previous studies show that p16 expression has high specificity and sensitivity in HPV-dependent vulvar cancer diagnosis and only 1% of HPV-independent vulvar cancers show false positive to p16 stain (Lerma et al., 2002).

1.1.8.2 HPV-independent vulvar cancer

Development of HPV-independent vulvar cancer from dVIN takes a longer period of time (over 10 years), as compared to HPV dependent tumorigenesis (Sideri et al., 2005). About 60% of vulvar squamous cell carcinomas are not free from HPV infection and a majority of HPV-independent VSCC patients are diagnosed at post-menopausal age, especially patients with the previous history of LS and dVIN (Hørding et al., 1994). HPV-independent VSCCs are generally classified into keratinizing types of vulvar cancer which present large, well-differentiated keratinocytes and keratinized pearl structure in epidermoid carcinoma (Canavan and Cohen, 2002). Compared with HPV-dependent vulvar cancer, HPV-independent type has a much higher metastatic risk and its pre-cancerous lesion (dVIN) always demonstrates benign-like morphology or is completely asymptotic until it progresses to cancer (McAlpine et al., 2017). The characteristics of HPV-independent vulvar cancer described above mean that this form of cancer is often not diagnosed until later stages (Modesitt et al., 1998, van de Nieuwenhof et al., 2011). Further, the direct etiological cause of HPV-independent vulvar cancer remains unclear. One hypothesis is an accumulation of inflammatory mutations in previous differentiated VIN and LS (van de Nieuwenhof et al., 2009c). Somatic mutation of TP53 gene in chromosome 17p has been identified in over 80% of HPV-independent VSCC cases, while other genetic changes, including loss of 4p and 3p, have also been detected (Gordinier et al., 1997). Expression of non-functional TP53 protein could facilitate tumour cell escape from apoptosis and overexpression of mutated p53 is frequently reported in keratinizing vulvar carcinoma patients with poor prognostic outcomes (Vanin et al., 2002). Mutation of p53 protein is an important classification parameter of HPV dependent/independent vulvar cancer (Hay et al., 2016). However, high abundance of TP53 proteins are also found in LS and dVIN patients without vulvar cancer and p53 mutation may not trigger malignant progression of HPV-independent carcinogenesis (Muller and Vousden, 2013). For example, DNA sequencing of HPV-independent vulvar cancer and adjacent LS tissue samples indicated that overexpressed p53 proteins in LS are wild-type, whereas non-functional TP53 proteins with loss of another gene in 17p were found in malignant tumour regions (Vanin et al., 2002). Identification of other biomarkers for HPV-independent vulvar cancer has been explored. Immunohistochemical studies show that estrogen receptor alpha and beta are significantly

decreased in VSCC regions, as compared to the healthy epithelium and further analysis suggests estrogen receptor signalling pathways play an important role in HPV-independent vulvar cancer development, rather than p53 (Zannoni et al., 2011). Also, ubiquitin-proteasome and interferon signalling regulated proteins have been identified with different abundance in VSCC during mass spectrometry vulvar cancer profiling and future mass spectrometry analysis may highlight HPV-independent signalling cascades to provide a better understanding of HPV-independent VSCC progression and potential therapeutic targets (Sandberg et al., 2012).

1.1.9 Vulvar cancer diagnosis

Lack of awareness and sensitive diagnostic methods are the top two factors that contribute to vulvar cancer prevalence. LS and dVIN may develop without significant symptoms and patients often self-medicate for itching or bleeding vulva skin, rather than seeking clinical assistance (Board, 2016). Thus, most vulvar cancer patients present at a clinic when significant symptoms are present and cancer have progressed to stage IB or II. Biopsy of suspected warts on the vulvar skin is the most direct method to identify vulvar cancer. This biopsy process involves sampling small pieces of vulvar skin of at least 1mm depth for pathological microscopy (Ansink, 1996). However, sometimes multiple skin biopsies and sentinel node biopsy are required to accurately identify cancer stage and such invasive sampling causes unnecessary pain for patients (De Cicco et al., 2000). While biopsy can offer direct histological evidence regarding patient vulvar cancer lesions, discrimination of VSCC stages and grades are highly dependent on pathologist observation and experience. Alternative diagnosed methods, such as imaging examinations, are accessible in most hospitals for vulvar cancer diagnosis, but X-ray and CT technologies only detect larger sized primary tumours and MRI or PET are mainly used to trace lymph node metastasis (Hauspy et al., 2007). However, all of these imaging examination methods are not sensitive to small tumours in early stage cancer. Therefore, imaging methods are not recommended to diagnose vulvar cancer unless vagina, anus and urethra invasion is suspected (Board, 2016).

Considering the sensitivity and cost-effectiveness of current vulvar cancer diagnosis, applying molecular markers to vulvar cancer early diagnosis may be a complementary approach to microscopic or imaging tests. Using genetic or proteomic biomarkers has demonstrated advantages in classifying two different cancer etiologic phenotypes and can discriminate pre-cancerous stage lesions with high sensitivity and specificity (Hørding et al., 1994, Kokka et al., 2011a). Many proteins have shown the potential to discriminate different vulvar cancer phenotypes and precursor lesions from biopsy samples. Proteins that regulate tumor growth, including p16, p53, VEGF, CDC1 and EGFR, and other gynecological cancer biomarkers, such as HER2 and CA125, have demonstrated different levels of correlation with cancer progression stages (Gordinier et al., 1997, Lerma et al., 2002, Hefler et al., 1999, Olawaiye et al., 2007, Borras et al., 1995). A recent study

reported the expression of cell proliferation biomarker ki-67 in the early stage of squamous cell carcinoma, rather than expression in VIN lesions and normal epithelium (Modesitt et al., 2000b). Identification of p16 in HPV-dependent vulvar cancer in clinical practice has been identified, as over-expression of p16 is almost 100% specific to HPV-positive lesions (Lerma et al., 2002). On the contrary, HPV-independent vulvar cancer with LS and dVIN frequently show a high abundance of p53 protein expression, as compared to benign epithelium or HPV-infected tissues (Vanin et al., 2002). However, precursor dVIN and vulvar squamous cell carcinoma show an abnormal increase of p53 expression in lesions. Thus, wide local excision is required to eliminate potential tumour remains which cause additional damage and pain for vulvar cancer patients (Hay et al., 2016). Few biomarkers that distinguish VIN from VSCC have been identified in recent genetic studies. DNA methyltransferases (DNMT) have the ability to distinguish cancer from VIN lesions and over-expression of DNMT3A is strongly associated with potential tumours remaining in HPV-dependent and independent vulvar cancer patients (Leonard et al., 2016). Cytokeratin 5, as an intermedia filament protein expressed in keratinocytes, has been revealed as a potential biomarker of abnormal tumour proliferation in vulvar adenocarcinoma and cancer phenotypes (Prince et al., 2007). In the published manuscript, MALDI-MSI results show a higher abundance of cytokeratin 5 in dVIN, in contrast to the significantly lower expression level in HPV-independent vulvar squamous cell carcinoma (Zhang et al., 2016a). As an abundant protein, cytokeratin 5 can be easily detected in dVIN by multiple methods, which presents a potential solution to current misdiagnosis of HPV-independent vulvar cancer.

1.1.10 Vulvar cancer prognosis and metastasis

In general, vulvar cancer is not a deadly disease compared to high mortality gynecological cancers, such as breast or cervical cancer. The 5 year survival rate of VSCC patients was 71% for 2008-2012 and survival rate amongst negative lymph node patients exceeded 80% (AIHW, 2016) However, later stage VSCC patients have worse prognostic outcomes than dVIN and stage I VSCC patients, as survival rate may drop to less than 41% when inguinal positive nodes reported and only 10-15% of patients with iliac or pelvic positive nodes survive the first 5 years post-treatment (Edwards et al., 1996). Poor prognosis of vulvar cancer is often related to older patient age, terminal cancer stage, smoking and inguinal lymph node metastasis, which are potential determining factors for vulvar cancer prognosis. To improve vulvar cancer prognostic results, patients diagnosed with stage IB or above cancer receive compulsory vulvectomy and lymphadenectomy without positive node detection (Aviki et al., 2015). Aggressive groin node dissection improves overall prognostic outcomes, although extra treatments on lymph nodes are unnecessary for patients with *in-situ* carcinoma. Thus, accurate prediction of potential lymph node metastasis risk for primary VSCC tumour is a promising strategy to avoid VSCC patient overtreatment. Few histopathological characteristics of vulvar squamous cell carcinoma are highlighted in connection to the presence of positive lymph nodes. Previous studies indicate that grade 1 and 2 tumours have a higher metastasis-free survival rate, while grade 3 tumours are more commonly found in patients with positive nodes (Raspagliesi et al., 2003). There is no evidence to suggest that HPV infection favours tumour metastasis and HPV-dependent vulvar cancer shows better prognostic outcomes due to younger patient age (Munoz et al., 2006). However, HPV-independent vulvar cancer is always associated with the high possibility of lymph node metastasis, even when patients have received necessary treatment (Rasmussen et al., 2018).

The application of proteomic or genomic profiles to classify LNM positive and LNM negative patients has not yet been fully addressed (Han et al., 2014). Tumours with potential metastasis risk may demonstrate different morphological patterns during invasive growth as these tumours can either push tumour stroma borders or infiltrate single cells. Tumours with this type of infiltration pattern more easily invade lymph nodes (Wan et al., 2013). Currently, LNM positive and LNM

negative VSCCs are difficult to determine by microscopic analysis, but recent research suggests that epithelial to mesenchymal transition (EMT) may be an important biological process triggering malignant squamous cell carcinoma invasive growth to sentinel lymph nodes (***Figure 1.4***) (Frederick et al., 2007, Tsai et al., 2012). Tumour EMT changes demonstrate stem cell-like patterns including low cell-cell adhesion, disassembly, loss of polarity, increasing cell mobility and apoptosis resistance (Roth et al., 2016). EMT-related proteins have been highlighted in the tumour progression of other gynecological cancer (Zhou et al., 2014). For example, E-cadherin functions as a cancer invasion and metastasis suppressor and insufficient E-cadherin expression result in reduced cell-cell adhesion in sporadically epithelial tumours (Zannoni et al., 2011). Expression E-cadherin also regulates estrogen signalling and the epidermal growth factor receptor (EGRF) pathway activity in tumour infiltrative growth (Andl et al., 2003). Further, up-regulation of another EMT marker, beta-catenin, was reported in tumour cell proliferation and migration in metastasised vulvar carcinoma (Rodrigues et al., 2013b). In HPV-independent vulvar cancer, alternation of Wnt/beta-catenin pathway is thought to promote tumour invasion and metastasis genes via accumulation of beta-catenin in tumour cytoplasm (Eckey et al., 2014). Thus, E-cadherin/beta-catenin alternation may lead to the cellular cytoskeleton and intermedia filament change inducing delocalisation of primary vulvar squamous cell carcinoma (Rodrigues et al., 2013a). EGFR is also an EMT marker found in most vulvar squamous cell carcinoma and abnormal EGFR expression is correlated with inguinofemoral lymph node metastasis in other cancer types (Tang et al., 2016). Vimentin is another protein down-regulated by EGRF in the EMT process. The high abundance of vimentin is shown in immunohistochemical staining from metastasis vulvar cancer with poor prognostic results (Sand and Thomsen, 2018). Other EMT hallmarks, like Snail1, do not show a significant change in metastatic progression of vulvar cancer (Prieske et al., 2016). However, increased expression of collagen I was found in vulvar and other gynecological cancers with positive node status. The receptor of collagen I stabilised Snail 1 protein and induced Snail 1 accumulation in nuclei, further promoting tumour invasion (Gonzalez et al., 2015). With the identification of EMT-like events in vulvar cancer progression, using EMT-related proteins to predict potential metastasis risk could significantly improve prognostic outcomes of vulvar cancer patients.

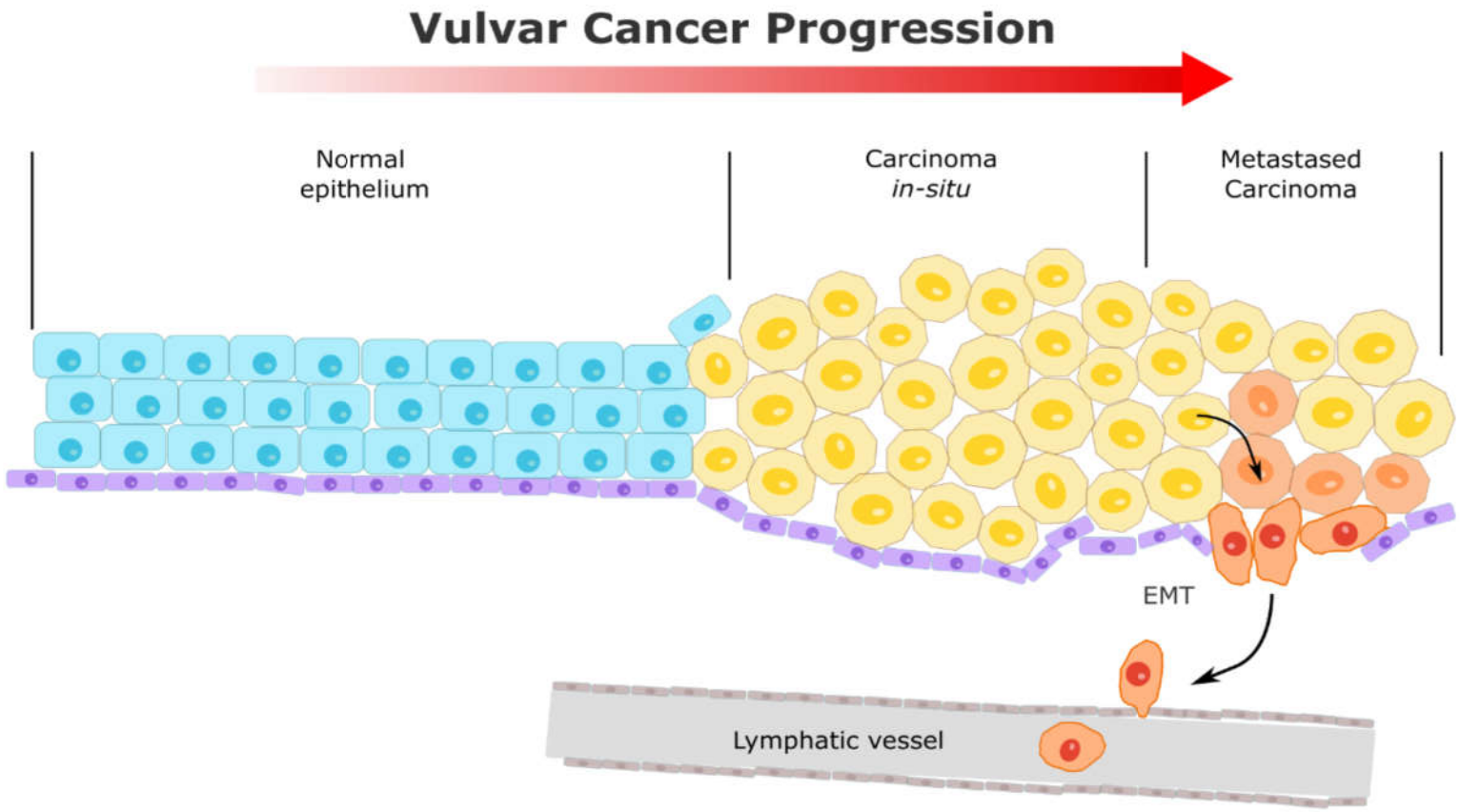


Figure 1.4: Model of HPV-independent vulvar cancer tissue types presenting in different cancer development stages, including benign vulvar epithelium (colored in blue), vulvar squamous cell carcinoma *in-situ* (colored in yellow), metastasized vulvar squamous cell carcinoma in vulvar (colored in red) and invaded in sentinel lymph nodes (colored in grey). Normally, metastasized and non-metastasized vulvar squamous cell carcinoma is hard to be separated based on their morphological appearance.

1.1.11 Vulvar cancer metastasis prediction

Determination of lymph node status is vital to improving the survival rate of vulvar cancer patients. However, over-treatment of non-metastasis patients is widespread as no effective metastatic predictor is applied in current vulvar cancer practice. Most general practitioners apply palpation to determine inguinofemoral lymph nodes (Yamaoka et al., 2016). While this is a direct and simple method, it has a very low sensitivity to detect lymph node metastasis under fatty tissue or surgical scarring. Ultrasound is a highly accepted non-invasive method for lymph node metastasis detection in most vulvar cancer patients. A previous study showed 84% sensitivity and 97.2% specificity using ultrasound instruments to classify 79 vulvar cancer patients (Pinto et al., 2016). However, ultrasound tests sometimes result in a false negative prediction of positive nodes due to issues with instrumental resolution. Other imaging scans, namely CT and MRI, are also used as non-invasive methods for deep inguinal node screening, but these methods are challenged by relatively higher false negative rates, as compared to ultrasound, in high-risk patients (Thaker et al., 2015). Positron emission tomography has been frequently used to determine vulvar cancer node status and has high accuracy in tracing metastases over 10mm, but a recent study showed that only 23% of small size (<5mm) metastasis nodes are detected (Collarino et al., 2015). Thus, sentinel node biopsy is the most accurate and reliable method for detecting lymph node metastasis of any size, meaning that this invasive method results in a painful experience for vulvar cancer patients (Robison et al., 2014).

Recent studies have attempted to define a quantitative relationship between histopathological parameters and risk of lymph node metastasis, but failed to make conclusions regarding appropriate false discovery rates (Oonk et al., 2010b). Therefore, application of biomolecular parameters for metastatic prediction is currently considered a sensitive and non-invasive method for vulvar cancer practice. Genetic analysis indicates that p14 represents a better connection with HPV-dependent vulvar cancer metastasis than p16 and p53 (Lavorato-Rocha et al., 2013). Further, upregulation of L1CAM (cell adhesion regulator) has better indicative results regarding primary tumour cell delocalization compared to prediction results using p53 and Wnt signalling proteins (Trietsch et al., 2016). Cytokeratin proteins are ideal candidate parameters due to high abundance

in mutated skin keratinocytes of HPV-independent vulvar cancer (Zhang et al., 2016a). IHC of 82 lymph nodes showed that cytokeratin 19 has 100% specificity for stage IV vulvar cancer, presenting a promising predictor for future clinical trials (Pappa et al., 2014). Expression of Rho-associated coiled-coil-containing protein kinase (ROCK1), which controls the actin-remodelling process in EMT, increased in patients with lymph node metastasis (Akagi et al., 2014). Previously, overexpression of flotillin-1 in tumour region was reported as a prognostic marker of cervical cancer and correlated with cancer FIGO stage, further IHC stain with flotillin-1 protein showed a promising result on stage classification of 70 HPV-dependent VSCC samples (Gao et al., 2016). Transcription factors, like Twist, Snail1, Slug and Zeb1/2, are hallmarks of the cancer EMT process which can be used to estimate metastasis risk (Rodrigues et al., 2013a). A better understanding of proteomic profile changes during vulvar cancer EMT and metastasis processes is required to establish more reliable prediction models of vulvar cancer development.

1.1.12 Treatment of vulvar cancer

Surgery in conjunction with chemotherapy and radiotherapy are classic treatment strategies for vulvar cancer patients (Thaker et al., 2015). Vulvar surgery, including laser excision, local excision, vulvectomy and lymph node removal, can treat a majority of vulvar cancer lesions. Local excision is used to treat most VIN and *in situ* carcinoma lesions and an extra large piece of skin is removed to prevent further malignant lesion development. Chemotherapy, such as fluorouracil ointment, is a non-invasive treatment choice, but has a higher risk of local recurrence as compared to surgical lesion removal. FIGO stage IB lymphadenectomy requires removal of all adjacent nodes to reduce the risk of vulvar cancer recurrence (Board, 2016). Even when no positive nodes are found in node biopsy, patients with a greater than 2mm size tumour or 1mm stromal have adjuvant nodes removed. As described above, a lack of clinical reference regarding prognostic outcomes is the key reason for current over-treatment. Previous clinical reports suggest that adjuvant radiotherapy on VSCC patients can significantly improve prognostic outcomes when less than 2-3 nodes are present (Woelber et al., 2011). In the terminal stage of vulvar cancer, mortality risk is significantly high even the following surgery with radiotherapy or chemotherapy (Hacker et al., 1986). Currently, distant metastasized tumours are very difficult to detect and remove and identification of new therapeutic targets may create new solutions for future vulvar cancer treatment.

1.1.13 Indigenous vulvar cancer

Although vulvar cancer is generally a rare type of gynecological cancer in Australia, the incidence rates amongst Indigenous women is significantly higher compared to the non-Indigenous population (Condon et al., 2005). The prevalence and prognostic outcomes of different cancers for Australia's Indigenous population has not been well-studied until more recently. Compared with the non-Indigenous population, Indigenous vulvar cancer patients suffer medical disadvantage in cancer prevention and treatment, especially women living in remote communities (Anderson et al.). Indigenous Australians present with distinctive incidence patterns for different cancers. For example, incidence rates of colon, prostate and breast cancer are much lower than national levels, while liver, cervical and vulvar cancer are much higher than that observed in the non-Indigenous population (Condon et al., 2003). Medical records show that Indigenous patients are frequently diagnosed at a younger age and much more vulnerable to metastasis cancer (Cunningham et al., 2008, Valery et al., 2006). Unfortunately, vulvar cancer and other vulvar diseases are common amongst women in Indigenous communities. Amongst women under the age of 64, vulvar cancer incidence for Indigenous women from the Northern Territory was 12 times more than the national average levels between 1991 and 2001, which is the highest cancer rate including liver, lung and other gynecological cancers (Condon et al., 2005). There is no evidence that common cancer risk factors, like smoking and alcohol, are responsible for the high prevalence of vulvar cancer in Indigenous communities. However, clinical records suggest that the prevalence of VIN lesion and vulvar cancer are correlated with specific geographic distributions in northern Australia (Condon et al., 2009b). In East Arnhem land, the Indigenous incidence rate is over 31.1 per million, which is significantly higher than the rate of 5 per million in nearby regions and 50 times greater than the non-Indigenous population (Condon et al., 2009a). The HPV infection is frequently detected in Indigenous cervical cancer patients, so HPV prevalence in Indigenous communities is suspected as a major risk factor for vulvar cancer (Chen et al., 1999, Grulich et al., 2010, Rosenthal et al., 2007). However, there is no difference in the number of women with HPV genotype16/18 infection between Indigenous and non-Indigenous populations (Garland et al., 2011). Even though infection rates are similar, Indigenous women have a much higher risk of developing vulvar cancer after HPV infection (Condon et al., 2005). It has been shown that the local environment and HPV

infection cannot explain the high prevalence of dVIN lesions and HPV-independent vulvar cancer in the Indigenous population (Rumbold et al., 2012a). The successful Human Genome Project has facilitated an increasing number of studies investigating genomic profile differences between Indigenous and non-Indigenous patients. However, there is no unique gene family or specific genes that have been identified as being involved in the high incidence rates of Indigenous vulvar cancer and VIN lesions.

Previous research on vulvar cancer suggests that HPV infections, tobacco consumption and other unhealthy lifestyle factors could explain the high prevalence of vulvar cancer and precursor lesions among Indigenous communities from East Arnhem Land (Scurry and Vanin, 1997). Initially, the younger age of Indigenous patients and higher prevalence of VIN indicates that oncogenic HPV infection, such as HPV16 or 18, might be a major risk factor for Indigenous vulvar cancer as HPV infection ranges from 1.7% to 10.9% in different Indigenous communities (Muñoz et al., 2003). However, no significant change of HPV prevalence has been reported among Indigenous patients from East Arnhem Land compared with other Australian patients in HPV16 gene PCR identification or p16 protein IHC staining (Rumbold et al., 2012b). In the HPV test of 2152 Australian women, HPV16, HPV18 and other HPV genotypes did not show any obvious difference between Indigenous and non-Indigenous young women and HPV68 was the only high-risk HPV type that showed higher prevalence among Indigenous women (Garland et al., 2011). Interestingly, high-risk HPV56, 58 and low risk 71, 72, and 81 showed a higher incidence rate in remote areas, but those genotypes were relatively rare among vulvar cancer patients due to weak carcinogenesis effects. Currently, no evidence suggests HPV infection is associated with higher Indigenous vulvar cancer incidence. Thus, tropical weather and inherited oncogenic effects of HPV infection may be alternative explanations for higher carcinogenesis risk associated with HPV in East Arnhem Land. Although no evidence shows the risk of vulvar cancer related to smoking, cancers associated with smoking always demonstrated higher incidence and much worse prognostic outcomes among Indigenous Australians (Coory et al., 2000). Tobacco and alcohol consumption among Indigenous consumption, as well as limited housing and access to medical treatment, are also strongly connected with vulvar cancer development. Smoking is a vital health problem among the Indigenous population with a long history. Nearly half of Indigenous women were addicted to tobacco and over 70 % of Indigenous were reported with smoking or chewing tobacco in northern

Australia (Condon et al., 2005). The number of Indigenous smokers is increasing in the NT and other Indigenous Australian communities. Unfortunately, the connection between tobacco consumption and vulvar cancer development is poorly understood. Lacking necessary medical assistance and education also contributes to the higher vulvar cancer mortality rates among Indigenous women. Although hospitalization rates of Indigenous and non-Indigenous Australians are similar, cancer-related hospitalizations are very low for Indigenous patients which results in vulvar cancer advancement. Previous data reported that 24% of Indigenous cancer patients were less likely to receive surgery or other treatments after diagnosis (McGrath and Rawson, 2013).

Currently, no genetic or proteomic analyses have been completed that are able to fully explain the higher incidence rates of vulvar cancer in the Indigenous Australian community, however, such analyses have helped explain the prevalence of other pressing conditions. Previously, whole genome analysis of Indigenous Australians has been performed for type 2 diabetes-associated genes highlighting that the Indigenous population may harbour unique genetic mutations related to diabetes and other diseases in chromosome 2q (Busfield et al., 2002). The mitochondrial genome also showed different genetic patterns among the Indigenous population compared with the non-Indigenous population (van Holst Pellekaan et al., 2006). Genome-wide homozygosity clustering was performed on Indigenous vulvar cancer patients from East Arnhem Land, but results showed that loss of homozygosity failed to highlight a gene related with high vulvar cancer prevalence and mortality (McWhirter et al., 2014). Mannose-binding lectin (MBL) is an important protein activating the complement system during in anti-cancer immunoreactions. However, DNA sequencing of the MBL gene and its promoters demonstrated that Indigenous Australians have unique genetic changes which interfere with MBL structure and expression, and deficiency of MBL proteins has been shown to increase risks of infections and other autoimmune diseases, like cancer (Turner et al., 2000). Currently, regulation of inflammation during Indigenous vulvar cancer development is poorly studied. But the frequency of cytokine gene single nuclear polymorphisms and other genetic variation changes show differences compared with the Caucasian population which may explain the higher risk of chronic inflammatory diseases, like cancer or diabetes, within the Indigenous population. For example, the tumour necrotic factor (TNF) and other anti-inflammatory cytokines of Indigenous people have special genotypes rarely found in non-Indigenous populations (Cox et al., 2014a). A recent proteomic study reported two

serum inflammatory markers, C-reactive protein (CRP) and soluble E-selectins, were overexpressed in Indigenous serum samples collected from remote communities (Brodt et al., 1997). Although the relationship between inflammation and vulvar cancer is still poorly understood, high CRP level has been frequently detected in serum samples from Indigenous patients with renal and cardiovascular problems, while overexpression of E-selectin in endothelial cells could facilitate tumour metastasis through the lymph-vascular system by changing cell-cell adhesion (Rowley et al., 2003, McDONALD et al., 2004). Interestingly, the geographical location of remote Indigenous communities, East Arnhem Land as an example, results in higher CRP level and high prevalence of HPV-independent vulvar cancer (Shemesh et al., 2007). Further, inadequate nutrition and physical inactivity were also found in Indigenous populations from remote areas, factors which may accelerate the development of vulvar cancer. Therefore, identification of unique proteomic patterns of Indigenous patients necessary to explain the high prevalence of dVIN and HPV-independent vulvar cancer amongst Indigenous Australians.

1.1.14 Previous molecular biomarkers of vulvar cancer

In past decades, genetic, proteomic, glycomic and metabolic analyses have studied vulvar cancer carcinogenesis and metastasis. Multiple mutations have been observed in the genetic sequencing of vulvar cancer tissues. In past clinical diagnosis, atypical verruciform proliferation was regarded as a benign lesion without tumorigenesis risk on the HPV-negative vulvar skin, as verruciform lesions do not show a TP53 mutation in DNA sequencing (Lee et al., 1994). However, alternative oncogenes, such as PIK3CA which activates the PI3K/AKT/mTOR pathway, have shown significantly higher expression in atypical verruciform lesions with vulvar squamous cell carcinoma, and HPV-dependent vulvar cancer recurrence has been associated with PIK3CA and AKT1 mutations (Watkins et al., 2016). Thus, vulvar verruciform lesions may be another pre-cancerous lesion alongside the VINs. In addition to significant chromosome loss and the TP53 mutations mentioned above, genotype variation between patients results in distinctive outcomes for cancer treatment. Single nucleotide polymorphisms (SNPs) analysis of 32 potential gene targets related to the toll-like receptor and nuclear factor kB pathways showed that women with different allele SNPs had different results of HPV-induced tumorigenesis (Bodelon et al., 2014). Whole genome SNP tag searching shows that genetic variation of lymphocyte receptors, such as interleukin 2, may result in higher vulvar cancer risk among the smoking population with unique SNPs. Even without genetic mutations, DNA methylation status may regulate vulvar cancer carcinogenesis and unique methylation-specific sequences have been revealed in different vulvar lesions. Aberrant DNA methylation has shown gene regulation effects without mutation. Hypermethylation of RASSF2A and MGMT genes restricted to LS lesions coexisted with vulvar squamous cell carcinoma, but were absent in independent vulvar LS lesions (Frost et al., 2016). Further research indicated that the hypermethylation of TSP-1 proteins may be used to determine poor prognostic cases from other vulvar cancer patients (Trietsch et al., 2015). Recent studies of Indigenous Australian vulvar cancer samples revealed unique single nucleotide polymorphism (SNP) frequencies in eight important cytokine genes which regulate immune response of intensive inflammatory diseases (Cox et al., 2014a). However, no other genetic mutation has been identified as unique to Indigenous vulvar cancer patients which suggests other risk factors, such as environment or imbalanced nutrition, could be the potential cause of higher vulvar cancer

incidence rates in Indigenous communities, with chronic inflammation being one of the most important causes of HPV-independent vulvar cancer among aged women.

While there is no HPV-independent vulvar cancer genome or transcriptome database, many protein and metabolite expression profile changes have been identified in previous proteomic and metabolomic work. Protein expression changes of some tumour biomarkers, including CD9, CD63 and CD82, have been found in VSCC patients, especially down-regulation of CD82 gene expression in most poor prognosis cancer cases (Bodelon et al., 2012). Another biomarker, S100P protein, showed stronger immunoreaction in the invasive vulvar cancer front compared with tumour cores, which may be used to predict potential metastasis risk of primary VSCC (Botelho et al., 2015). Currently, the unique metabolism of metastatic vulvar cancer remains unclear. However, decreased blood estrogen levels and high incidence rates amongst the aged population have been reported in many gynecological cancers, including vulvar cancer (Sherman et al., 1994). Matrix metalloproteinase expression may be correlated with FIGO stages of VSCCs, and high abundance of MMP-2 and 9 promote cancer invasive behaviour (Binder et al., 2015a). Escape from host immune system surveillance is the key mechanism for cancer metastasis into other tissues. In a recent mass spectrometry based study, lower tryptophan and higher kynurenine were reported in the serum of HPV-dependent and independent vulvar cancer patients, which indicates that the kynurenine metabolic pathway and tryptophan degradation are present in malignant VSCCs. Generally, the number of ordinary T cells and Tim-3 positive T cells in common VIN or HPV-dependent vulvar cancer are much higher than that of benign tissue, resulting in higher expression of indoleamine 2,3-dioxygenase, galectins and other coinhibitory molecules in serum (Hung et al., 2008, Kohrenhagen et al., 2010). Galectins are a series of glycoproteins expressed in the cytomembrane and extracellular matrix to protect metastasized tumours in the lymphatic system, and abundant levels of galectin-1 show a strong correlation with histopathological vulvar cancer grade (Kohrenhagen et al., 2010). Further research of identified metabolic pathways is needed to deliver better classification between metastasis and non-metastasis vulvar cancer and provide new drug targets in vulvar cancer chemotherapy.

1.2 MASS SPECTROMETRY

Mass spectrometry (MS) is a rapid analytical tool for the identification of biomolecules, including nucleic acids, peptide, proteins, lipids, glycans, metabolites and therapeutic drugs, from biological samples (Gessel et al., 2014, Gika et al., 2014, Holst et al., 2016). In current proteomic studies, MS has shown the advantages (***Table 1.2***) of being a high throughput profiling and quantification tools applicable to various sample types such as plasma and tissues (Schöne et al., 2013, Bantscheff et al., 2007). Briefly, mass spectrometers have three components: an ion source that ionizes peptides, proteins or other analytes; a mass analyzer which separates analytes based on mass to charge ratios; and a detector that receives separate analyte signals. Currently, matrix-assisted laser desorption/ ionization (MALDI) and electrospray (ESI) are two dominant types of ion sources in MS analysis. In the case of protein research, ion sources are frequently used to transfer intact proteins or peptides into analytical ions (Aebersold and Mann, 2003, Lim et al., 2003), and are generally coupled with rapid time of flight (TOF) analyser, ion trap analyser, high mass resolution Orbitrap and Fourier-transform ion cyclotron resonance mass analysers in MS instruments (Kellie et al., 2010). To explore the proteomic patterns of different stages of vulvar cancer development, this study selected MALDI-TOF-MS and *nano*-LC-ESI-MS (qTOF and Orbitrap) to reveal protein abundance and spatial distribution differences in vulvar cancer patient samples.

Table 1.2: Comparing different analysis approaches in the current vulvar cancer study.

	Genetic analysis	**Histopathology**	**Immunohistochemistry**	**Mass Spectrometry**	
				MALDI-TOF-MS	***nano*-LC-ESI-MS/MS**
Methods	PCR DNA/RNA microarray	H&E stain	Enzyme or fluorophore-conjugated antibody staining	Protein/peptide profiling or imaging	labelling or label-free identification or quantification
Applications	Detect genetic mutations	Tissue morphological patterns	Candidate protein distribution on tissue	Protein identification with spatial distribution	Protein identification quantification
Advantages	Sensitive Well established Reproducible	Quick Easy Direct	Sensitive Specific Spatial information	Rapid High-throughput No database required Spatial information	High sensitivity High throughput High resolution De novo sequencing
Disadvantages	Previous knowledge required Limited resolution Unable to find protein post-transitioned	Non-specific Highly descriptive Bias to pathologist	Limited protein detection Expensive Time and labour intensive Required dilution adjustment	Low protein fraction Limited identification Spatial resolution Non-quantitative	Time intensive Expensive Labour intensive No spatial information

1.2.1 MALDI-TOF-MS and MS/MS

The application of MALDI-TOF-MS and tandem MS (MS/MS) to characterize proteins and other biomolecules from various biological samples has been successfully established over recent decades (Caprioli et al., 1997). Target analytes are co-crystallised with matrices and vaporized using absorbed laser power in the MALDI source. To detect peptides or proteins from cancer tissue samples, cyano-4-hydroxycinnamic acid (CHCA) and 2, 5-Dihydroxybenzoic acid (DHB) are frequently selected for digested peptide or other low mass molecule analysis, due to better analyte ionisation efficiency, whereas sinapinic acid (SA) and other matrices display better performance in large molecular weight proteins over 10kDa (Caprioli et al., 1997). Digested peptide fingerprinting and sequencing are two major strategies of protein or peptide identification in the MALDI-TOF instrument. Peptide mass fingerprint (PMF) is a direct peptide identification approach that searches acquired intact mass of tryptic peptides against a library of *in-silico* digested proteins (Suckau et al., 2003). However, PMF is rarely used as a current identification toll given many peptides share the same or close masses, and sequence modifications and in source decay may result in false positive matches within the *in-silico* database. On the other hand, peptide sequencing via MS/MS offers reliable protein or peptide identification results from complex biological samples (Vitorino et al., 2008). Currently, the MALDI-TOF instrument employs a novel LIFT technique that provides high sensitivity for fragmented ions from 2D gel or tissue samples. Unlike peptide identification by PMF, performing peptide *in-situ* MALDI-TOF-MS/MS on tissue sections has a relatively low false discovery rate and is able to reveal potential modifications of the peptide. Previous experiments show that over 80% of protein spots in a 2D gel could be identified through MALDI-TOF-MS/MS sequencing of digested peptides (Neubert et al., 2004).

Although it is a high-cost instrument, MALDI-TOF-MS is a high throughput and high resolution technique that allows for the accurate detection of proteins from samples such as bacteria, parasites, plants and mammalian samples (Villanueva et al., 2004, Sandrin et al., 2013). Application of MALDI-TOF-MS to microbe identification, including bacteria and fungi detection, has many advantages including speed, high throughput, low cost and accurate identification, as compared with current PCR and other genetic methods (Patel, 2015). For example, the MALDI biotyper

system can discriminate spectral signatures of different bacteria strains via searching the pre-built microorganism MS spectra database (Sogawa et al., 2011). In plant proteomics, MALDI-TOF-MS proves a more reliable method to characterize high molecular weight glutenins in different wheat genotypes compared with other proteomics methods, such as SDS-PAGE and liquid chromatography (Gao et al., 2010). Along with proteomic characterization, MALDI-TOF-MS can be used to highlight different lipids from complex samples, including hydrocarbons, plant pigments, carotenoids, fatty acids, cholesterol, sphingolipids, glycolipids and mixtures (Fuchs et al., 2010). Peptide *de novo* sequencing using MALDI-TOF-MS/MS has many advantages in proteolytic peptide sequencing from incomplete or non-genome organisms. In a recent study, 158 unique proteins were digested, sequenced and identified from dinoflagellate species which had not been identified and recorded in the genome database previously (Wang et al., 2011). Furthermore, MALDI-TOF-MS/MS may also be used in monitoring the synthesis of peptides or peptide-like compounds in biomolecular affinity experiments due to analytical speed and high salt tolerance (Lee et al., 2009).

Protein or peptide profiling of patient samples using MALDI-TOF-MS and MS/MS is a potentially important biomarker discovery method in cancer research. Compared with protein and peptide detection in the SDS-PAGE and IHC methods, MALDI-TOF-MS and MS/MS show a rapid and accurate identification of peptide masses with very simple preparation steps. Further, analysis can be widely performed on the majority of cancer sample types, such as urine or reserved tissues (Bryan et al., 2011, Gustafsson et al., 2011b). Using MALDI-TOF-MS spectra acquired from 66 colorectal cancer serum samples, researchers were able to determine colorectal cancer patients from control groups with 95.2% sensitivity and 90.0% specificity (de Noo et al., 2006). In a gastric cancer study, fibrinopeptide A peptides were identified using MALDI-TOF-MS/MS and abundance of fibrinopeptide A showed significant differences between gastric cancer, high risk and normal tissue samples (Ebert et al., 2006). Therefore, MALDI-TOF-MS and MS/MS present as suitable methods for the rapid profiling and identification of proteins and peptides from vulvar cancer patient samples.

1.2.2 MALDI-Mass Spectrometry Imaging

MALDI-TOF-MS based peptide profiling proved to be a powerful high-throughput analytical tool in cancer and other life sciences research. It allows for the rapid label-free multiplexed detection of different molecules, such as lipids, peptides, proteins and their modified bio-products (Merrill Jr et al., 2005). With increasing requirements for protein spatial information and advancement of analytical abilities, MALDI-TOF mass spectrometry imaging (MALDI-MSI) combines mass spectrometry with classic histological tissue analysis (Gustafsson et al., 2011a). MALDI-MSI is based on MALDI-TOF mass spectrometry and is able to visualize and identify peptides and other macromolecules on tissues without extra labelling (***Figure 1.5***). Compared to protein visualization based on immunohistochemical (IHC) staining, MALDI-MSI is a high throughput analytical approach that does not require previous knowledge of the sample genome and histology (Longuespée et al., 2014). Furthermore, MALDI-MSI can reveal spatial information of small molecules, such as therapeutic drug or metabolites, which cannot be directly labelled by antibodies or other chemical compounds (Marko-Varga et al., 2011). Although there are many MSI methods, including secondary ion mass spectrometry (SIMS), MALDI-MSI displays very good reproducibility in sample imaging and is able to couple with high mass resolution analysers, such as Orbitrap or Fourier-transform ion cyclotron resonance mass spectrometry (FT-ICR) systems, which make it a very useful tool in protein spatial analysis (Benabdellah et al., 2010).

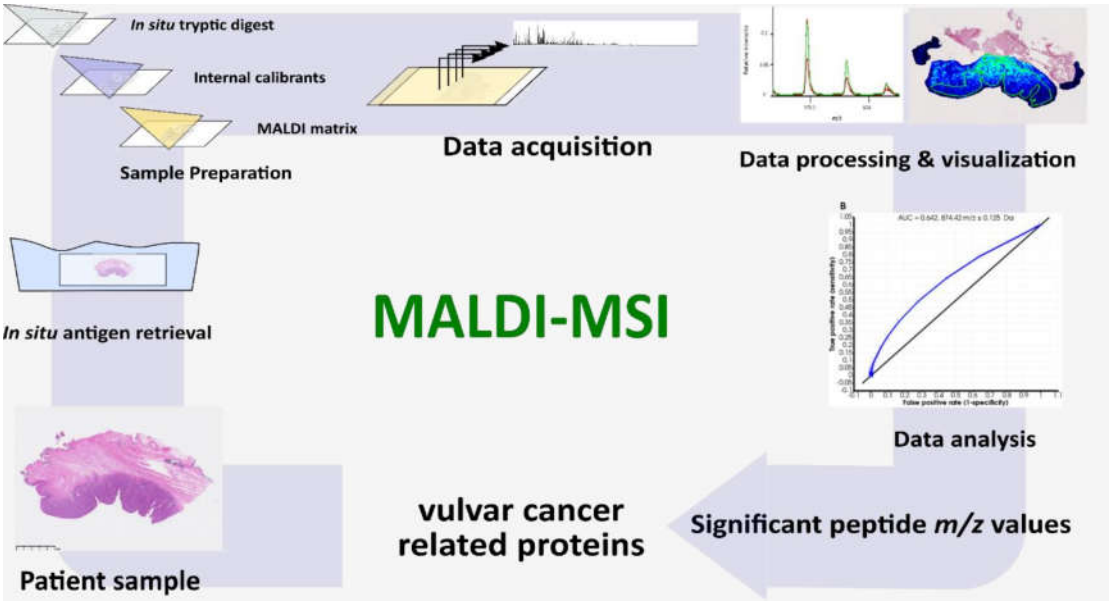


Figure 1.5: Workflow of peptide MALDI mass spectrometry imaging (MALDI-MSI) of vulvar cancer tissue sections. After co-registered with sample histological H&E stains, the relative ion abundance of peptides highlighted in different vulvar cancer tissue types revealed their potential relationship with vulvar cancer development.

MALDI-MSI has been used to study biological content from various samples ranging from bacteria colonies to whole mice bodies (Moree et al., 2012, Stoeckli et al., 2007). Previous experiments show that MALDI-MSI is able to highlight different morphological patterns of gastrointestinal, lung, brain and gynecological cancers (Franck et al., 2009). For example, a recent MALDI-MSI study of pancreatic cancer compared peptide images acquired from 90 adenocarcinoma and normal control samples, and digested peptides were successfully identified by MALDI-TOF-MS/MS sequencing (Goodwin et al., 2008a). The development of MALDI-TOF analysis on reserved patient tissues, including formalin-fixed, paraffin-embedded (FFPE) tissue samples, extends potential application of MALDI-MSI for cancer research. A series of new biomarkers have been identified using MALDI-MSI from 22 non-small cell lung cancer FFPE samples (Groseclose et al., 2008). The analytical power of MALDI-MSI can be improved when using tissue microarray (TMA) comprised of multiple different tissues (Song et al., 2010). In

gastric carcinoma peptide imaging, a recent study reported a 300 specimen TMA sample classification experiment using MALDI-MSI data, where 54 digested peptide peaks showed strong discriminative power in the identification of tumour samples (Morita et al., 2010). Based on the MALDI-MSI results, 14 proteins, including histone H4, were verified by immunohistochemical experiments as gastric cancer-specific markers. Therefore, MALDI-MSI is a suitable tool to study different morphological characters and inner biological changes of vulvar cancer.

Previously, no MALDI-MSI analyses have been performed with the aim of characterizing vulvar cancer. Additionally, to date, the only MALDI-TOF-MS studies that have been performed on vulvar cancer have been limited to the HPV-dependent type, rather than targeting on high-risk HPV-independent vulvar cancer (Trietsch et al., 2014, Kriegsmann et al., 2017). In chapter 3, MALDI-MSI was applied to characterize the proteome of benign vulvar skin, dVIN and VSCC from 12 HPV-independent vulvar cancer tissue samples, which highlighted cytokeratin 5 as a potential biomarker to discriminate VSCC from dVIN and other lesions (Zhang et al., 2016a). The following thesis chapters describe the method and results of applying MALDI-MSI to detect proteomic changes between LNM positive and LNM negative VSCC samples, and the examination of unique Indigenous Australian vulvar cancer tissue samples.

1.2.3 *nano*-LC-ESI-MS/MS

Electrospray (ESI), which ionizes in-solution proteins or other analytes via a high voltage electric field, is another frequently used type of MS ion source other than MALDI. The ESI source is frequently coupled with a liquid chromatography system which uses capillary action to separate complex biological samples before MS measurement. *Nano*-ESI-liquid-chromatography-mass-spectrometry (*nano*-LC-ESI-MS/MS) has been widely performed for high throughput profiling and quantification of different biological samples (Gustafsson et al., 2011b). Performing *nano*-LC ESI-MS/MS analysis on cancer samples parallel to MALDI-MSI significantly improves the efficacy of biomarker discovery. Although protein identification can be achieved by MALDI-TOF-MS/MS sequencing, the outcomes of MALDI-TOF-MS/MS based protein identification are subject to sample complexity and technological barriers that require pre-treatment and optimization (Ebert et al., 2006). Protein or peptide-based sequencing using *nano*-LC-ESI-MS/MS is a global protein identification method that does not require previous knowledge of a sample. Highly abundant digested peptides can be relatively easily identified and sequenced using MALDI-TOF-MS/MS, however, the identification by sequencing of lower abundant peptides is difficult (Balluff et al., 2011b). Thus, performing HPLC separation before MS measurement is a significant advantage of *nano*-LC-ESI-MS/MS, which offers better protein coverage and more comprehensive sample profiling. Another advantage of *nano*-LC-ESI-MS/MS are the built-in high mass resolution analyzers (over 50000 resolution), such as the Orbitrap and quadrupole-TOF, which have significantly high sensitivity and mass accuracy of digested peptides through LC separation (Krauss et al., 2010). In the vulvar cancer study, the application of *nano*-LC-ESI-MS/MS using qTOF and Orbitrap MS instruments could identify the intact peptide masses that detected in the MALDI-MSI data acquisition.

Identifying a large number of proteins or digested peptides using *nano*-LC-ESI-MS/MS has been shown to be possible from a wide range of various biological samples, including biopsies tissue, serum, plasma, urine and cell-lines, in fresh frozen and FFPE forms (Rose et al., 2012, Li and Tse, 2010, Matuszewski et al., 1998). *Nano*-LC-ESI-MS/MS is a very effective tool in cancer-associated protein discovery. More than 250 secreted proteins have been identified from the breast

isogenic cell line, which covered over 90% of proteins recorded in previous transcriptome databases (Mbeunkui et al., 2007). LC-MS based plasma profiling also revealed that CAGE1, SPAT9 and TEX28 expression levels in the patient blood are closely related to lung squamous carcinoma progression stages (Liu et al., 2017b). *Nano*-LC-ESI-MS/MS has also shown advantages in detecting post translational modification and small metabolites in patient samples. Metabolic profiling of prostate and cancers showed that patient FFPE samples can be used for untargeted metabolite identification and quantification in the LC-MS/MS analysis (Cacciatore et al., 2017, Armitage and Ciborowski, 2017). The LC system of *nano*-LC-ESI-MS/MS cannot only separate complex proteins or peptide contents with similar masses, but is also able to separate different protein structures or isomers in special columns (Wojakowska et al., 2013). Recently, in-depth glycome profiling of colorectal cancer and benign regions reported that more than 91 N-glycans demonstrated different isomers between two tissue types, especially EGFR-specific N-glycosylation for which abundance can be used to classify different colorectal cancer stages (Sethi et al., 2015). There are many other post translational modifications of cancer proteins, including phosphorylation and methylation, that have been profiled and studied using *nano*-LC-ESI-MS/MS analysis (Xiao et al., 2014, Wu et al., 2006, Ko et al., 2016).

1.2.4 Label-free quantification in *nano*-LC-ESI-MS/MS

Current *nano*-LC-ESI-MS/MS data acquisition has extended from simple protein identification to accurate and reproducible quantification. There are many different LC-MS protein quantification strategies developed for cancer proteome studies. Isotopic labelling of target analytes is an accurate protein quantification approach used in LC-MS. Labels can be introduced through cell metabolic labelling (SILAC), chemical tags (ICAT, iTRAQ, dimethyl and TMT) or using modified enzyme labelling ($O18^8$ label) (Ong et al., 2003, Mann, 2006, Smith et al., 2002). However, the high cost of consumable isotopic labels and other chemicals inhibit the application of labelling methods in large-scale protein quantification experiments. Unlike isotopic labelling methods, an increasing number of label-free quantification (LFQ) methods have been developed for cancer sample LC-MS/MS analysis (***Figure 1.6***). Quantification of analytes, such as protein, can be achieved by counting the intensities of eluted analytes or numbers of targeted fragmentation ion spectra in MS/MS analysis (Ng et al., 2017). Label-free quantification via *nano*-LC-ESI-MS/MS has many advantages; the methodology is able to identify a large number of proteins from samples with high protein sequence coverage, dynamic range of quantifiable samples with good reproducibility, the sample preparation and data acquisition consume less time and effort, as compared with methods that use labels, which frequently introduce bias to selective labelling strategies (Asara et al., 2008).

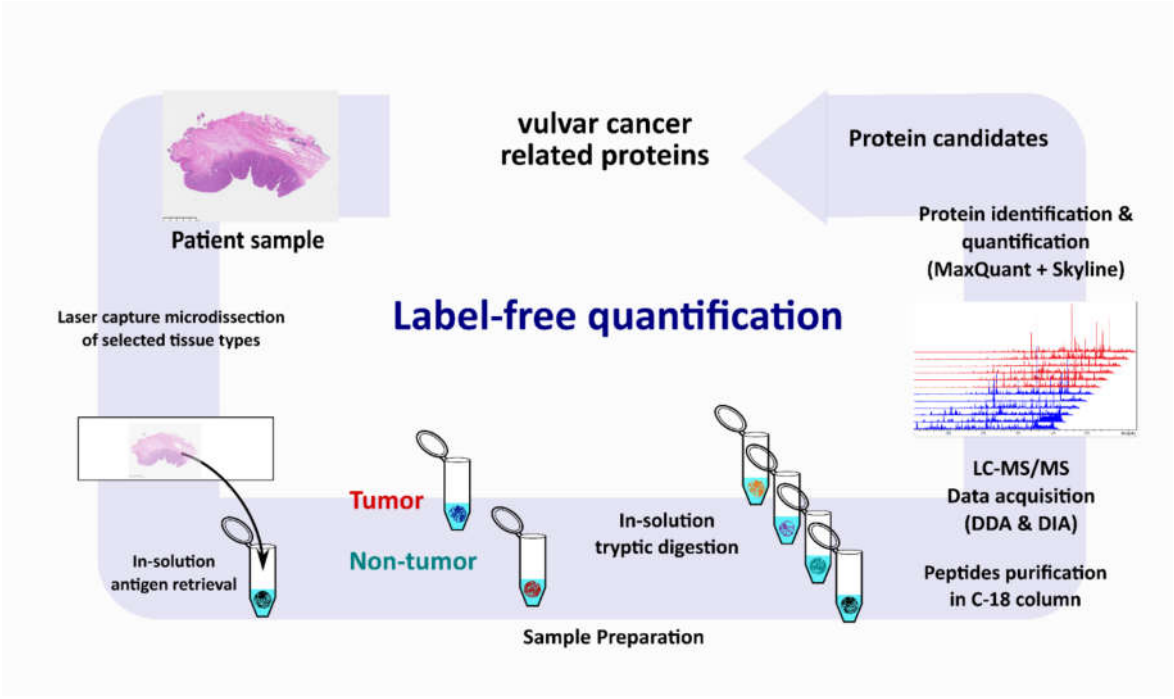


Figure 1.6: Workflow of vulvar cancer peptide sample analysis via label-free quantification approach in *nano*-LC-ESI-MS/MS. Regions of interests on vulvar cancer tissues were separated by the laser capture microdissection and vulvar cancer related protein label-free quantification analysis were achieved by software, such as MaxQuant and Skyline.

Based on different data acquisition modes used by *nano*-LC-ESI-MS/MS instruments, current LFQ techniques can be categorized into data-dependent acquisition (DDA) and data-independent acquisition (DIA) (Ng et al., 2017). Both have unique advantages and disadvantages in protein LFQ analysis. Previously, DDA based LFQ has been achieved in a majority of MS instruments, including Orbitraps and qTOFs (Välikangas et al., 2017, Lombard-Banek et al., 2016). In the DDA mode, only those top-ranked peaks from high abundant peptides or other analyte precursor ions in the MS scan are selected and sent for fragmentation, whereas the rest eluted peptide signals are arbitrarily ignored due to the limited instrumental analytical speed (Cosme et al., 2016). DDA based LFQ is an untargeted protein quantification method which is able to identify and relatively quantify a large number of proteins in a single analysis. With the development of analytical software packages, such as MaxQuant, DDA based LFQ has become a powerful tool for studying cancer proteome. For example, DDA based LFQ analysis of the GALNT3 knockdown ovarian

cancer cell line highlighted glycoprotein expression changes in comparison with wild-type cell lines in MaxQuant (Sheta et al., 2016). However, the efficiency of DDA based LFQ on complex sample analysis is challenged by the analytical speed of MS instrument, it is difficult to determine whether the absence of peptide detection is genuine or due to undersampling. Therefore, the detection and quantification of low abundance peptides in DDA based LFQ experiments sometimes may show lower levels of reproducibility (Goeminne et al., 2016).

To improve the LFQ efficiency for low abundant proteins, DIA based LFQ was developed as a targeted protein quantification approach using *nano*-LC-ESI-MS/MS (Röst et al., 2014). As a new next-generation targeted proteomics method, DIA based LFQ quantification merges the advantages of untargeted data acquisition methods and classic targeted approaches during the processes of data extraction and identification. DIA can be applied to a number of different MS instruments (Gillet et al., 2012). For example, each DIA scan cycle across the entire mass range in the Impact HD qTOF mass spectrometry from Bruker could finish within 2.6 seconds. Compared with the DDA method, DIA based LFQ transfers all precursor ions to fragmentation without selection, and different peptide MS/MS scans are convoluted with coeluting peptides. DIA based LFQ avoids the sampling biases encountered with traditional DDA LC-MS/MS and has achieved higher proteome coverage and reproducibility in LFQ analysis results of Skyline and other software (Bernhardt et al., 2012). Acquired DIA datasets can also be re-mined for the new protein targets analysis after the data acquisition, which increases efficiency and reproducibility of the technique. For example, the DIA method has shown a more comprehensive and reproducible investigation of insulin signalling target kinases (Parker et al., 2015). As DIA fragmentation ion spectra contain multiple peptide precursor ions, a pre-built spectral library acquired in DDA mode is required for spectral deconvolution and matching to target proteins (Gillet et al., 2012).

The application of *nano*-LC-ESI-MS/MS to label-free quantification (LFQ) analysis has achieved significant outcomes for studying tumour development, biomarker and therapeutic drug discovery of many cancers, including lung, colorectal, prostate, ovarian, breast and cervical cancers (Zhu et al., 2009, Alluri et al., 2016, Ramberg et al., 2017). For example, Serpin H1, Transgelin 2 and other less-known vascular biomarkers were found to have different abundances between malignant breast tumour and healthy controls in an LFQ comparison (Hill et al., 2011). LFQ also displays

powerful analytical abilities in high-throughput biomarker identification and untargeted quantification of rare brain cancer types, like glioblastoma (Heroux et al., 2014). Recently, some researchers have performed LC-MS analysis on vulvar cancer samples, but these studies were targeted toward well-known gynecological cancer biomarkers using labelling methods and were limited to HPV-dependent type vulvar cancer (Holthoff et al., 2016, Sandberg et al., 2012, Sandberg et al.). No LFQ analyses targeting high-risk HPV-independent vulvar cancer have been performed as yet, even though the identification of vulvar cancer-associated proteins could potentially have an impact on diagnosis and therapeutic drug development. Furthermore, the molecular mechanism of HPV-independent oncogenesis from the precursor lesions dVIN, and then metastasis from the primary tumour to adjacent lymph nodes, are still unknown. By combining MALDI-MSI and *nano*-LC-ESI-MS/MS platforms, a better understanding of HPV-independent type vulvar cancer can be achieved.

1.2.5 Tissue samples for mass spectrometry analysis

In current cancer MS studies, a tiny tissue section between 5 and 40μm thickness or a small amount of digested protein solution in the range of 0.5 to 5μg may provide a sufficient amount material for the analysis of tryptic peptides in the MS. Compared with other proteomics based approaches, such as IHC or protein microarrays, MS analysis shows significantly higher throughput for studying proteome profiles, which may promote the discovery of vulvar cancer biomarkers or drug targets (Lemaire et al., 2007). Present vulvar cancer samples are mainly reserved in two forms: fresh-frozen in liquid nitrogen or formalin fixed-paraffin embedded (FFPE) at room temperature. Both sample forms require a series of preparation steps before MS analysis, but vulvar cancer samples preserved as FFPE tissues are more readily available due to their routine collection in daily clinical practice.

1.2.5.1 Fresh-frozen tissue samples for MS study

Fresh frozen tissue samples have been primarily used in MALDI-MSI and *nano*-LC-ESI-MS/MS cancer proteomic or metabolic studies (Mirnezami et al., 2014, Rohner et al., 2005). Ideally, fresh frozen tissues should be used given they have rich biomolecule components, especially for small metabolites and signalling peptides, which are easily lost in preservation processes (Yuan et al., 2012, Smolders et al., 2013, Tamara et al., 2016). Sample preserved conditions, including frozen temperature and time, and potential contaminations, such as polymers or lipids, are two determining factors of using the fresh frozen sample for MS based proteomics analysis. Currently, most fresh tissues are rapidly frozen after surgery and stored at -80°C before MS experiments. Although proteomic characterization of fresh frozen samples generally represents *in-vivo* tissue conditions, there is a considerable decrease in protein or peptide quality of frozen samples if left untreated or allowed to thaw at room temperature (Goodwin et al., 2008a). Along with sample self-digestion, several extra chemical modifications, such as oxidation and loss of phosphorylation, may occur to proteins within fresh frozen samples, which leads to significant loss or change in the original proteomic information (Goodwin et al., 2008b, Ozeki et al., 2011). Another disadvantage of using fresh frozen samples for cancer MS-based proteomic research is that water-soluble compounds, including lipids and polysaccharide, remain in the sample during preparation. Remaining lipids could dramatically suppress peptide ionization and generate strong background noise (Goodwin, 2012). Increasing washing times in chemical solutions is an effective way to remove lipids from samples, but sample washing may lead to loss of soluble protein targets (Zavalin et al., 2015).

1.2.5.2 The FFPE samples for MS study

Although fresh frozen samples are suitable for vulvar cancer proteomic study, the lack of accessible frozen samples makes it difficult to perform MS analysis on fresh frozen vulvar cancer tissues. However, a majority of cancer patient samples are stored as FFPE tissue, with a considerable number of samples available with detailed pathology reports. Development of preserved sample recovery methods has resulted in the wide application of MS-based proteomic characterization of FFPE tissues (Kokkat et al., 2013, Wisztorski et al., 2010). A previous study reported that FFPE samples provided enough proteomic information for quantitative analysis, including a wide variety of cancer-associated factors such as p53, Cyclin D1, cyclin E, and cyclin-dependent kinases (CDKs) (Ikeda et al., 1998). FFPE samples have great potential for use in MS-based cancer studies, but the direct usage of FFPE tissues is challenged by cross-linked proteins or peptides in formalin, as amino acids have a strong reaction with formaldehyde and create a methylene bridge for inter or intra protein cross-linking (Magdeldin and Yamamoto, 2012). To recover proteins or peptides from cross-linking, heat-induced antigen retrieval (HIAR) in citric acid followed by proteolytic digestion have been applied to overcome peptide cross-linking within FFPE samples (Gustafsson et al., 2010). For example, recent studies show that tryptic digestion with modified matrices improved peptide signal quality and sequence coverage in MALDI-TOF-MS/MS-based cancer peptide identification and profiling (Groseclose et al., 2008, Berg et al., 2010, Fowler et al., 2011). Development of other fixing protocols, such as ethanol fixation, present alternative choices for long-term sample preservation without protein cross-linking (Balluff et al., 2011b, Masir et al., 2012, Kap et al., 2011). These non-toxic and non-crosslinked sample preservation methods potentially extend the viability of analytical samples in cancer MS research.

Vulvar cancer, especially HPV-independent type vulvar cancer, is a rare disease with very few samples available for MS analysis. Thus, patient FFPE samples archived previously in clinical practice were used in this MALDI-MSI and *nano*-LC-ESI-MS/MS-based vulvar cancer proteomic characterization studies.

PERSPECTIVE

MALDI-MSI peptide imaging and *nano*-LC-ESI-MS/MS LFQ have been widely performed on tissue proteomic characterization and patient classification in gynecological cancers. In the field of vulvar cancer, most MS based proteomics analysis was performed on studying the HPV-dependent vulvar cancer rather than the higher metastatic risk HPV-independent vulvar cancer. Therefore, the primary aim of this study was to perform MALDI-MSI analysis to characterize the recognizable tissues types involved in the HPV-independent vulvar cancer development, including benign epithelium, dVIN, VSCC and LNM, using FFPE patient samples. The following aim was to classify primary tumours with and without LNM in VSCC TMAs using MALDI-MSI data and identified potential protein predictors for the metastatic vulvar cancer. This study also tried to reveal the differential protein abundances between Indigenous and non-Indigenous vulvar cancers using *nano*-LC-ESI-MS/MS LFQ and highlighted signalling pathways potentially involved with high Indigenous vulvar cancer prevalence. Proteins of interest identified in this study could be used as molecular evidence indicating different HPV-independent vulvar cancer progression stages and explain the abnormally high incidence rate amongst Indigenous Australian women.

HYPOTHESIS

The hypothesis of this thesis is that MALDI mass spectrometry imaging (MALDI-MSI) and *nano*-LC-ESI-MS/MS label-free quantification (LFQ) could display different protein abundances of different HPV-independent vulvar tissue types that reflected in the morphology, primary tumours with and without lymph node metastasis, and Indigenous and non-Indigenous vulvar cancers which have not been described in previous vulvar cancer studies. Most importantly, highlighted proteins of interests identified in this MS analysis yield complementary clinical information to distinguish pre-cancerous lesion, vulvar squamous cell carcinoma with and without metastasis, and could answer the clinical question of high Indigenous vulvar cancer prevalence. The aims of this study were:

AIMS:

I. To characterize proteomic patterns of different vulvar cancer histopathological tissue types, including, benign epithelium, dVIN, VSCC and LNM through peptide MALDI-MSI and to validate proteins with differential abundance in vulvar tissues by IHC staining and DIA based LFQ.

II. To classify HPV-independent VSCC with and without LNM from the primary tumour tissue microarrays using MALDI-MSI data and validate potential protein predictors by IHC staining and DDA based LFQ.

III. To detect differential protein abundances between four Indigenous and four non-Indigenous VSCC samples using DDA based LFQ. The signalling pathways involved with proteins of interest are highlighted in the Reactome pathway analysis and protein spatial distributions were validated by MALDI-MSI and IHC staining.

CHAPTER 2: MATERIALS AND METHODS

2.1 BASIC MATERIALS, REAGENTS AND INSTRUMENTS

All experimental materials, reagents and instruments used in this thesis are listed in the following tables below.

2.1.1 Solvents and chemicals

Chemical or Solvent	*Manufacturers*	*Product numbers*
Acetonitrile (ACN) - HPLC grade	Merck, Darmstadt, Germany	1000302500
Ammonium hydrogen bicarbonate (NH_4HCO_3)	Merck, Poole, England	103025E-1 KG
Citric acid monohydrate	Sigma-Aldrich, Japan	C0706-500 G
α-Cyano-4-hydroxycinnamic acid (CHCA)	Bruker Daltonics, Bremen, Germany	201344
1,4-Dithiothreitol (DTT)	Sigma, Steinheim, Germany	10708984001-10 G
EDTA	Sigma, Steinheim, Germany	431788-100 G
Eosin	Sigma-Aldrich, Steinheim, Germany	HT110332-1 L
Ethanol (EtOH)-HPLC grade	Merck, Victoria, Australia	1117272500
Ethanol (EtOH)-AR grade	VWR, Le Périgares, France	20821330
Formic acid (FA)	Sigma-Aldrich, Steinheim, Germany	21909098
Haematoxylin	Sigma-Aldrich, Steinheim, Germany	GH5232-1 L

Hydrogen peroxide	Merck, Darmstadt, Germany	1072090250
2-Iodoacetamide (IAA)	Sigma-Aldrich, Steinheim, Germany	I6125-25 G
Isopropanol-HPLC grade	Merck, Darmstadt, Germany	1096342500
Isopropanol-AR grade	Merck, Darmstadt, Germany	8187661000
Methanol (MeOH)-HPLC grade	Merck, Darmstadt, Germany	1060182500
Phosphate-buffered saline (PBS)	CFU, Adelaide, Austalia	INV10118-500
Pertex mounting medium	Meditate Midzintechnic, Burgdorf, Germany	PEP40000
Sodium hydroxide (NaOH)	Merck, Darmstadt, Germany	1064980500
Trifluoroacetic acid (TFA)	Merck, Darmstadt, Germany	1082620100
Urea	Merck, Darmstadt, Germany	1084870500
Xylene	Chem-Supply, Gillman, South Australia	XA003-2.5 L

2.1.2 Consumable materials

Consumables	*Manufacturers*	*Product numbers*
C18 spin columns	Thermo-Fisher Scientific, Rockford, IL	89870
C18 ziptip	Thermo-Fisher Scientific, Rockford, IL	87784
Microcon-30kDa	Merck, Millipore, Darmstadt, Germany	MRCF0R030
Stainless steel cryotome blades (35×80 mm)	Feather Safety Razor Co., Japan	S35
ImagePrep spray generators	Bruker Daltonics, Bremen, Germany	261614
Indium Tin Oxide coated glass slides	Bruker Daltonics, Bremen, Germany	237001
PEN membrane slides	Thermo-Fisher Scientific, Rockford, IL, USA	LCM0522
Safe-Lock polypropylene tubes (0.5/1.5/2.0 mL)	Eppendorf, Hamburg, Germany	022363611/022363204/022363352
Reloaded Tips (10/200/1000 µL)	Eppendorf, Hamburg, Germany	0030073207/0030000870/0030000919
Superfrost slide	Thermo-Fisher Scientific, Rockford, IL	SF41296SP
ImagePrep solution vials	Bruker Daltonics, Bremen, Germany	N/A
HPLC vials	Thermo-Fisher Scientific, Rockford, IL	03-PPSV
Microscope coverslips (24×40/50/60 mm)	Waldemar Knittel Glasbearbeitungs, Germany	VD1 2240 Y100A (40/50/60 mm)

2.1.3 Proteolytic enzymes

Enzyme	*Manufacturers*	*Product Numbers*	*Working Concentration*
Sequencing Trypsin	Promega, Madison, USA	V5111	100 ng/µL
Trypsin Gold	Promega, Madison, USA	V5280	500 ng/µL

2.1.4 Peptide calibration standard

Standard	*Sequence*	*Manufacturers*	*Product numbers*	*Working Concentration*
ACTH (1-24)	SYSMEHFRWG KPVGKKRRPVK VYP	Sigma-Aldrich, St. Louis, MO	A0298	2 pmol/µL
ACTH (18-39)	RPVKVYPNGAE DESAEAFPLEF	Bruker Daltonics, Bremen, Germany	225593	2 pmol/µL
Angiotensin I (34-43)	DRVYIHPFHL	Auspep, Victoria, Australia	C10-00002	0.4 pmol/µL
Dynorphin A (1-17)	YGGFLRRIRPKL KWDNQ	Auspep, Victoria, Australia	2032	2 pmol/µL
[Glu]-Fibrinopeptide B	EGVNDNEEGFF SAR	Auspep, Victoria, Australia	F3261	0.4 pmol/µL
Peptide calibration standard II	Proprietary peptide mixture	Bruker Daltonics, Bremen, Germany	222570	1 µL

2.1.5 Matrices

Matrix	*Manufacturers*	*Product numbers*	*Working Concentration*	*Dissolved solvent*
α-Cyano-4hydroxycinnamic acid (CHCA)	Bruker Daltonics, Bremen, Germany	201344	7 g/L	50% ACN in water, 0.2% TFA
			10 g/L	70% ACN in water, 0.1% TFA

2.1.6 Primary, secondary antibodies and IHC reagents

Primary antibodies	*Manufacturers*	*Product Numbers*	*Working dilution*
Alpha-Actin	Abcam, MA, USA	ab5694	1/500
Alpha-actinin 4	Sigma-Aldrich, St. Louis, MO	a5044	1/1000
Beta-Actin	Abcam, MA, USA	ab8226	1/500
Annexin-A1	BD Science, CA, USA	610066	1/5000
Annexin-A2	BD Science, CA, USA	610069	1/500
Plectin	Abcam, MA, USA	ab32528	1/250
Collagen Type I	Abcam,MA,USA	ab34710	1/500
Cytokeratin 1/10	Merck, Millipore, Germany	CBL266	1/100

Cytokeratin 5	Abcam, MA, USA	ab52635	1/200
Cytokeratin 5/6	Dako, Australia	790-4554	1/500
Vimentin	Abcam, MA, USA	ab8978	1/200
NDRG1	Abcam, MA, USA	ab63989	1/400
Annexin-A8	Thermo Fisher, Scientific, Rockford, IL	PA5-50196	1/100
NARS	Abcam, MA, USA	ab197326	1/100
RAP1A	Thermo Fisher, Scientific, Rockford, IL	MA1-147	1/500
POSTN	Abcam, MA, USA	ab14041	1/1000
Goat Anti-Rabbit	Dako, Australia	P0448	1/500
Goat Anti-Mouse	Dako, Australia	P0447	1/500
Streptavidin	Dako, Australia	F0422	1/500
Goat serum	Thermo Fisher, Scientific, Rockford, IL	16210064	5% in PBS

2.1.7 Instruments

Instrument	*Manufacturers*
Ultimate 3000 RSLC system	Agilent Technologies, Japan
Centrifuge 5810R	Eppendorf, Hamburg, Germany
ImagePrep station	Bruker Daltonics, Bremen, Germany
Leica AS LCM microscope	Leica Instruments, Germany
Milli-Q® Reference Water Purification	Merck, Darmstadt, Germany
MTP 384-800um AnchorChip target	Bruker Daltonics, Bremen, Germany
MTP slide adaptor II	Bruker Daltonics, Bremen, Germany
NanoZoomer automated microscope	Hamamatsu, Beijing, China
LTQ XL Orbitrap	Thermo-Fisher Scientific, Rockford, IL
Dry block heater	Ratek, VIC, Australia
Impact HD qTOF	Bruker Daltonics, Bremen, Germany
Microwave	LG, Whirlpool, VIC, Australia
Canon, CanoScan5600F, flatbed scanner	Canon, Japan
Ultraflex III MALDI-TOF/TOF	Bruker Daltonics, Bremen, Germany

Ultraflextreme MALDI-TOF/TOF	Bruker Daltonics, Bremen, Germany
Autoflex speed MALDI-TOF/TOF	Bruker Daltonics, Bremen, Germany
SpeedVac SVC100	Savant Instruments, Farmingdale, NY
Ultra-low -86^{0}C freezer	SANYO, Osaka, Japan

2.1.8 Software

Program	*Manufacturers*	*Versions*
flexAnalysis	Bruker Daltonics, Bremen, Germany	4.0
flexControl	Bruker Daltonics, Bremen, Germany	3.4
flexImaging	Bruker Daltonics, Bremen, Germany	4.0
BioTools	Bruker Daltonics, Bremen, Germany	3.1
Mascot	Matrix Science, Boston, MA, USA	2.3.02
Proteome Discoverer	Thermo-Fisher Scientific, Rockford, IL	1.4
DataAnalysis	Bruker Daltonics, Bremen, Germany	4.2
R	R Foundation for Statistical Computing	3.0.1
SCiLS Lab	GmbH, Breneb, Germany	2016b

SkylineMS	Washington, USA	3.4
MaxQuant	Max Planck institute, Germany	1.6.0.1
Perseus	Max Planck Institute, Germany	1.6.0.2
ImageJ	NIH, USA	1.50i
GraphPad	La Jolla, CA, USA	6.0
NDP View 2	Hamamatsu, Beijing, China	2.3.1
Xcalibur	Thermo-Fisher Scientific, Rockford, IL	2.2
Reactome	Reactome teams	V62

2.2 METHODS

All experimental methods performed in this thesis, including both data acquisition and data analysis methods, are described in the following sections.

2.2.1 Sample preparation methods

Protein and peptide sample preparation methods performed in this thesis are described below.

2.2.1.1 Tissue samples for experiments

The vulvar cancer formalin fixed paraffin embedded (FFPE) tissue blocks obtained from HPV-independent patients underwent surgery at Royal Adelaide Hospital (Adelaide, South Australia) and Burnside War Memorial Hospital (Adelaide, South Australia). Human sample studies were approved by the ethics committee of Royal Adelaide Hospital (Approval No.120908). The histopathological report, including cancer stages, associated lesions, HPV infection, and other related clinical information, of all vulvar cancer patients collected in this study, were listed in ***Supplementary S2.1***. Sample tissues were first embedded in OCT after surgery, then fixed in formaldehyde solution overnight, dehydrated in ethanol and embedded in paraffin blocks. All FFPE blocks contained one or more vulvar tissues types, including benign epithelium, Lichen Sclerosis (LS), differentiated Vulvar Intraepithelial Neoplasia (dVIN), vulvar squamous cell carcinoma (VSCC), lymph node metastasis (LNM) or connective tissue. All sample slides were sectioned at 6μm and applied hematoxylin and eosin (H&E) stains. Different morphological patterns of vulvar tissues on sample slides were identified and circled by experienced pathologists.

2.2.1.2 Tissue microarray construction (TMA) for chapter 4 experiments

Vulvar squamous cell carcinoma tissue microarrays were constructed in the department of SA Health (Adelaide, South Australia). In short, all selective VSCC tissue blocks were punctured one or two 1.5mm diameter sampling cores and assembled into two 60 core TMAs. In total, 28 VSCC primary tumour blocks, including 16 LNM positive samples and 12 LNM negative samples, from the cohort of 42 HPV-independent patients were incorporated in two TMA blocks, named TMAI and TMAII (***Supplementary S2.2***). Sampling cores from benign vulvar epithelium and necrotic tumours blocks were also included as control samples. After TMA construction, morphological patterns of all TMA sampling cores were examined by experienced pathologists. Consecutive TMA sections were cut at 6μm thickness and mounted onto indium titanium oxide (ITO) coated slides, polyethylene naphthalate (PEN) membrane slides, and Superfrost glass slides for the Chapter 4 experiments.

2.2.1.3 Experimental tissue sample preparation and heat-induced antigen retrieval (HIAR)

Vulvar cancer tissue FFPE samples for MALDI-MSI were sectioned at 6μm thickness and mounted onto ITO coated slides, then dried at room temperature overnight and stored at 4°C. Sample slides were subjected to heat-induced antigen retrieval (HIAR) as described previously (Gustafsson et al., 2010). In brief, slides were placed on a heating block (60°C) for an hour, then washed in 100% xylene for 5mins twice, 100% Ethanol for 2mins twice and 10mM NH_4HCO_3 for 5mins twice. Afterwards, sample slides were immersed into 10mM citric acid buffer (pH = 6), rapidly heated to boiling point in a microwave and then simmered for 10mins. Slides were kept in citric acid buffer at 98°C for another 30mins. Slides were removed from the citric acid solution and cooled down at room temperature. Then, slides were washed in 10mM NH_4HCO_3 for 5mins twice. Air dried slides and applied to teach marks using water-based white-out. MALDI-MSI sample slides were scanned at negative film mode, 2400 dpi image resolution, in the flatbed Image scanner (Cannon. Japan).

2.2.1.4 TMA sample reduction and alkylation for chapter 4 experiments

TMA sample slides for chapter 4 experiment were first performed HIAR using the same method as described previously (Gustafsson et al., 2010). Afterwards, slides were washed in 10mM NH_4HCO_3 for 5mins twice, then air dried and applied 100µl 20mM DDT (dissolved in 5mM NH_4HCO_3) solution on the top of the section, heated up to 56°C on the heating block for 30mins. Removed DDT and washed the slides again in 10mM NH_4HCO_3 for 5mins twice. Dissolved 55mM IAA into 5mM NH_4HCO_3 and performed 100µL solution on the top of the section, then left the slides in darkness at room temperature for another 30mins. Washed away the IAA solution on the slides using 10mM NH_4HCO_3 twice and air dried sample slides. Applied slide teaches marks using water-based white-out. MALDI-MSI sample slides were scanned at negative film mode, 3200 dpi image resolution, in the flatbed Image scanner (Cannon. Japan).

2.2.1.5 *in-situ* tryptic digestion

Trypsin Gold (Promega, USA) was dissolved in 5mM NH_4HCO_3 to 500ng/µL working concentration on the ice and stored in 40µL aliquot forms at -80°C freezer. For the MALDI-MS experiment *in-situ* tryptic digestion, an aliquot of 40µL Trypsin Gold was further diluted into 160µL of 10mM NH_4HCO_3 and sprayed onto tissue sections using the default trypsin deposition method (1.30s spray time, 30s drying time, and 30 cycles) and 38% deposition power of the ImagePrep station (Bruker Daltonics, Germany). The spray power was fined tuned so that more than 30 cycles could be deposited. After 15 spray cycles, the direction of sample slides was reversed and then finished the rest trypsin spray. This was done to assure a homogenous distribution of trypsin on all tissue sections within the slide. Samples were digested for 2 hours at 37 °C for in a homemade humid box.

2.2.1.6 Internal calibrants application

The internal calibrant standard peptides, including Angiotensin I, [Glu]-Fibrinopeptide B, ACTH (1-24), and Dynorphin-A, diluted into 100µL working solution (section 2.1.4) and sprayed on the tissue after 2 hours tryptic digestion. The internal calibrants standard was sprayed using the default trypsin deposition method (1.30 s spray time and 30 s drying time) and 38% deposition power of the ImagePrep station (Bruker Daltonics, Germany).

2.2.1.7 Matrix deposition

The α-cyano-4-hydroxycinnamic acid (CHCA) matrix used for the MALDI-MSI experiments were performed at the 7g/L working concentration and roughly 4mL matrix working solution which deposited by the ImagePrep station (Bruker Daltonics, Germany) according to the internal CHCA matrix deposition method of the instrument (20% deposition power and 30% modulation as the instrumental spray set).

2.2.1.8 Laser capture microdissection (LCM)

Vulvar tissue sections mounted on the PEN membrane slides (Thermo-fisher Scientific, USA) were first heated up to 60°C for an hour on the heating block, and then removed paraffin on the sample slides using 100% xylene for 90 s and 100% ethanol for 60 s. The selective tissue area on the sample was excised using a Leica AS LCM microscope (Leica Instruments, Germany) with the following parameters: 50 intensity, 1 speed, medium bridge, and 20 power setting. Isolated tissue sections were dropped into 80µL of 10mM citric acid buffer in the 0.5mL Eppendorf tube.

2.2.1.9 In-solution tryptic digestion, antigen retrieval and purification for the peptide library construction

Tissue samples collected in the 0.5mL tubes were spun down and filled up to 200µL of 10mM citric acid buffer. Sample tubes were incubated at 98°C on the heating block for 45 mins antigen retrieval and cooled down at the room temperate. Removed HIAR buffer and washed samples with 10mM NH_4HCO_3. Spun down samples again and re-filled with 100µL of 5mM NH_4HCO_3, then added 1µL Trypsin Gold (Promega, USA) working solution into each tube and incubated samples at 37°C overnight. Terminated trypsin digestion by adding 1µL 10% FA. Digested peptides were cleaned by C-18 columns (Thermo-Fisher Scientific, USA). In a short, wet resined the C-18 columns with 200µL 100% ACN for 5 mins twice; equilibrate resined columns with 200µL 5% ACN with 0.5%TFA for 5 mins twice; loaded sample solution into columns; centrifuged 5 mins and reloaded; washed columns with 200 µL 5% ACN with 0.5%TFA for 5 mins three times; Eluted digested peptides with 20µL 70% ACN three times. Transferred eluted peptides into HPLC vial and vacuum dried in the speed vacuum centrifuge (Savant Instrument, NY). Re-suspended digested peptides in 10µL 2%ACN with 0.1% FA and stored in the -80 ^{0}C freezer.

2.2.1.10 In-solution tryptic digestion, antigen retrieval and purification for peptide label-free quantitation (LFQ)

Tissue samples collected in the 0.5mL tubes were spun down and filled up to 200µL with 10mM citric acid buffer. Samples were incubated at a 98°C heating block for 45 mins antigen retrieval and cooled down at room temperate. Peptide sample preparation was based on published filter-aided sample preparation (FASP) protocol (Lipecka et al., 2016). In a short, HIAR treated samples were mixed with 7M Urea in 25mM NH_4HCO_3 in Microcon-30k filter tubes. Spun down samples and applied 100µL 20mM DTT (dissolved in 7M Urea) on the filter for 1h; Spun down again and added 100µL 55mM IAA (dissolved in 7M Urea). Washed the filter with 100µL 7 M Urea and centrifuged 15 minutes three times. Then, washed the filter again with 100µL of 50mM NH_4HCO_3 and centrifuged 15 minutes three times. Added 10µL 100ng/µL sequencing trypsin dissolved in fresh 10mM NH_4HCO_3 in the filter and incubated samples overnight in the 37°C room. Centrifuged the filter 10 mins, and added 50µL 10mM NH_4HCO_3 spun down again. Re-suspend samples in 50µL 1% FA spun down again to elute peptide samples from the filter. Transferred peptide elution into HPLC vial and vacuum dried in the speed vacuum centrifuge (Savant Instrument, NY). Peptide samples were reconstituted into 10µL solution using 2% ACN, 0.1% TFA and peptide sample concentration were measured in the Nanodrop2000 (Thermo-fisher Scientific, USA) with a UV light wavelength set of 205nm.

2.2.2 Data acquisition methods

Methods used for MALDI-MSI and *nano*-LC-ESI-MS/MS data acquisition in this thesis were described below.

2.2.2.1 MALDI-MSI for chapter 3 and 5 experiments

The sample spectra were acquired by the UltrafleXtreme MALDI-TOF-MS instrument (Bruker Daltonics, Germany) according to the published protocol (Gustafsson et al., 2011a). In brief, the detection regions on sample slides were circled in flexImaging (v4.0). The MALDI instrumental acquisition settings were: 1000Hz laser repetition rate; 45 to 55% laser power, 500 shots per position; 100μm acquisition resolution, 800-3500 *m/z* detection range; 800 Da mass deflection; 2.573 kV detector gain, 25 kV ion source voltage; 7.6 kV lens voltage; 29.5 kV reflector voltage. The acquisition progress was monitored by the flexControl (v3.4). Detected peak masses from acquired spectra were recalibrated with the Internal Calibrants standard and tryptic peptide masses in the flexAnalysis batch processing (v4.0) (Gustafsson et al., 2012).

2.2.2.2 MALDI-MSI for chapter 4 experiments

The sample spectra were acquired by the UltrafleXtreme MALDI-TOF-MS instrument (Bruker Daltonics, Germany) according to the published protocol (Gustafsson et al., 2011a). In brief, the detection regions on TMA slides were circled in the flexImaging (v4.0). The acquisition settings were: 2000 Hz laser repetition rate; 55 to 60% laser power, 500 shots per position; 60μm spatial acquisition resolution, 800-4000 *m/z* detection range; 750 Da mass deflection; 2.798 kV detector gain, 25 kV ion source voltage; 8.25 kV lens voltage; 26.38 kV reflector voltage. The acquisition progress was monitored by flexControl (v3.4). Detected peak masses from acquired spectra were recalibrated with the Internal Calibrants standard and tryptic peptide masses in the flexAnalysis batch processing (v4.0) (Gustafsson et al., 2012).

2.2.2.3 Raw spectra processing and peak picking

The raw spectra were automatically processing into peaklists in the flexAnalysis batch analysis (v4.0) based on the algorithms of SNAP peak picking, Gauss spectra smoothing, TopHat baseline subtraction, and a signal cutoff of over 2 Signal to Noise (S/N) ratio. The performance of MALD-MSI data acquisition was examined by plotting the detected signals of Internal Calibrants standard peaks with their detected regions on sample slides using the in-house R script (section 2.2.4.2). After re-annotated H&E stained sample images with the detected regions in the flexImaging (v4.0), tissue-specific regions of interest (ROIs) on the slides were circled and exported into XML format files for further bioinformatics analysis in section 2.2.3.1.

2.2.2.4 *in-situ* tandem mass spectrometry (MALDI-TOF/TOF-MS)

The peptide masses of interest were manually selected and input into the precursor ion window of flexControl (v3.4) for the fragmentation of selective precursor peaks. The *in-situ* MS/MS analysis was performed with an optimized method referred as LIFT (Bruker Daltonics, Germany) mode using the same tissue sections for MALDI-MSI (Neubert et al., 2004). Tandem MS spectra were calibrated with the default Immonium Ions list of the flexAnalysis (v4.0) and sent to BioTools (v3.1) for database searching. The result was searched in the in-house MASCOT database search engine with the following parameter sets: Taxonomy, Homo sapiens; MS tolerance, 100 ppm; MS/MS tolerance, 0.8Da; human database, Swiss-Prot (version 2016-02) and NCBI (version 20150401); enzyme, trypsin/P; Carbamidomethyl of cysteine as fixed modification; oxidation of methionine, N-terminal acetylation, and glutamine to pyroglutamate as variable modifications, missed cleavages set at 2; the significant threshold less than 0.05 and score of 20 were selected as the identification cutoff.

2.2.2.5 *nano*-LC-ESI-Orbitrap-MS/MS for the peptide library construction

The *nano*-LC-ESI-MS/MS was performed using an Ultimate 3000 RSLC system (Thermo-Fisher Scientific, USA) coupled to an LTQ XL Orbitrap mass spectrometer via an Electrospray source (Thermo-Fisher Scientific, USA). Peptide samples were pre-concentrated onto a C18 trapping column (Acclaim PepMap100 C18 75µm × 20 mm, ThermoFisher Scientific) at a flow rate of 5µL/min in 2% ACN 0.1% TFA for 10 min. Peptide separation was performed using a 75µm ID C18 column (Acclaim PepMap100 C18 75µm × 50 cm, Thermo-Fisher Scientific) at a flow rate of 0.2µL/min using a linear gradient from 5% to 45% B (A: 5% ACN 0.1% FA, B: 80% ACN 0.1% FA) over 130 min, followed by a 20 min wash with 90% B, and a 20 min equilibration with 5% A. MS scans were acquired in the mass range of 200 to 2200 *m*/*z* in a data-dependent fashion using default proteomics profiling method. Singly charged precursor ions were excluded from acquisition. Collision cell energy was determined by the *m*/*z* of the precursor ion. Acquired spectra were subjected to peak detection, deconvolution, and re-calibration, and preprocessed using Thermo Xcalibur (v2.2). Further data processing and bioinformatics analysis were in section 2.2.3.4 and 2.2.3.5.

2.2.2.6 *nano*-LC-ESI-qTOF-MS/MS for the peptide library construction and data dependent acquisition (DDA) based label-free quantification

The *nano*-LC-MS/MS was performed using an Ultimate 3000 RSLC system coupled to an Impact II™ QTOF mass spectrometer via an Advance CaptiveSpray source. Peptide samples were pre-concentrated onto a C18 trapping column (Acclaim PepMap100 C18 75µm × 20 mm, ThermoFisher Scientific) at a flow rate of 5µL/min in 2% ACN 0.1% TFA for 10 min. Peptide separation was performed using a 75µm ID C18 column (Acclaim PepMap100 C18 75µm × 50 cm, Thermo-Fisher Scientific) at a flow rate of 0.2µL/min using a linear gradient from 5% to 45% B (A: 5% ACN 0.1% FA, B: 80% ACN 0.1% FA) over 130 min, followed by a 20 min wash with 90% B, and a 20 min 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, deconvolution, and re-calibration according to a lock mass using DataAnalysis (v4.2). Further data processing and bioinformatics analysis were in section 2.2.3.5.

2.2.2.7 Data independent acquisition (DIA) based label-free quantification

Samples were prepared for data independent *nano*-LC-ESI-MS/MS by laser microdissection, citric acid antigen retrieval, and trypsin digestion as described in section xxx. Data acquisition was carried out 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 75µm × 20 mm, Thermo-Fisher Scientific) at a flow rate of 5µL/min in 2% ACN 0.1% TFA for 10 min. Peptide separation was performed using a 75µm ID C18 column (Acclaim PepMap100 C18 75µm × 50 cm, Thermo-Fisher Scientific) at a flow rate of 0.2µL/min using a linear gradient from 5% to 45% B (A: 5% ACN 0.1% FA, B: 80% ACN 0.1% FA) over 130 min, followed by a 20 min wash with 90% B, and a 20 min equilibration with 5% A. The *nano*-LC was performed as described above with the Impact II™ QTOF mass spectrometer set to acquire 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. Further data processing and bioinformatics analysis were in section 2.2.3.7.

2.2.2.8 Hematoxylin and eosin (H&E) staining

All sample slides, including MALDI-MSI and PEN membrane slides, were put into 70% EtOH for 1 min to remove matrix. Slides were dipped in deionized water and immersed into hematoxylin solution (dissolved 1 mg/mL in 100% ethanol) for the 30s, then tap water for 2-5 min and immersed into eosin (dissolved 1 mg/mL in 100% ethanol) for 2 mins. The slides were washed with 70%, 90%, and 100% EtOH and 100% Isopropanol, and 30s for each solution. Then they were transferred into 100% xylene for 1 min and into fresh 100% xylene for at least another minute. Coverslips were mounted with Pertex. The slides were dried at 37°C until they were ready for microscope scanning by the Nanozoomer Scanner (Hamamatsu, Beijing, China) with 40× magnitude and brightness background.

2.2.2.9 Immunohistochemistry (IHC) staining

For the analysis of target proteins by IHC, 6 µm tissue consecutive sections were placed on superfrost slides, 5 mins dewaxed in 100% xylene twice, rehydrated 2 mins in 100% ethanol twice, washed slides, and blocked slides with 100µL 10% H_2O_2 dissolved in 100mL 1X PBS. Washed slides again with 1X PBS 5 mins twice and subjected to microwave antigen retrieval in 10mM citric acid buffer (pH = 6) for 10 min at 100 °C in a steam microwave (Sixth Sense). Afterwards, slides were washed with 1X PBS 5 mins twice and incubated with 5% goat serum blocking buffer 30 mins. Primary antibodies were first diluted to working concentration (section 2.1.6) in 5% goat serum blocking buffer, then incubated overnight in the 4 °C incubation room. Washed away primary antibodies with 1X PBS biotinylated anti-rabbit or mouse secondary antibodies were diluted to the working concentration and performed on slides for 1-hour incubation. Removed secondary antibodies on slides and added working streptavidin-HRP solution on slides for 1-hour incubation. Washed slides with 1X PBS 5 mins twice, mixed Diaminobenzidine (DAB) with H_2O_2 and incubated with slides for 6 mins and washed away the reagent. Washed slides with 1X PBS 5 mins twice, stained slides in hematoxylin 10 s, washed slides with running water and dehydrating in 70 %, 100% ethanol, and 100% xylene. Mounted stained sample slides with cover slides by

Pertex glue. The IHC stained sections images were scanned using a NanoZommer (Hamamatsu, Beijing, China) with 40× magnitude and brightness background. Further positive stain data analysis was in section 2.2.3.8.

2.2.3 Data analysis methods

The analytical pipelines and software for acquired MS and other data are described below (***Figure 2.1***).

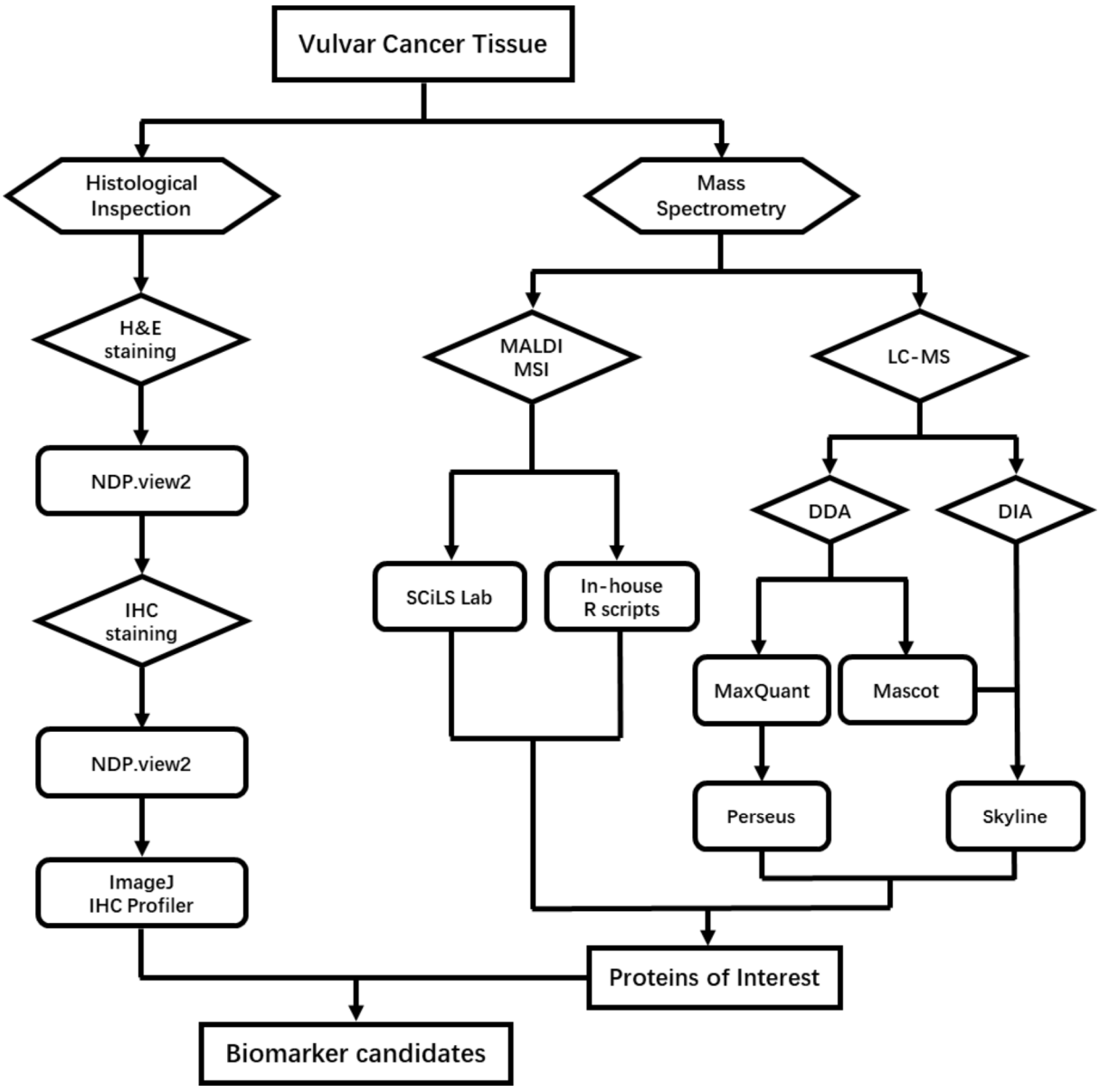


Figure 2.1: The flow chat of analytical software used in this vulvar cancer proteomics characterization study.

2.2.3.1 MALDI-MSI peaks grouping and statistical analysis

Peaklists from ROIs exported from the flexAnalysis (v4.0) were combined and Density-based spatial clustering of applications with noise (DBSCAN), with the radius of neighbourhood (eps) of 0.02 and a minimum density of 500 peaks was used to cluster peaks into peak groups based on their *m*/*z* values, the abundance weighted mean (AWM) of each peak group was calculated. Initially, only peak groups containing at least 10,000 peaks were considered for further analysis (section 2.2.4.3). This strict criterion was put in place to ensure that all peak groups analyzed could be reproducibly detected in many spectra. A Kruskal-Wallis test was performed on the tissue type specific S/N values to determine among different tissue types in chapter 3, 4, and 5, while Wilcoxon signed-rank test performed on TMA data of peak groups coming from LNM positive and LNM negative VSCC patients (section 2.2.4.4). All hypothesis test *p*-values were adjusted according to Benjamini-Hochberg (Benjamini, 2010). Potential peak groups of interests are performed boxplot and receiver operating characteristic curve (ROC) analysis of their intensity or S/N based in-house R script (section 2.2.4.6). Matching of Mascot and MaxQuant identified peptides with peak groups was also done by in-house R script (section 2.2.4.5). The number of peaks for matching threshold was set as minimum 100 spectra per AWMs and calculated mass from *nano*-LC-ESI-MS/MS seat within MALDI-MSI peakgroups mass range was believed as a positive match. The median log intensity in each of tissue type was calculated for each peak group and for each patient. Peak groups were heuristically ranked by the largest difference in median log intensity between tissue types.

2.2.3.2 Canonical correlation analysis and Linear discriminant analysis (CCA-LDA) patient classification

A detailed description of the CCA-LDA method using for patient classification was described in the published methods (Winderbaum et al., 2016). In summary, arbitrary peak binning with a width of 0.25 Da were used to discretize *m/z* domains in order to group peaks from MALDI-MS. 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 analyzed patient and rows corresponding to *m/z* bins. A CCA dimension reduced submatrix consisting of the top-ranked rows was then analyzed using linear discriminant analysis (LDA) in order to predict node status of the 28 patients from duplicated experiments. The rows of this matrix were then ranked using the developed LDA based method for their ability to distinguish between primary carcinomas with positive and negative nodes status. All analyses were replicated in parallel using -0.25/3 and +0.25/3 location shifted bin location, resulting in three analyses in total. The leave one out cross validation was used to evaluate the outcomes of classification. A majority rule was used to combine the data and top 30 *m/z* bins were selected as candidates for peptide identification.

2.2.3.3 MALDI-MSI spectra processing, statistical analysis and visualization in the SCiLS Lab

For MALDI-MSI data analysis and result visualization, the raw sample spectra were also imported into the SCiLS Lab software (v2016b). Here the data was pre-processed with TopHat baseline removal, total ion count (TIC) normalization, peak alignment and peak picking, and data spectra segmentation (medium denoise) using internal analytical functions. The annotated aligned spectra from ROIs has performed classification analysis based on internal LDA functions (no CCA dimensional reduction) and the leave one out cross validation was used to evaluate the training of classification results in SCiLS Lab. All picked peaks were interpreted as *m/z* intervals with 0.100 or 0.125 Da width for data analysis and peak visualization. Data hypothesis analysis were down

by software hypothesis test function using internal Kruskal-Wallis and Wilcoxon signed-rank test script, results were exported from SCiLS Lab report table. The spatial expression map of *m/z* intervals of interest and receiver operating characteristic (ROC) plots between two comparing ROI regions were generated by the SCiLS lab result visualization and discriminative analysis.

2.2.3.4 Peptide identification in the MASCOT

Processed *nano*-LC-ESI-MS/MS spectra were then exported as the Mascot generic format and submitted to Mascot (Version 2.3.02) for peptide identification. Search parameters used in this study were as follows: SwissProt Homo sapiens database (version *2016-02*) was searched, the digestion enzyme was specified as trypsin with up to 2 missed cleavages, fixed modification of carbamidomethylation of cysteine, variable modification of oxidation of methionine, N-terminal acetylation, and glutamine to pyroglutamate as variable modifications. MS/MS mass tolerance of 40 ppm, and MS mass tolerance of 0.2 Da. In Mascot, the peptide false discovery rate (FDR) was set to 1% using Percolator, the peptide ion score cut-off was set to 20, and the peptide significance score was set to less than 0.05.

2.2.3.5 Peptide identification and LFQ data processing in the MaxQuant

Raw spectral data acquired in the data-dependent acquisition was exported to the MaxQuant (v1.6.0.1) software and peptides identification was done by searching the UniProt human-reviewed protein fasta file (generated from version 2017_06, human taxonomy, 20196 entries) in software Andromeda searching engine with following parameters: trypsin/P with up to 2 missed cleavages fixed modification of carbamidomethylation of cysteine, variable modification of oxidation of methionine, N-terminal acetylation, and glutamine to pyroglutamate as variable modifications.; 40ppm MS/MS mass tolerance and 1% False Discovery Rate (FDR). The peptide identification cut-off was set with MaxQaunt PEP value less than 0.01. Proteins identified by unique peptides are grouped in one protein group and quantified together in the MaxQuant, and the proteins of the

group are used to represent this identified protein group in the following study. The protein LFQ intensities of data-dependent data acquisition LC-MS/MS were automatically calculated when the MaxQuant LFQ tool was selected. In general, the LFQ intensities were normalized based on extract-ion-current (XIC) among different comparing runs using the internal MaxLFQ with minimum LFQ count ration of 2.

2.2.3.6 DDA LFQ data analysis in the Perseus

The protein LFQ results were summarized into a 'proteingroups.txt' file in the resulting folder of MaxQuant and exported into Perseus software (v1.6.0.2) for the statistical analysis. The LFQ results were first filtered out those unidentified peptides and hit the reverse database, then transferred protein LFQ intensities into $\log_2$ form. Regrouped patient samples according to comparing classes and removed proteins only presenting in one patient of the class. Performed the missing values imputation functions of Perseus according to the normal distribution (width = 0.3, shift = 1.8), assuming these proteins were close to the detection limit. For chapter 5 small cohort patients' statistical analysis, the 2 class student t-test was applied on those comparing patient classes with the significant threshold (*p*-value) less than 0.05 and minimum 2 fold of change between 2 comparing classes; For chapter 4 large TMA cohort patients' statistical analysis, the 2 class student t-test was applied on those comparing patient classes with the significant threshold (*p*-value) less than 0.01 and minimum 2 fold of change between 2 comparing classes, besides, the permutation-based false discovery rate (FDR) was used for the t-test result correction. The Volcano plot was created in the Perseus based on protein $\log_2$ form LFQ intensity difference between two classes (X-axis) and $-\log_{10}$ form t-test *p*-values (Y-axis) and protein LFQ intensities among each patient were interpreted in heat maps created by the in-house R script (section 2.2.4.7)

2.2.3.7 DIA LFQ analysis in the Skyline MS

Data were analyzed in the Skyline MS software against a spectral library generated from the previous *nano*-LC-ESI-MS/MS experiments. Unique peptides from the protein spectral library were selected for target protein quantification. The peptide and transition settings during analysis were as follows; trypsin was specified as the cleavage enzyme with a maximum of 1 missed cleavage, the minimum and maximum peptide lengths were set to 7 and 25 amino acids respectively, precursor charge states 2 and 3, ion charges 1 and 2, ion types y and b, from ion 3 to 6 ions, 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 of spectral library MS/MS identification was used.

2.2.3.8 IHC staining analysis in the IHC Profiler

For each tissue section, the Nanozoomer scanned images were viewed in NDP Viewer (v2.3.1) and three representative photomicrographic images from ROIs were captured at 40× magnification and exported. The analysis was carried out using the IHC Profiler plugins in the ImageJ with cytoplasmic stained image mode and H DAB vectors which determines stain pixel into a high positive, positive, low positive, and negative levels. The levels of antibody-positive staining for each tissue type were summed and compared across patients using GraphPad Prism 6 (v008), where the median of sample total positive stains score, score standard error, and significance in paired or unpaired t-test was calculated.

2.2.4 Analytical R Scripts

2.2.4.1. The script of basic R functions supporting other R scripts in this thesis. This script includes the following functions. 1. Read peaklist files in the default data folder, then extract detected region numbers and X, Y coordinates of each data points from its peaklist file names and transfer that spatial information into a new datatable. 2. Import all acquired peaklist files from the folder into the datatable and merge them into one comprehensive peaklist. 3. Plot spatial information of imported peaklists based on their X, Y coordinates with experimental observations, including acquisition, intensity and SN, using the ggplot2 package.

2.2.4.2. R script for MALDI-MSI data quality evaluation. This script first reads into peaklist files and save them into the comprehensive peaklist based on basic functions described above, and then summarize the number of peaks, the number of spectra acquired, and the number of spectra without picked peaks in the diagnostic report table. The numbers of total peaks and peaks detected with 100ppm of Internal Calibrants (1296.685, 1570.677, 2147.199, 2211.105, and 2932.588) masses in each data point are plotted with their coordinates, and then printed out diagnostic table and images in the data folder.

2.2.4.3 R script for peak grouping. This script provides three different peak grouping strategies, including using a linkage distance of 0.01 Da, a linkage distance of 0.05 Da, and Density-Based Spatial Clustering of Applications with Noise (DBSCAN) which group peaks with grouping radius of the neighbourhood (eps) at 0.02 and the minimum number of peaks (mnspts) set at 200 (Tissue dataset) or 500 (TMA dataset) in each group, and then the abundant weight means (AWM) of each peakgroup are calculated and save with the list of peakgroup result (Campello et al., 2013). The following script functions read tissue annotations that saved in the XML format with the comprehensive peaklist into dataframe and then exported peak grouping results as one huge peaklist file with peakgroup numbers and peak spatial coordinates.

2.2.4.4 R script for hypothesis tests. This script first processes the huge peaklist file grouped above and summarizes grouping results into a short peakgroup list. Within each peakgroup, peaks

acquired from regions of different vulvar cancer tissue types are performed the Kruskal-Wallis (KW) test to examine the significance of peak intensity difference in those tissue types. While for peak grouping data acquired from TMA samples, the Wilcoxon Rank Sum (WX) test is performed on each peakgroup to examine the significance of peak intensity difference between LNM positive and negative patients. False discovery rates of KW and WX tests are further controlled by the Benjamini-Hochberg correction (F test) and results are saved into the peakgroup short list.

2.2.4.5 R script for identified peptide library construction, peak matching with and peak intensity boxplots. This script first uses the hypothesis test result to filter all peakgroups with significance threshold ($p<0.05$) and a number of peaks per group (>1000), and then the median intensity difference between selective tissue types in each peakgroup is calculated and written into the shortlist for further peptide matching. This following script reads the MASCOT or MaxQuant peptide identification result and combined them in one list called peptide library. To match AWM of each peak groups with identified peptide masses in the library, calculated protonated masses of tryptic peptides that fitted in the mass range of selective peakgroups are recorded as positive hits and all peakgroup matching results are summarized in a new list with mass errors, matched peptide masses, sequences, and other details. The last function of this script is boxplot which plot the peak intensities of selective peakgroup with different vulvar cancer tissue types or LNM status in the ggplot2 package.

2.2.4.6. R script for Received Operation Curve (ROC) analysis. This script first tests discriminative powers of LNM using peak intensities from selective peakgroups in a logistic regression model using internal functions in R, and then this model is further used to predict sample LNM status and the model performance is interpreted as a ROC curve using the ROCR package. The value of Area Under Curve (AUC) of each group is also calculated in R function from this package.

2.2.4.7. R script for heatmap creation. This script first read the datatable of protein LFQ intensities exported from Persues results (Chapter 2.2.3.6), and then re-organizes protein intensities of different patent samples to a data matrix and draw a hierarchy clustering heatmap based on this matrix using the ggplot2 package.

2.2.5 Reactome Pathway Analysis

The Reactome Pathway Analysis (http://www.reactome.org) was queried using the Pathway Browser (version 3.4) as the searching tool. Parameters used for species was *Homo sapiens* and protein candidates were integrated into signalling networks with their biological significance.

This page intentionally left blank

CHAPTER 3: PROTEOMIC CHARACTERIZATION OF HPV-INDEPENDENT VULVAR CANCER DEVELOPMENT

3.1 INTRODUCTION

Vulvar cancer is the fourth most common gynecological cancer globally, constituting 5% of all gynecological cancers and 0.6% of total cancer cases in 2016 (Board, 2016). Although vulvar cancer is relatively rare compared to endometrial or cervical cancer, recent incidence rates of human papillomavirus (HPV) independent invasive vulvar squamous cell carcinoma (VSCC) and its precursor lesion, differentiated vulvar intraepithelial neoplasia (dVIN), have significantly increased amongst young women, especially Indigenous young women in East Arnhem Land in the Northern Territory (Condon et al., 2009a, Joura, 2002). Previously, tumour suppressor proteins, such as p16 and pRb, have been examined in vulvar cancer tissues as diagnostic biomarkers, but they are only sensitive to HPV-dependent VSCCs and have not shown any diagnostic value in HPV-independent patients (Santos et al., 2004, Lerma et al., 1999). The overexpression of p53 is an important hallmark of many cancers, but the high abundance of the p53 is also reported in LS and dVIN patients without vulvar cancer (Muller and Vousden, 2013, Pinto et al., 2010). However, the proteomic characterization of vulvar cancer tissues has been poor, in part due to sample rarity, hence the understanding of HPV-independent VSCC progresses from healthy tissue to the precursor lesion at the protein level is low.

Previous studies in gynecological cancers have shown that MALDI-MSI could acquire a comprehensive proteomic image of the tissue whilst retaining the important spatial information (Gustafsson et al., 2011b, Mittal et al., 2016). Here, vulvar cancer tissues from 12 representative HPV-independent patients were examined using MALDI-MSI and the results revealed proteomic fingerprints of benign epithelium, dVIN, VSCC and metastatic lymph node (LN). In previous vulvar cancer histopathological studies, pathologists noticed many keratinized epithelial structures disappear in VSCCs, which suggest a potential relationship between cytokeratin alternations and tumour proliferation in the vulva (Canavan and Cohen, 2002). In the published manuscript that forms the majority of this chapter, the abundance level of the intermediate filament protein cytokeratin 5 (CK5 or K2C5) was found to be significantly decreased in VSCCs compared with dVIN and benign epithelium (Zhang et al., 2016b). As CK5 is an important component of the cytoskeletal scaffold and helps to maintain epithelial cell shape and differentiation, decreased CK5 levels are likely related to tumorigenesis in the squamous cell carcinoma (Jacob et al., 2018, Ratushny et al., 2012). Additionally, β-actin

(ACTB) levels were found to be increased in metastatic LN and VSCC tissues compared with benign epithelium and dVIN, which was further validated by the immunohistochemistry (IHC) staining and label-free quantification (LFQ) mass spectrometry analysis on consecutive tissue sections. Interestingly, a recent study analyzing the proteome of several cancer cell lines suggests the remodelling of cytoskeleton proteins, such as ACTB or CK5, facilitates malignant tumour cell migration and metastasis to distance sites (Mouneimne et al., 2012, Shankar and Nabi, 2015). In this chapter, proteins associated with tissue morphological patterns of the different stages of vulvar cancer development are highlighted and provide an insight into the molecular pathology behind HPV-independent vulvar cancer progression.

3.2 MATERIAL AND METHODS

3.2.1 Sample Cohort: Tissue blocks from 12 Formalin-fixed paraffin-embedded (FFPE) patients diagnosed as advanced HPV-independent VSCC were retrieved from the archive of SA Pathology (Supplementary S2.1). All sample blocks were annotated by pathologists, and each of the tissues analyzed in this study contained one or more tissue areas of healthy vulvar epithelium, dVIN, VSCC and metastatic lymph nodes (LN).

3.2.2 MALDI-MSI Preparation and Acquisition: MALDI-MSI was performed on tissue from six VSCC patients as previously described in Chapter 2.2.3.3. In brief, FFPE tissues were sectioned and mounted on ITO slides. The FFPE tissues sections were deparaffinized, then subjected to heat-induced citric acid antigen retrieval followed by digestion with trypsin gold. Peptide Internal calibrants and CHCA matrix were overlaid onto the tissue sections. An ultrafleXtreme MALDI-TOF/TOF MS system was used for MS and *in-situ* MS/MS acquisition.

3.2.3 MALDI-MSI Data Analysis: Acquired spectra were recalibrated using the internal calibrants and pre-processed using flexAnalysis and flexImaging software (v4.0) in order to produce peak lists. During the pre-processing the Snap Algorithm with a signal to noise (SN) ratio of 2, a baseline subtraction was performed using the TopHat method, and baseline smoothing were performed using the Gauss algorithm. Tissue-specific regions of interest (ROI) were selected and exported into flexImaging (v4.0). Peaklists from all ROI were combined and density-based clustering (DBSCAN*, with an epsilon of 0.02 and a minimum density of 100 peaks) was used to cluster peaks into peak groups based on their *m/z* values. Only peakgroups containing at least 10,000 peaks were considered for further analysis. The median log intensity in each tissue type was calculated for each peak group and for each patient. Peak groups were heuristically ranked by the largest difference in median log intensity between tissue types. The raw data was also analyzed using the SCiLS lab software where the processing steps of baseline removal and normalization were carried out. Ion intensity maps for the peptides of interest were generated using the SCiLS lab software (Chapter 2.2.3.3).

3.2.4 Identification of *m/z* values by *in-situ* MALDI MS/MS and *nano*-LC-ESI-MS/MS: In order to achieve peptide identifications for the *m/z* values of interest acquired by MALDI-MSI, *in-situ* MS/MS was performed directly from the tissues analyzed in the imaging

experiments and was searched using Mascot (v2.3.2) as previously described (Chapter 2.2.2.4) . For matching back to peptide sequences obtained by data-dependent acquisition, *nano*-LC-ESI-MS/MS, ROIs were collected from tissue sections using laser capture microdissection (LCM), subjected to HIAR, and digested with trypsin. The *nano*-LC-ESI-MS/MS was performed using an Ultimate 3000 RSLC system coupled to an Impact II QTOF mass spectrometer via an Advance CaptiveSpray source. Peptide separation was performed using a 75µm C18 column over 180 min. MS scans were acquired in the mass range of 300 to 2200 *m/z* in a data-dependent fashion using Bruker's Shotgun Instant method (Chapter 2.2.2.6). Acquired spectra were subjected to Compass DataAnalysis for OTOF (v1.7). Processed spectra were then exported to Mascot generic format and submitted to Mascot (v2.3.2) for identification. Search parameters were as follows; SwissProt Homo sapiens database search, trypsin digestion with up to two missed cleavages, variable modification of oxidation of methionine, MS mass tolerance of 40 ppm and an 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 of <5% (Chapter 2.2.3.4). Matching between the MALDI-MSI and *nano*-LC-ESI-MS/MS was done by comparing the experimental *m/z* values of the *nano*-LC-ESI-MS/MS sequenced peptides to the *m/z* values from the compiled peak bins in R (Chapter 2.2.4.5).

3.2.6 Immunohistochemistry: CK5 and ACTB were found to be proteins of interest in the MALDI-MSI experiments, and hence were chosen for further validation by IHC on consecutive tissue sections (Chapter 2.2.2.9). In brief, FFPE tissue sections were de-waxed, rehydrated and subjected to antigen retrieval, then incubated overnight at 4°C with CK5 (1/500) and ACTB (1/1000) followed by incubation with biotinylated anti-mouse immunoglobulin (1/500) and streptavidin-HRP (1/500). Immunoreactivity was detected using diaminobenzidine (DAB)/H_2O_2 and streptavidin-HRP (1/500). Diaminobenzidine (DAB)/H_2O_2 substrate was added before counterstaining with hematoxylin, dehydrating and mounting in Pertex. Sections slides were scanned in Nanozommer and viewed in NDP view. The analysis was carried on IHC Profiler in ImageJ, for each tissue, three random representative images at 40x magnification were analyzed

3.2.7 Label-free quantification: The data-independent acquisition *nano*-LC-ESI-MS was performed on cancer tissues from four patients with benign, dVIN, VSCC and metastatic LN. The *nano*-LC-MS was performed as described above using an Ultimate 3000 RSLC system

coupled to an Impact II QTOF mass spectrometer. Data was acquired data using Bruker's Middle Band CID method where a mass range of *m/z* 375 to 1206 is scanned in 26Da increments and peptides within the isolation mass window are fragmented with increasing collision energies of 20 to 36. The collected data were analyzed in the Skyline software against a spectral library generated from the previous experiments. The peptide and transition settings during analysis are described in Chapter 2.2.3.7. In a short, trypsin was specified as the cleavage enzyme with a maximum of one 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 min retention time window of spectral library MS/MS identification was used. Summed area intensities for the analyzed peptides were calculated from y and b ions 3 to 6 (starting from ion 3). For each peptide analyzed, the relative intensity in each tissue type was normalized to the relative intensity of the normal tissue from each patient (Chapter 2.2.3.7).

3.3 RESULTS

3.3.1 Results of the manuscript

In the published manuscript (Zhang et al., 2016b), an examination of adjacent benign vulvar epithelium, dVIN, and VSCC from six HPV-independent patients by peptide MALDI-MSI is presented (***Table 3.1***). The results reveal the differential abundance of multiple peptides from the protein CK5 across the benign epithelium, dVIN, and VSCC. The sequence identities of the peptides are confirmed by matching the MALDI-MSI result, which provide information at the intact peptide level, with peptide sequence results obtained by *nano*-ESI-LS-MS/MS and *in-situ* MALDI-MS/MS. The differences observed in the relative abundance of CK5 by MALDI-MSI was further analyzed by immunohistochemistry (IHC) in tissue from eight VSCC patients. A decrease in CK5 immunostaining was observed in the VSCC compared to the healthy epithelium and dVIN. These results provide an insight into the molecular fingerprint of the vulvar intraepithelial neoplasia that appears to be more closely related to the healthy epithelium than the VSCC. Results of the published manuscript were provided below:

Table .3.1: Pathology details of the HPV-independent vulvar cancer patients (n=12) analyzed in the published manuscript. Full clinical and pathological details for each patient are provided in Chapter 2.2.1.1.

Patient No.	Tissue Type	dVIN	LNM	Lichen sclerosus	Age at diagnosis	FIGO Stage	Grade	Max diameter (mm)	Depth invasion (mm)	Lympho-vascular Space invasion (LVSI)	HPV Infection
1	VSCC	1	Pos.	1	49	III	3	37	0	0	Neg.
2	VSCC	0	Pos.	unknown	74	III	3	21	4	0	Neg.
3	VSCC	1	Pos.	0	76	III	3	45	10	1	Neg.
4	VSCC	1	Neg.	unknown	48	IB	1	32	6	0	Neg.
5	VSCC	1	Neg.	0	67	IB	2	20	6	0	Neg.
6	VSCC	1	Pos.	0	61	III	2	35	7	1	Neg.
7	VSCC	1	Neg.	0	74	IB	2	27	9	0	Neg.
8	VSCC	1	Pos.	1	87	III	3	30	10	1	Neg.
9	VSCC	1	Pos.	1	75	III	3	17	2.5	0	Neg.
10	VSCC	1	Pos.	unknown	66	III	2	18	5	0	Neg.
11	VSCC	1	Neg.	0	57	IB	2	27	4.4	0	Neg.
12	VSCC	1	Pos.	0	64	III	2	42	5	0	Neg.

3.3.1.1 Matrix-assisted Laser Desorption/Ionization Mass Spectrometry Imaging (MALDI-MSI)

MALDI-MSI was carried out on tissue sections containing regions of healthy vulvar epithelium, dVIN, and VSCC from six late-stage carcinoma patients. Peak groups, representing the imaged peptides across all of the analyzed tissue sections with the same *m/z* (assuming a mass accuracy of +/− 0.02 Da), were detected using density-based clustering of peaks (DBSCAN) (Campello et al., 2013) with a radius of neighborhood (eps) of 0.02 Da and a minimum of 100 points. Additionally, the stringent criteria of requiring each peak group to have a minimum of 10,000 peaks were implemented. This was done to ensure that each of the peak groups analyzed could be reproducibly detected across a majority of spectra. With this criteria in place, 31 peak groups were detected (***Table 3.2***). When the minimum number of peaks per peak group was lowered to 1000, 316 peak groups were detected. For each peak group, the abundance weighted mean (AWM) was calculated, representing the apex of the peak group and the *m/z* of the most intense peptide within the group. The mass widths of the 31 detected peak groups ranged from ±0.075 to 0.9 Da. The maximum difference, d, in median log intensities between the three tissue regions was calculated for each of the 31 peak groups of interest across the patient cohort (***Table 3.2***). Of the 31 peak groups, 19 were found to have a difference in fold change intensity of ≥1.4 fold across the three tissue types.

Table 3.2: Matrix-assisted laser desorption/ionization mass spectrometry imaging peak groups ranked heuristically by the largest difference in median log intensity between the healthy epithelium, differentiated vulvar intraepithelial neoplasia, and vulvar squamous cell carcinoma.

Peptide AWM [M + H] of the Peak Group [1]	Da Range of the Peak Group	No. Spectra in Peak Group [2]	Maximum Difference Median Log Intensity [3]
1198.71	±0.62	23099	0.75
1028.62	±0.56	15354	0.66
1410.72 *	±0.82	18784	0.58
1905.95	±0.52	10956	0.56
1095.58	±0.82	21704	0.55
944.54	±0.69	19354	0.53
957.63	±0.65	10464	0.52
857.60	±0.68	14956	0.50
976.51	±0.65	15534	0.49
878.56	±0.68	11213	0.46
810.48 *	±0.76	17095	0.45
958.59	±0.69	15432	0.42
943.62	±0.82	20030	0.41
865.42 *	±0.84	15617	0.41
1217.66	±0.55	10906	0.39
884.52 *	±0.66	14337	0.37
856.67	±0.74	15483	0.36
1143.65 *	±0.64	17685	0.36
1045.62	±0.65	16370	0.35
2147.19	±0.14	20418	0.33
1105.59	±0.50	11479	0.30
1296.68	±0.10	22823	0.29
870.64	±0.85	22749	0.29
842.60	±0.90	25210	0.22
1267.65	±0.32	11686	0.20
2932.59	±0.07	15500	0.19
816.50	±0.73	10983	0.19
971.56 *	±0.65	15998	0.18
1111.58	±0.51	13473	0.09
1570.70	±0.17	12029	0.05
864.63	±0.70	11114	0.05

[1] The abundance weighted means (AWM) *m*/*z* was calculated for the overall peak group; [2] Number of Spectra in each peak group compiled from all acquisition spectra across the patient cohort; [3] Peak groups were heuristically ranked based on the maximum difference, d, in median log intensity between the tissue regions of interest; * CK5 peptides.

3.3.1.2 Cytokeratin 5 (CK5) Identified as a Protein of Interest

In order to gain peptide identifications for the MALDI-MSI peak groups of interest, nanoflow liquid chromatography-tandem mass spectrometry (*nano*-ESI-C-MS/MS) was performed on laser microdissected regions of the healthy epithelium, dVIN, and VSCC. Matching between the MALDI-MSI peak groups and *nano*-LC-ESI-MS/MS data was done by aligning the experimental *m/z* values of the sequenced peptides that fell between the minimum and maximum *m/z* of each of the MALDI-MSI peak groups of interest. A table containing all of the matching results for the 31 peak groups of interest is provided in ***Supplementary S3.1***. Of the 31 MALDI-MSI peak groups of interest, six matched to sequenced peptides from cytokeratin 5 (CK5) (***Table 3.3***), hence CK5 was selected for further analysis. One of the peak groups (AWM [M + H] of 1410.72) matched to two unique CK5 peptides that share the same mass to within 0.009 Da, SFSTASAITPSVSR (1409.7203) and TTAENEFVMLKK (1409.7295). A MALDI-MSI annotated ion intensity map for the peak group *m/z* 1410.67, which matches to the CK5 peptides SFSTASAITPSVSR and TTAENEFVMLKK (both with an *m/z* 1410.67), is shown in ***Figure 3.1***. With regards to the remaining results in ***Table 3.2,*** only three other proteins were detected with more than one unique peptide match as shown in ***Supplementary S3.2***. Four other peptides were detected from the protein Neuroblast differentiation-associated protein AHNAK, three peptides were detected from the protein Annexin A4, and two peptides were detected from the protein Nicotinate phosphoribosyltransferase, but these peptides could not be sequenced by *in-situ* MS/MS. As the *in-situ* MS/MS was performed on tissue sections directly after the MALDI-MSI data acquisition using the same instrument, the identification results obtained much more reliable as compared with matching peptide masses detected using another MS platform, like *nano*-LC-ESI-MS/MS.

Table 3.3: CK5 matrix-assisted laser desorption/ionization mass spectrometry imaging peak groups of interest with matched nanoflow liquid chromatography-tandem mass spectrometry results.

Peptide AWM [M + H] [1]	***d*** [2]	**LC-MS/MS *m/z*** [3]	**LC-MS/MS MR** [4]	**Peptide Score** [5]	**Peptide Significance** [6]	**Peptide Sequence** [7]
810.48	0.45	405.7087	809.4028	21.44	0.0072	QSSVSFR
865.42	0.41	433.1998	864.3851	26	0.0025	SGGGGGGGFGR
884.52	0.37	442.7286	883.4427	24.74	0.0034	TSFTSVSR
971.56	0.18	486.2429	970.4712	25.1	0.0031	FVSTTSSSR
1143.65	0.36	572.3175	1142.6205	39.59	0.00011	LAELEEALQK
1410.72	0.58	705.8674	1409.7203	115.28	3×10^{-12}	SFSTASAITPSVSR
1410.72	0.58	705.872	1409.7295	150.55	8.8×10^{-16}	TTAENEFVMLKK

[1] The peptide abundance weighted mean (AWM) *m/z* calculated for each peak group; [2] Heuristically ranked maximum difference, d, between the healthy epithelium, dVIN, and VSCC; [3] Experimental *m/z* of the CK5 matched *nano*-LC-MS/MS detected peptide; [4] Experimental mass of the CK5 matched *nano*-LC-MS/MS detected peptide; [5] Mascot Ions Score; [6] Mascot Expected Score; [7] CK5 matched peptide sequence.

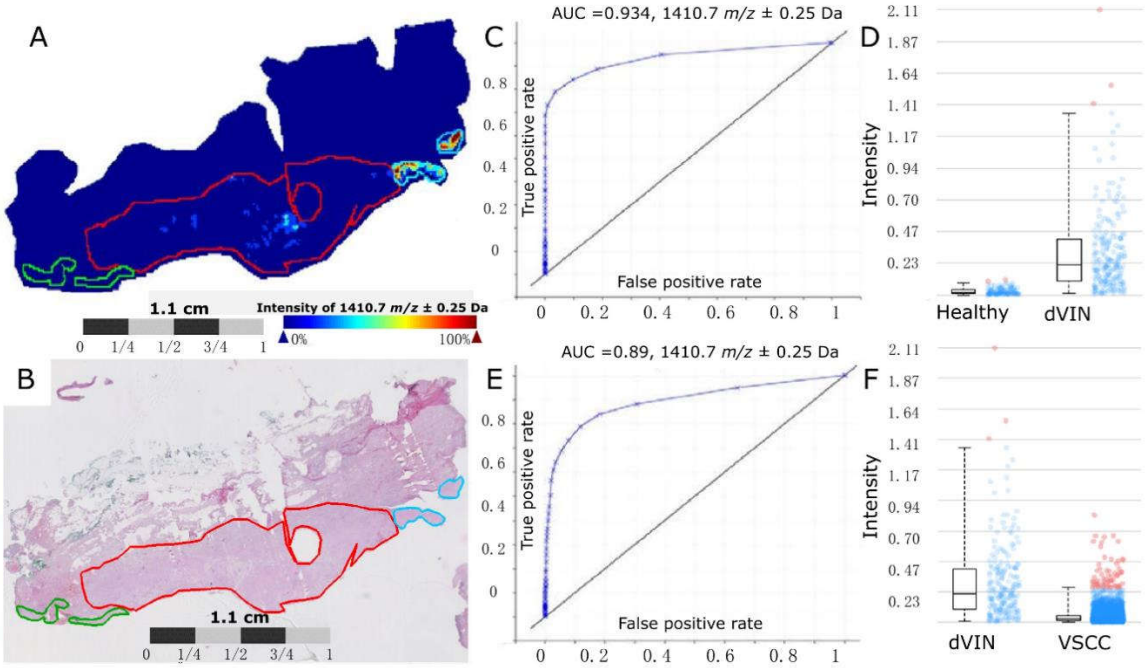


Figure 3.1: A MALDI-MSI annotated ion intensity map from a representative vulvar tissue section for the peak group *m/z* 1410.67. (**A**) An ion intensity map for *m/z* 1410.67, which matches to the CK5 peptides SFSTASAITPSVSR and TTAENEFVMLKK (both with a [M + H] *m/z* 1410.67). The regions of healthy epithelium are outlined in green, the dVIN is outlined in blue, and the VSCC is outlined in red; (**B**) H & E stain of Healthy (green), dVIN (blue) and VSCC (red) tissue post MALDI-MSI data acquisition; (**C**) A ROC curve and (**D**) boxplot analysis of the MALDI-MSI data for *m/z* 1410.67 comparing the annotated healthy epithelium to the dVIN region, AUC = 0.934; (**E**) A ROC curve and (**F**) boxplot analysis of the MALDI-MSI data for *m/z* 1410.67 comparing the annotated dVIN to the VSCC region, AUC = 0.89. Blue dots represent the *m/z* 1410.67 spectra intensity interval between the lower and upper quantiles. Red dots represent the spectra intensities outside of the intervals.

3.3.1.3 Immunohistochemistry (IHC) Analysis of CK5 comparing Healthy Epithelium, Differentiated Vulvar Intraepithelial Neoplasia, and Vulvar Squamous Cell Carcinoma

The differential detection of CK5 across the three tissue types as determined by MALDI-MSI was further investigated by IHC in tissue from eight VSCC patients. Of these eight tissues, three had been previously analyzed in the MALDI-MSI experiments. IHC analysis of CK5 revealed a significant difference in the level of staining between the dVIN (91.6 ± 6.1 %) and VSCC (75.4 ± 10.7 %) tissues ($p = 1.758 \times 10^{-7}$). A similar result was observed for the healthy as compared to VSCC tissue ($p = 0.004$), with increased staining observed in the healthy epithelium (86.7 ± 14.8 %). Overall, there was no significant difference in the level of staining between the dVIN and healthy epithelium ($p = 0.14$), although the mean staining intensity in the dVIN was slightly higher. Results for the IHC analysis and CK5 staining from two representative vulvar tissues are provided in ***Figure 3.2***. A direct comparison of the CK5 peptides SFSTASAITPSVSR and TTAENEFVMLKK, both *m/z* 1410.72, as detected by MALDI-MSI to the detection of CK5 by IHC on consecutive vulvar tissue sections is shown in ***Figure 3.3***.

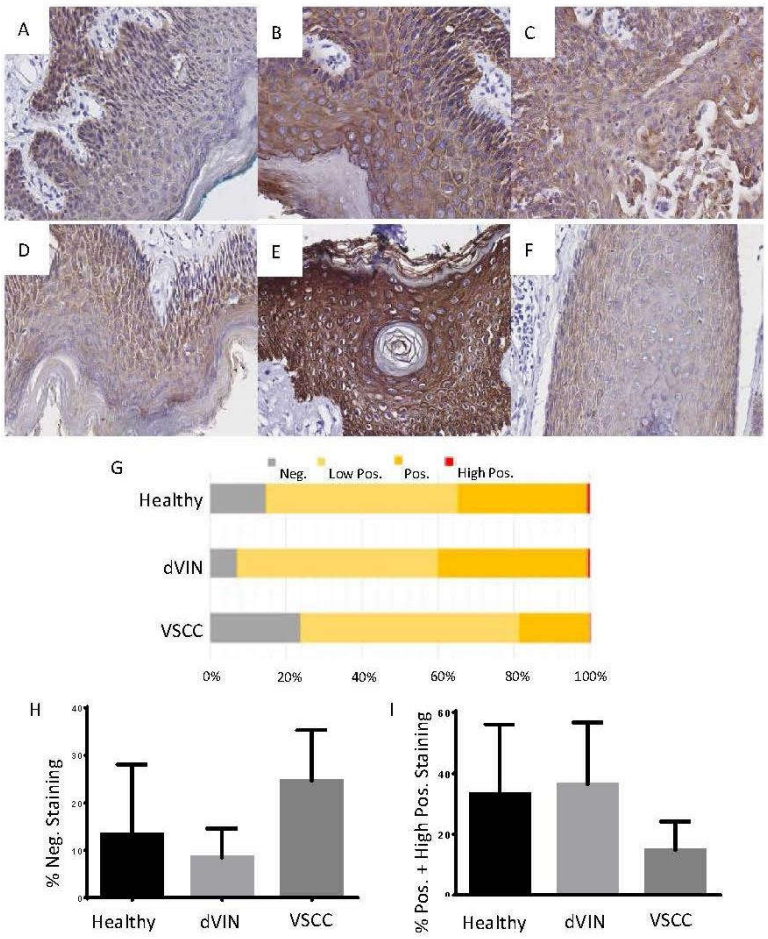


Figure 3.2: IHC analysis of CK5 on vulvar tissues containing regions of healthy epithelium, dVIN, and VSCC, all at 40× magnification. (A) CK5 staining of healthy epithelium; (B) dVIN; (C) VSCC from a single section; (D) CK5 staining of healthy epithelium; (E) dVIN; (F) VSCC from a single section of a different patient; (G) Quantitative analysis of CK5 staining across the tissue types was performed using IHC profiler-Image J. For each tissue section, three representative photo-micrographic images at 40× magnification were used and each image was assigned a staining score of negative, low positive, positive, and high positive; (H) The percent of negative staining was significantly different between the dVIN and the VSCC ($p < 0.0001$), but not between the dVIN and healthy epithelium ($p = 0.11$); (I) The per cent of positive and high positive staining was significantly different between the dVIN and the VSCC ($p < 0.0001$), but not between the dVIN and healthy epithelium ($p = 0.59$).

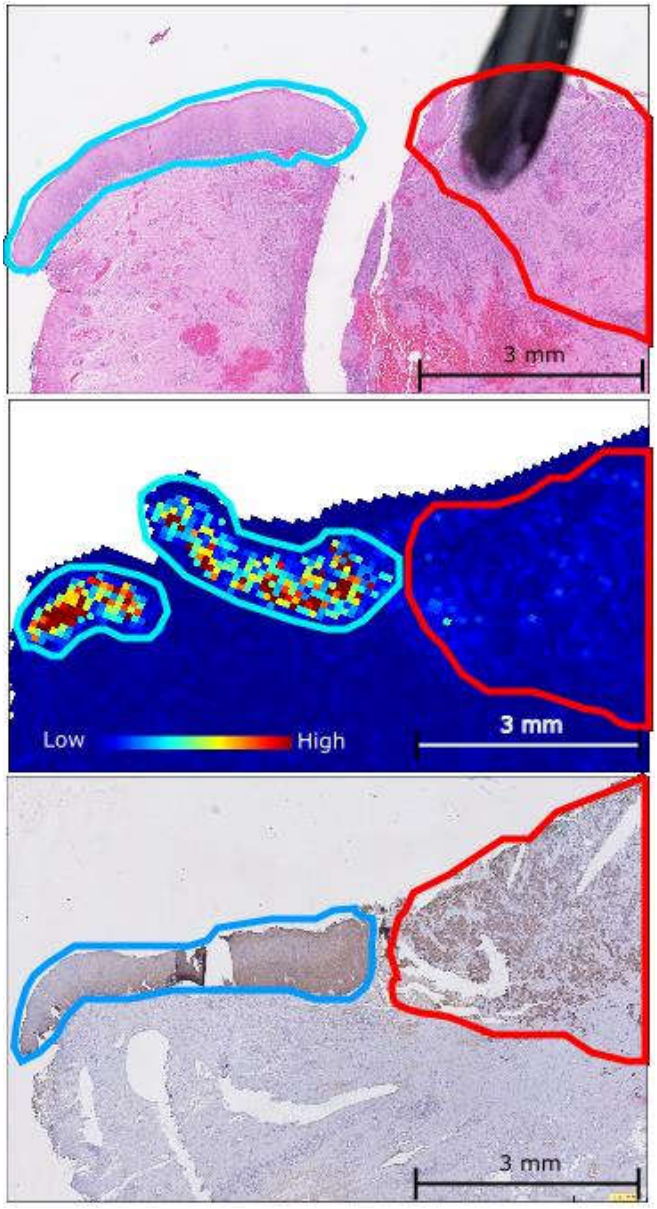


Figure 3.3: A comparison of the expression pattern observed for the CK5 peptides SFSTASAITPSVSR and TTAENEFVMLKK, both *m/z* 1410.72, in the dVIN (blue) and VSCC (red) using: MALDI-MSI ion intensity map (middle); H & E staining (top); and CK5 IHC staining (bottom) on consecutively cut vulvar tissue sections.

3.3.2 Proteomic characterization of metastatic vulvar cancer tissues

The characterization of benign epithelium, dVIN and VSCC by MALDI-MSI revealed a low abundance of CK5 in VSCC compared with the other two vulvar tissue types (Zhang et al., 2016a). In addition to the analyses described above, differences between metastatic lymph node (LNM) tissues were compared with benign epithelium, dVIN and VSCC in four patients. Across the four tissue types, MALDI-MSI analyses revealed the differential abundance of the five proteins, β-actin (ACTB), histone H2A type 1-B/E (H2A1B), hemoglobin (HBB), histone H4 (H4) and cytokeratin 10 (CK10). These findings were validated by IHC and data-independent acquisition (DIA) *nano*-LC-ESI-MS/MS on consecutive tissue sections.

3.3.2.1 MALDI-MSI

FFPE tissues from four HPV-independent vulvar cancer patients were sectioned at 6 µm and analyzed in duplicate using the protocol described in Chapter 2.2.2.1. ***Figure 3.4*** shows annotated hematoxylin and eosin (H&E) stained tissue sections delineating the different tissue types of benign epithelium, dVIN, VSCC and LNM.

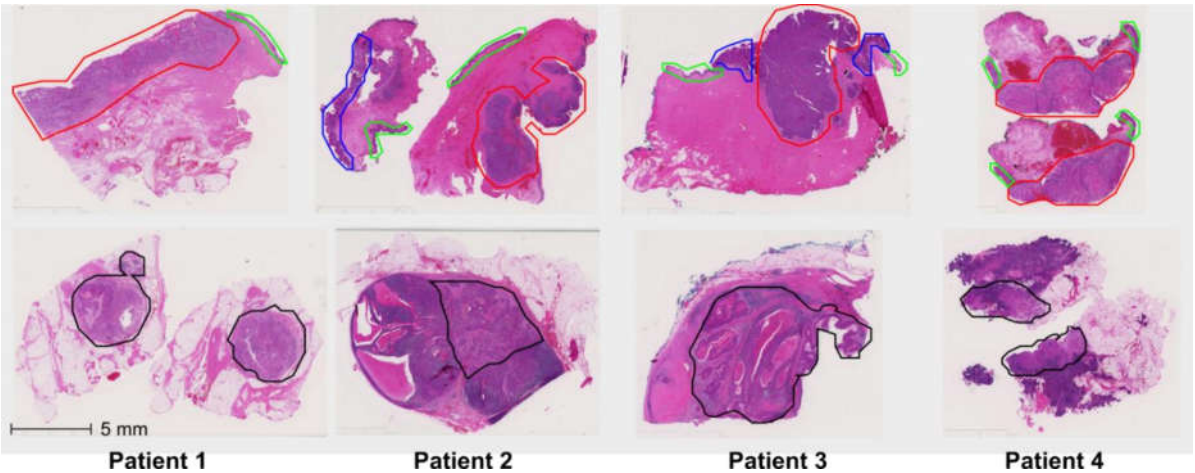


Figure 3.4: Haematoxylin and eosin (H&E) stain tissue sections from four patients. Consecutive sections were used for MALDI imaging and *nano*-LC-ESI-MS/MS. Different tissue types are encircled: benign epithelium

(green), dVIN (blue), VSCC (red) and metastatic LN (black). Annotations were made by a pathologist. Scale bar = 5 mm for all images.

A total of four benign, VSCC and LNM with two dVIN tissue samples were analyzed by MALDI-MSI, with 15000 to 25000 spectra acquired for each patient sample. Tissue specific regions of interests (ROIs, including healthy epithelium, dVIN, VSCC and LNM) were circled in flexImaging (v4.0) according to the pathological annotations. The mean spectra of four ROIs acquired from the same vulvar cancer tissue were normalized using Total Ion Count (TIC) and shown in ***Figure 3.5***. Peptide peaks were detected that demonstrated differential intensity levels as across the four vulvar cancer tissue types, such as the peak *m/z* 1198.6 (***Figure 3.5***).

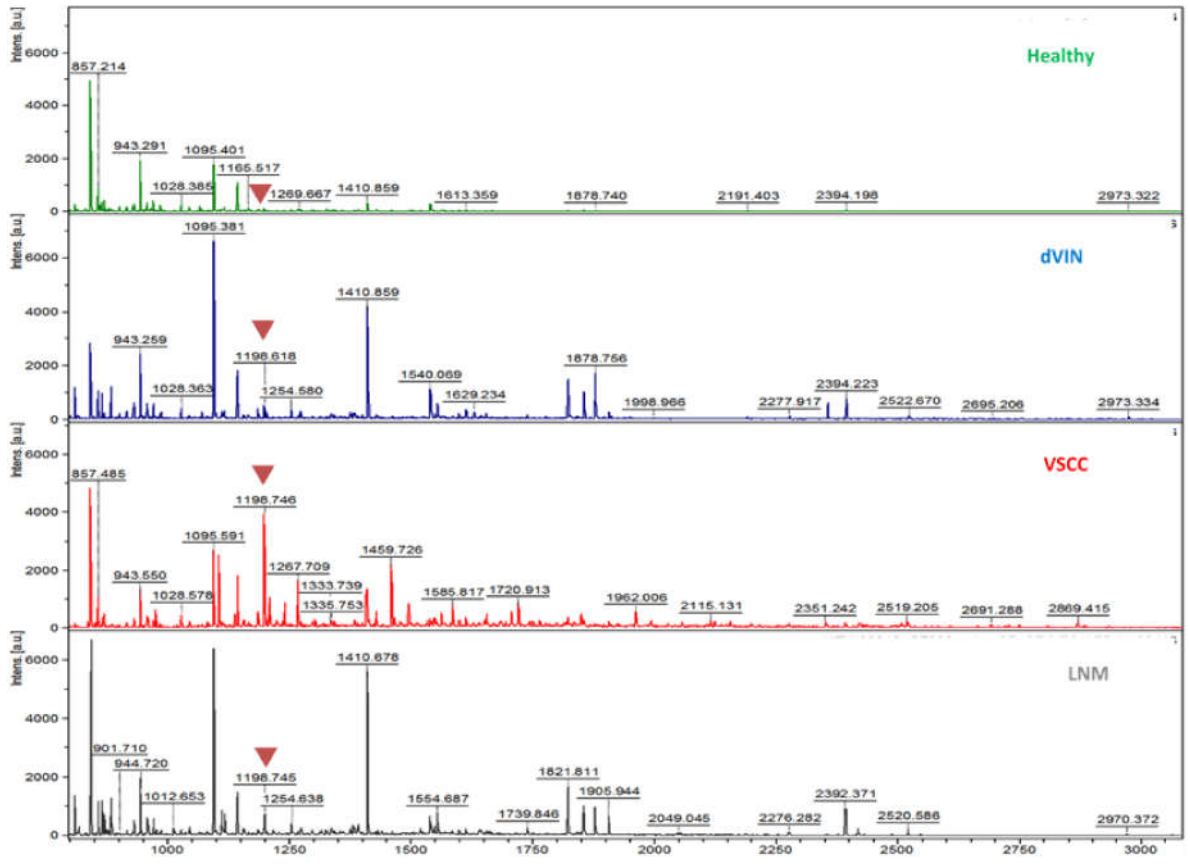


Figure 3.5: Representative benign (green), dVIN (blue), VSCC (red) and metastatic LN (black) mass spectrums obtained by MALDI-MSI detection from a reprehensive patient sample. Detective peptide peaks displayed differential intensity levels across four vulvar tissue types, *m/z* 1198.7 as an example (red triangle).

Raw spectra were processed into peaks in the flexAnalysis (v4.0) and grouped using the same parameters described in section 3.3.1.1. In total, 1078 peakgroups were detected from duplicate MALDI-MSI analyses. The AWM of each of these peakgroups was calculated. To evaluate the differences in peakgroup intensities across the four different tissue types a Kruskal-Wallis (KW) test was performed and the acquired *p*-values were adjusted using the Benjamini-Hochberg test. As a result, 844 peakgroups showed statistically significant differences across the different vulvar cancer tissue types. Details of this analysis, including the AMW, *p*-values of adjusted KW test, mass gaps and SN differences, are listed in ***Supplementary S3.3*****.** The top 10 significantly different peakgroups in terms of intensity variations across the tissues types are shown in ***Table. 3.4.***

Table 3.4: Top 10 MALDI-MSI peak groups ranked heuristically by KW test results among benign, dVIN, VSCC and LN tissue samples.

Peak Group No.[a]	*m/z min*	*m/z max*	*mass gap*	*number of Peaks*	*AWM*	*KW test p-value*[b]	*F-test p-value*[c]	*SN diff*[d]
143	909.292	909.471	0.179	7280	909.3951	1.42E-211	4.29E-241	96.52
335	1105.104	1105.168	0.064	1007	1105.139	1.22E-09	3.32E-09	90.9955
250	1011.376	1011.824	0.448	21099	1011.586	3.89E-37	1.60E-20	90.8245
366	1143.09	1143.219	0.129	2162	1143.133	2.51E-11	1.72E-12	77.0575
979	2183.541	2183.617	0.076	1558	2183.576	1.09E-23	8.79E-21	30.4585
802	1672.627	1672.867	0.24	11484	1672.723	8.07E-281	*p < 0.001*	28.367
412	1198.443	1198.85	0.407	137613	1198.687	5.25E-251	*p < 0.001*	27.478
580	1410.41	1410.908	0.498	96913	1410.706	7.62E-294	3.81E-130	25.619
789	1657.749	1657.839	0.09	3062	1657.788	2.38E-137	9.19E-115	25.1325
946	2050.225	2050.319	0.094	2780	2050.251	2.89E-59	3.98E-43	25.095

a. Peakgroups number came from DBSCAN peak grouping of comprehensive peaklists for TMA MALDI-MSI data.

b. The Kruskal-Wallis test *p*-values within each peak group.

c. *p*-value of Benjamini–Hochberg test of false discovery rate (FDR) control.

d. SN difference was based on median SN change among four vulvar cancer tissue types in each peak group.

3.3.2.2 Identification of discriminative *m/z* values

In order to obtain sequence identifications for the peptides in the peakgroups of interest, consecutive sections from the tissues used in the MALDI-MSI analyses were mounted on polyethylene naphthalate (PEN) slides. Areas of benign tissue, dVIN, VSCC and LNM were first deparaffinized, rehydration, and then collected using laser capture microdissection (LCM) according to methods described in section 3.2.4. The samples were then subjected to HIAR and tryptic digestion as previously described (Chapter 2.2.1.10). The pathologist annotations are shown in ***Figure 3.4***. Following analysis of the samples by *nano*-LC-ESI-MS/MS, a total of 8226 unique tryptic peptides from 2918 proteins were identified in the SwissProt human database using the MASCOT (v2.3.2) search engine (search parameters described section 3.2.4). Overall 3602 tryptic peptides from 1180 proteins were identified from the benign tissue, 5304 peptides from 1500 proteins were identified from the dVIN tissue, 7169 peptides from 2097 proteins were identified from the VSCC, and 5686 peptides from 2112 proteins were identified from the LNM. These results were used to create a comprehensive tryptic peptide library of vulvar cancer tissues for the following MALDI-MSI peak matching. The MALDI-MSI *m/z* values were matched to the vulvar cancer peptide library with less than 50 ppm error tolerance using in-house R scripts (Chapter 2.2.4.5). The MALDI-MSI *m/z* values that matched the peptide library are shown in ***Supplementary S3.4***.

To ensure the reliability of the MALDI-MSI peakgroup matching back to the *nano*-LC-ESI-MS/MS sequenced peptides, *in-situ* MS/MS was performed using the LIFT mode of MALDI-TOF-MS instrument to directly sequence and identify tryptic peptides from consecutive tissue sections according to the protocol described in Chapter 2.2.2.4 (Almeida et al., 2015). In total, six peptides from five proteins returned positive hits in the Swiss-Prot Human protein database searched using MASCOT(v2.3.2), with protein expected scores of less than 0.05 (***Table 3.5***). Details of the peptide sequencing results and the associated search parameters are provided in Chapter 2.2.3.4.

Table 3.5: Abundant peaks groups identifiable by *in situ* MALDI MS/MS found to have significantly different SN levels between the four tissue types ($p<0.01$)

in situ MALDI *m/z* value	Minimum *m/z* value of the peak group	Maximum *m/z* value of the peak group	Peptides AWM [M+H] of the peak group [1]	*p*-value from KW test	*in-situ* MS/MS identification Sequence	Accession name [2]	Protein Identification [3]	Ion scores [4]	Expected score [5]	Mass Error ppm
944.457	944.144	944.953	944.516	<0.001	AGLQFPVGR	H2A1B_HUMAN	Histone H2A type 1-B/E, P04908	38	0.01	11.2
976.475	976.119	977.02	976.457	<0.001	AGFAGDDAPR	ACTB_HUMAN	Actin, cytoplasmic 1, P60709	20	0.085	0.73
1198.687	1198.427	1198.879	1198.687	<0.001	AVFPSIVGRPR	ACTB_HUMAN	Actin, cytoplasmic 1, P60709	29	0.027	20.9
1314.663	1314.411	1314.774	1314.663	<0.001	VNVDEVGGEALGR	HBB_HUMAN	Hemoglobin subunit beta, P68871	39	0.0087	-5.19
1325.671	1325.459	1325.831	1325.671	<0.001	DNIQGITKPAIR	H4_HUMAN	Histone H4, P62805	38	0.0074	-44.35
1707.711	1707.346	1707.866	1707.711	<0.001	GSLGGGFSSGGFSGGSFSR	K1C10_HUMAN	Keratin, type I cytoskeletal 10, P13645	45	0.0014	48.9

[1] Abundance weighted means (AWM) *m/z* was calculated from the overall peak group

[2] SwissProt accession name

[3] Protein identification includes protein name, SwissProt accession number and gene name

[4] MASCOT ions score

[5] MASCOT expected score

In summary, there were 844 peakgroups with intensities that were significantly different amongst the four vulvar cancer tissue types ($p<0.01$), and six of them were matched with the same peptide sequences identified using both MS methods of *nano*-LC-ESI-MS/MS and *in-situ* MS/MS (***Table 3.5***). Among these identified peakgroups, two of them, *m/z* 976.475 and *m/z* 1198.687, were matched back to protonated masses of β-actin (ACTB) peptides, AGFAGDDAPR and AVFPSIVGRPR, in the peptide library. A MALDI-MSI annotated ion intensity map for matched ACTB peptide *m/z* 1198.687 across vulvar cancer patients (n=4), which matches to the ACTB peptide with higher protein abundances in the LNM and VSCC compared with the healthy and dVIN lesions of patients is shown in ***Figure 3.6***. The ROC curves and ion intensity boxplots of the same ACTB peptide (*m/z* 1198.687) revealed differential protein abundances across tissues are also shown.

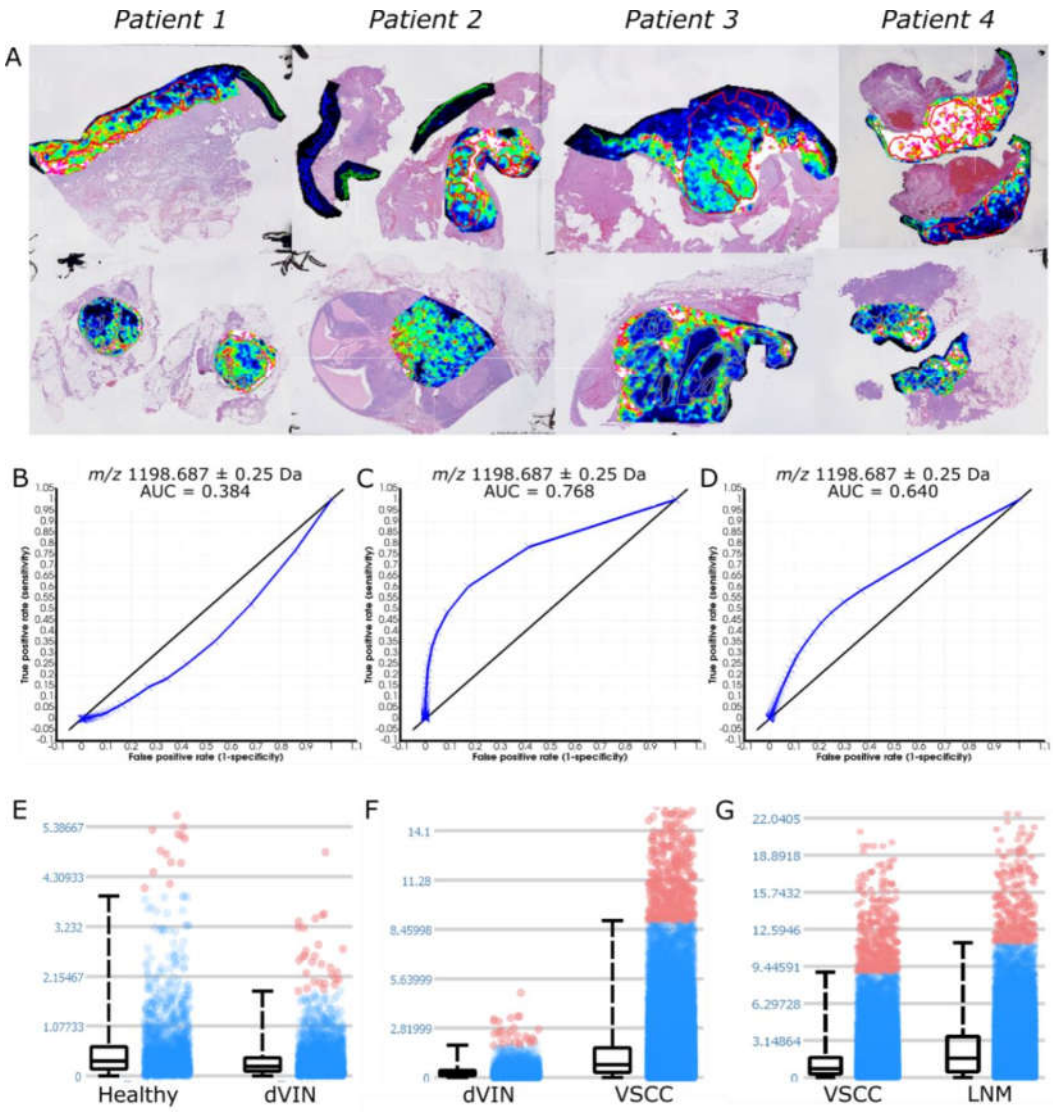


Figure 3.6: MALDI-MSI annotated ion intensity map from four vulvar cancer patients for ACTB peptide *m/z* 1198.687. (A) Ion intensity map for *m/z* 1198.687, which matches to ACTB peptide AVFPSIVGRPR. Regions of healthy circled in green, VSCC in red, dVIN in blue and LNM in grey; (B) ROC curve and (E) Boxplot analysis of MALDI-MSI data for *m/z* 1198.687 comparing dVIN and healthy samples, AUC = 0.384; (C) ROC curve and (F) Boxplot analysis of MALDI-MSI data for *m/z* 1198.687 comparing VSCC and dVIN samples, AUC = 0.768; (D) ROC curve and (G) Boxplot analysis of MALDI-MSI data for *m/z* 1198.687 comparing LNM and VSCC samples, AUC = 0.640. Blue dots represent spectra intensity interval between lower and upper quartiles. Red dots represent spectra intensities outside of intervals.

The AWM of two peakgroups, *m/z* 976.475 and 1198.687, were matched back to tryptic peptides from the protein ACTB using the peptide library, AGFAGDDAPR and AVFPSIVGRPR. Both peptides showed high peak intensity levels in LNM and VSCC tissues as compared to the intensities observed in the regions of dVIN and healthy tissues.

Other peakgroups identified as tryptic peptides of H2A1B, HBB, H4 and CK10 demonstrated their highest intensities in VSCC, dVIN, dVIN and metastatic LN respectively (***Supplementary S3.5***). To validate the ACTB protein abundance changes highlighted in the MALDI-MSI data, immunohistochemistry (IHC) analysis was carried out on consecutive tissue sections.

3.3.2.3 Immunohistochemistry analysis of ACTB

To further confirm the differential abundance of ACTB across the benign epithelium, dVIN, VSCC and metastatic LN tissues as observed by the MALDI-MSI, IHC was performed on the consecutive tissue section of the same four patients (***Figure 3.7***). Quantitative analysis of ACTB staining across the tissue types was performed using the IHC analysis software profiler-Image J and was further analyzed using GraphPad Prism 6 (v008). IHC analysis of ACTB revealed a significant difference in the level of positive staining between dVIN (31.21 ± 18.55%) and LNM (64.81 ± 8.09%) tissue samples ($p < 0.001$, Kruskal-Wallis test). A similar result was observed for LNM as compared to benign tissue ($p < 0.001$), with a weaker positive staining observed in healthy tissues (37.70 ± 9.06%). However, there was no significant difference in the level of staining between VSCC and LNM, although mean staining intensity in the LNM tissues was higher than VSCC samples (59.20 ± 16.20%). Comparison of the ACTB peptide *m*/*z* 1198.687, detected by MALDI-MSI, and results of the IHC analysis from two representative patient samples for ACTB are shown in ***Figure 3.8***.

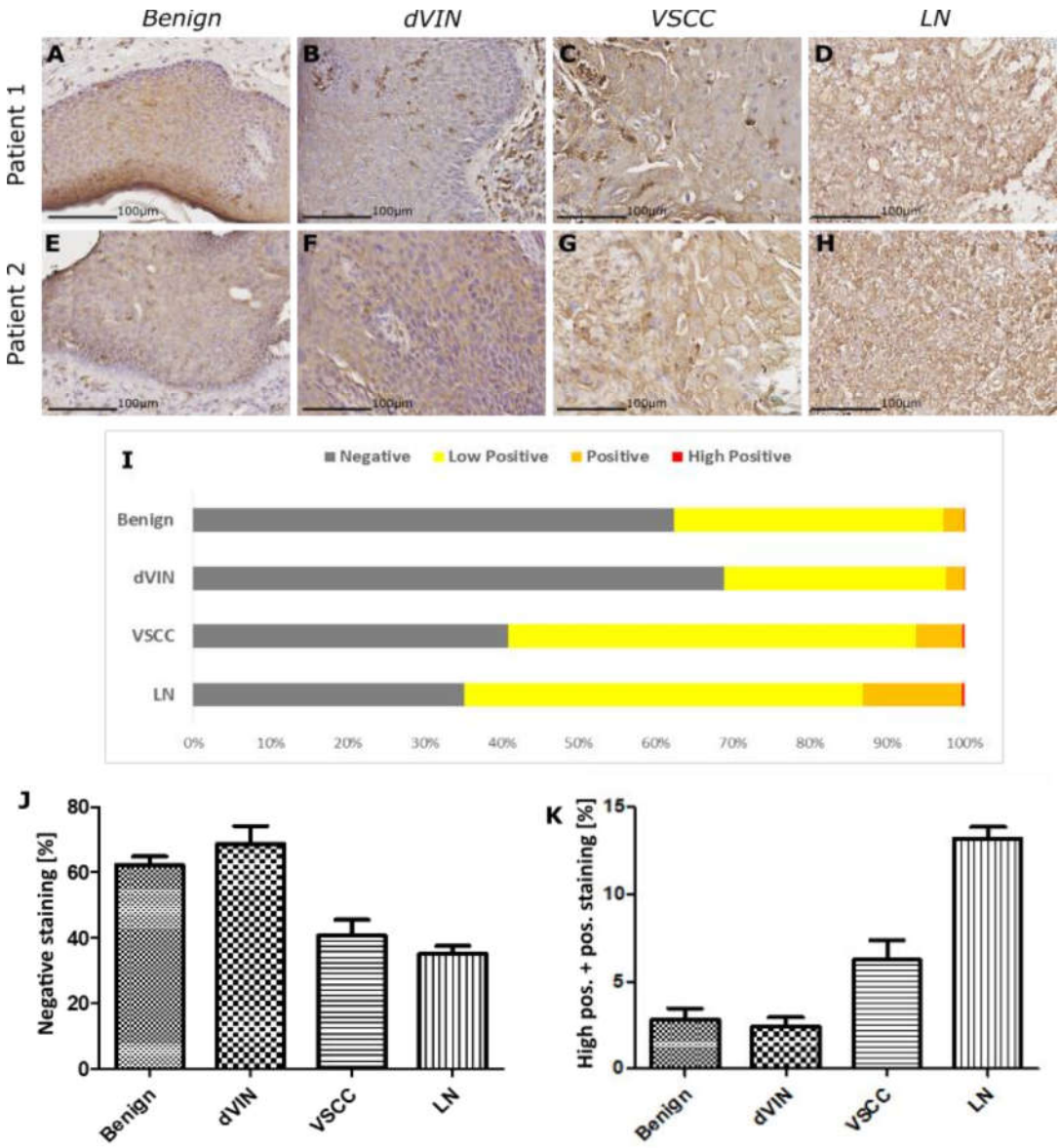


Figure 3.7: IHC analysis of ACTB on two representative vulvar cancer patient samples containing regions of benign epithelium, dVIN, VSCC and metastatic LN, at 40× magnification. (A) ACTB staining of benign epithelium; (B) dVIN; (C) VSCC from a single section; (D) LN from the same patient; (E) ACTB staining of benign epithelium; (F) dVIN; (G) SCC from a single section of a different patient; (H) metastatic LN from another patient; (I) Quantitative analysis of ACTB staining across tissue types from cohort patients (n=4) was performed using IHC profiler-Image J. For each tissue section, three representative photo-micrographic images at 40× magnification were used and each image was assigned a staining score of negative, low positive, positive and high positive; (J) Percent of negative staining was significantly different among benign, dVIN, VSCC and LN in KW test ($p< 0.001$); (K) Percent of positive and high positive staining was significantly among four tissue types in KW test ($p< 0.001$).

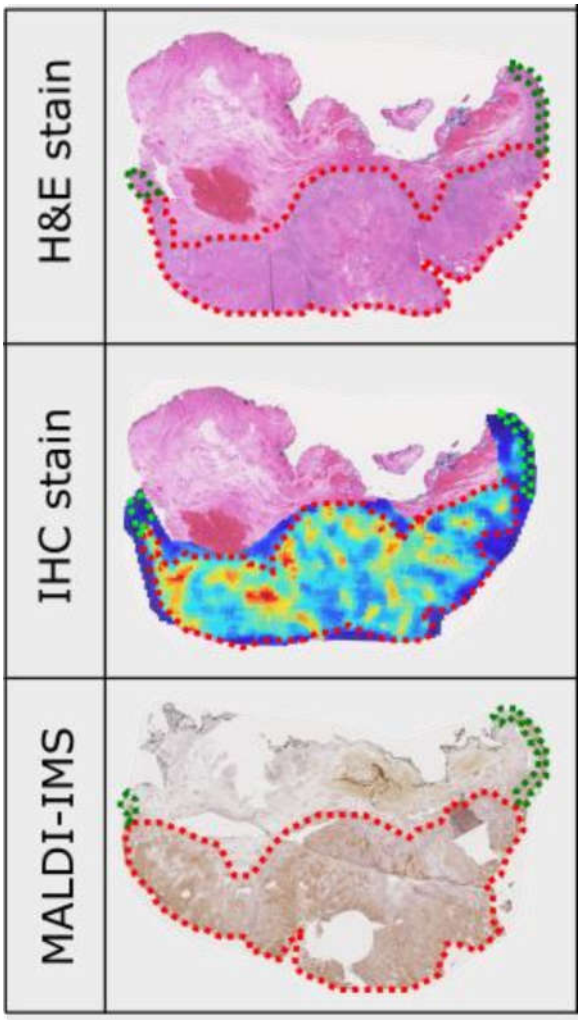


Figure 3.8: Comparison of expression pattern observed for the ACTB peptide, *m*/*z*1198.687, in the healthy epithelium (green) and VSCC (red) samples using MALDI-MSI ion intensity map (middle), H&E staining (top) and ACTB IHC staining (bottom) on consecutively cut vulvar tissue sections.

3.3.2.4 Verification of ACTB by data-independent acquisition *nano*-LC-ESI-MS/MS

Verification of the MALDI-MSI results for ACTB was further carried out using the data independent *nano*-LC-ESI-MS/MS on consecutive tissue sections from the same four HPV-independent vulvar cancer patients used in the MALDI-MSI analysis described above. In order to verify the differential expression of ACTB across the vulva cancer developmental stages, data independent *nano*-LC-ESI-MS/MS was performed on laser microdissected tissues as described in section 3.2.7. Summed area intensities for the analyzed peptides were calculated from y and b ions 3 to 6 (starting from ion 3). In total, six ACTB peptides, including those two tryptic peptides identified in *in-situ* MS/MS, were detected in data-dependent acquisition *nano*-LC-ESI-MS/MS were used in the spectral library peptides (including the two tryptic peptides identified by *in-situ* MS/MS, and four detected by data-dependent acquisition *nano*-LC-ESI-MS/MS) and candidates for quantification. Data for analysis was performed in the MS software Skyline MS software (v3.4). The summed intensities of these six ACTB peptides in the four vulvar tissues of each cohort patient are shown in ***Figure 3.9***.

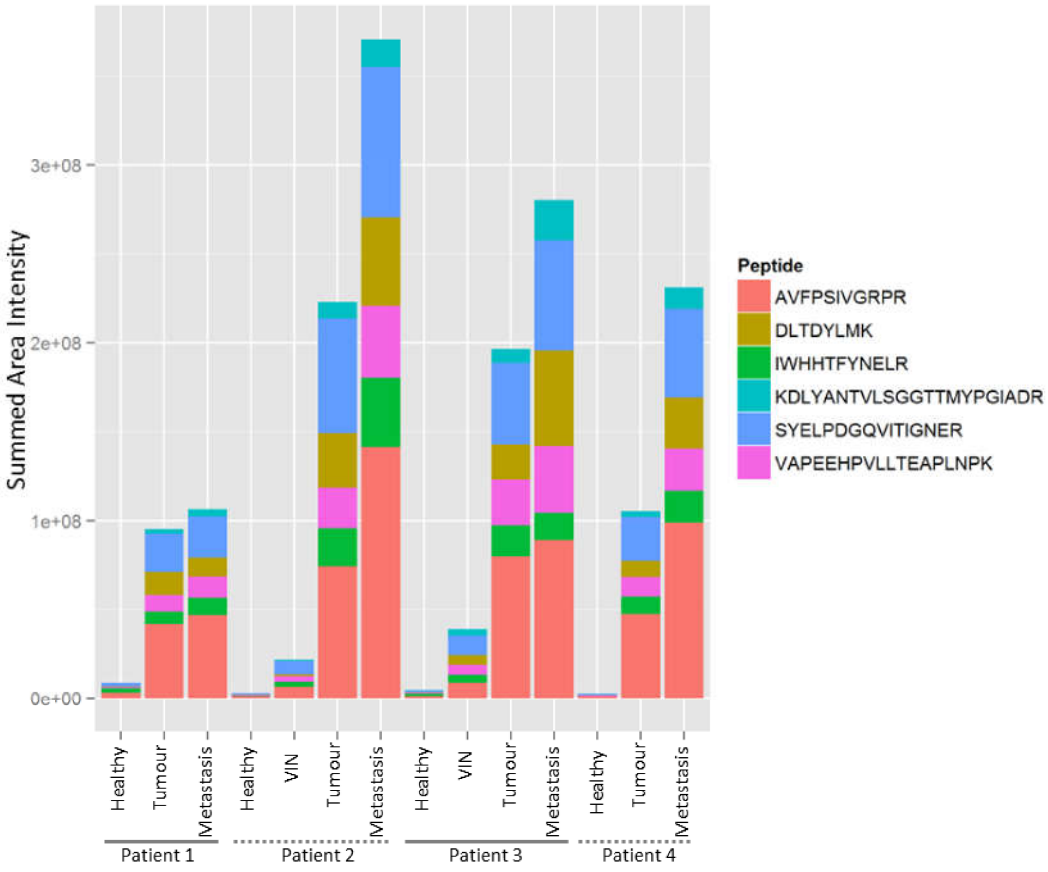


Figure 3.9: Quantification of ACTB by data independent *nano*-LC-ESI-MS/MS. Six ACTB peptides were used for quantification. Fragment area intensities of quantified peptides (colour coded) were summed for each patient.

3.4 DISCUSSION

Due to the relative rarity of vulvar cancer, characterization of the disease at the molecular level has been poor. A greater understanding of vulvar cancer progression is required, but especially with regards to dVIN as the precursor lesion to VSCC, which is known to have a high level of malignant potential (Balluff et al., 2011a) and is difficult to diagnose (Endo et al., 2015, Shin et al., 2014). Generally, dVIN is characterized by thickened parakeratotic epithelium with elongation and anastomosing rete ridges, little to no atypia above the basal or parabasal layers of the epidermis, with the basal cell layer containing abnormal squamous cells undergoing atypical mitosis and high levels of epithelial cell differentiation in the upper epithelium (Shin et al., 2014, Rothschild et al., 2006, Bharadwaj and Prasad, 2002). The aim of this study was to *in-situ* characterize the dVIN lesions at the peptide level using MALDI-MSI for comparison to the healthy vulvar epithelium and VSCC.

MALDI-MSI was carried out on tissue sections containing regions of healthy vulvar epithelium, dVIN and VSCC from six late-stage carcinoma patients. Thirty-one peak groups, representing the *in-situ* imaged peptides, could be reproducibly detected in the majority of acquired spectra (***Table 3.2***). The median log intensity for each of the 31 peak groups was calculated for the regions of healthy epithelium, dVIN and VSCC, and the maximum difference in intensity between the tissue types was calculated. This revealed 19 of the peak groups to have a change in intensity of ≥ 1.4 fold across the three tissue types, of which six matched to unique peptides from the protein CK5 (***Table 3.3***).

CK5 is a 58 kDa intermediate filament protein that dimerizes with cytokeratin 14 to form the cytoskeleton of basal epithelial cells (Yamaguchi and Condeelis, 2007). CK5 is known to have high levels of expression in squamous cell carcinomas and can be used as a marker of differentiation, with weaker to no expression in adenocarcinomas (Chu and Weiss, 2002a). Given squamous cell carcinomas account for over 90% of all vulvar cancers and that CK5 is a structural protein of basal epithelial cells (Khanna et al., 2016b), the detection of CK5 by MALDI-MSI was to be expected. A study characterizing the expression of CK5 in serious gynecological carcinomas of the ovaries, endometrium, fallopian tube, primary peritoneum and cervix observed the highest level of CK5 expression in the cervical and ovarian serous

carcinomas, with little to no reactivity in the endometrial, fallopian tube, or primary peritoneum serous carcinomas (Nofech-Mozes et al., 2008).

In order to determine if there was a genuine difference in CK5 levels across the three tissue types, IHC staining was performed on vulvar tissue from eight VSCC patients. Of these eight tissues, three had been previously analyzed in the MALDI-MSI experiments, six were moderately differentiated, one was well differentiated and one was poorly differentiated. A degree of keratinization was reported for all of the VSCC with the exception of the poorly differentiated tumour and one of the moderately differentiated tumours. Generally, the expression of CK5 appeared to be diffuse and cytosolic across the three tissue types, with no distinct tissue-specific staining patterns observed. IHC analysis revealed a significant increase in the level of CK5 staining in the dVIN as compared to the VSCC legions, and in the healthy epithelium as compared to the VSCC regions (***Figure 3.2***). There was a slight increase in the level of CK5 staining detected in the dVIN as compared to the healthy epithelium, but the difference was not significant.

To verify the spatial expression pattern of the peptides detected by MALDI-MSI, consecutive tissue sections that had been H&E stained and analyzed by CK5 IHC were compared to ion intensity maps generated for the CK5 peptides from the MSI experiments. A representative example of H&E staining, CK5 IHC, and an ion intensity map for the CK5 peptides SFSTASAITPSVSR and TTAENEFVMLKK, both *m*/*z* 1410.72, from consecutive vulvar tissue sections is shown in ***Figure 3.3***. The intensity and spatial distribution of the CK5 peptides as detected by MALDI-MSI corroborated well with the CK5 immunostaining observed. These results demonstrate the power of MALDI-MSI in the ability to combine classical mass spectrometry with histological tissue analysis. This study provides an insight into the molecular fingerprint of the VSCC precursor lesion dVIN, which appears to more closely resemble healthy epithelium than is does cancerous tissue, even in later stage patients.

All *in-situ* identified proteins found to be differentially expressed across tissue types have been implicated in cancer previously, with most known to be involved in chromatin and cytoplasmic filament restructuring processes which could facilitate reproduction and movement of cervical cancer cells (Barkati et al., 2013, Endo et al., 2015, Shin et al., 2014). In order to solidify the validity of the MALDI-MSI technique, the spatial expression pattern and differential abundance of ACTB across the VSCC stages was confirmed by IHC and data-independent

nano-LC-ESI-MS/MS. The relative abundance of ACTB across the different tissues, in particular between healthy/dVIN and VSCC/LNM regions, matched differences observed in signal to noise ratios of ACTB peptides detected by MALDI-MSI. The β-actin also called, actin, cytoplasmic 1, is one of the most conserved proteins that widely expressed in almost all eukaryotic cells. As a non-muscle cytoskeletal member, it has been reported directly involved with cell mobility, cell-cell/matrix interactions, cellular morphology and cytokinesis (Small et al., 1998). Previously, the expression level of ACTB was believed not affected by cancer progression and frequently used as an endogenous control in the protein quantification (Rubie et al., 2005, Kheirelseid et al., 2010). However, many studies have shown the expression level of β-actin increased in metastatic tumour cells, while the remodelling of actin filaments has also been highlighted as the key biological changes during malignant cancer transformation and invasion (Yamaguchi and Condeelis, 2007, Yamazaki et al., 2005, Rao and Li, 2004). For example, the remodelling of actin-filament could deliver a high motility and lower cell-cell adhesion for high-grade tumours of breast cancer (da Silva et al., 2015). In this study, the MALDI-MSI results revealed ACTB peptide peaks with higher abundances in VSCC and LNM than other tissue types. Further immunochemistry stains validated that conclusion that higher ACTB levels were strongly associated with primary and secondary tumours of the HPV-independent vulvar cancer. In addition, many actin-remodelling regulation proteins, including cadherin1, myosin II, Arp2/3 and vimentin, were also identified from the LNM samples of cohort patients (***Supplementary S3.4***), but their peptides did not show significant differences between VSCC and other vulvar tissues in the MALDI-MSI. Therefore, further analysis targeting these actin-remodelling regulation proteins will provide a better understanding of the actin filament change in the vulvar cancer progression.

3.5 CONCLUSION

In summary, MALDI-MSI could be used to highlight proteomic differences among benign epithelium, dVIN, VSCC and LNM in HPV-independent vulvar cancer patients. In the published manuscripts of this chapter, a decrease of CK5 peptide intensity level was observed in VSCC compared with dVIN and benign epithelium, while further analysis also showed a higher ACTB abundance in metastatic LN and VSCC than other tissue types. IHC staining of CK5 and ACTB antibodies validated protein expression changes detected in the MALDI-MSI experiment and the increase of ACTB was further validated DIA based LFQ analysis. Until now, this is the first time MALDI-MSI performed on vulva cancer study. Importantly, this study provides an insight into the molecular fingerprints of different vulvar cancer tissue types and offer a complete snapshot of disease progression.

This page intentionally left blank

CHAPTER 4: IDENTIFICATION OF PREDICTIVE PROTEIN MARKERS OF VULVAR CANCER LYMPH NODE METASTASIS

4.1 INTRODUCTION

Vulvar cancer, the fourth most common gynecological cancer globally, affects the outer part of the female reproductive system (Board, 2016). Over 90% of vulvar cancer cases are vulvar squamous cell carcinomas (VSCC), which can be further divided into two etiological subtypes: HPV-dependent and HPV-independent vulvar cancer (Hørding et al., 1993). HPV-dependent vulvar cancer patients normally have a better prognostic outcome and a very low risk of distant metastasis (Canavan and Cohen, 2002, Khanna et al., 2016a). Moreover, because the virus-induced carcinogenesis has been well understood and a vaccine has been widely made available, HPV-dependent vulvar cancer can be easily prevented by HPV vaccination. However, 60 to 70 % of all vulvar cancer patients, including patients diagnosed at postmenopausal age or with a history of precancerous lesions of differentiated VIN (dVIN) or lichen sclerosis (LS), present with HPV-independent vulvar cancer and experience worse clinical outcomes than HPV-dependent vulvar cancer patients (van de Nieuwenhof et al., 2009b). Currently, the presence of lymph node metastasis (LNM) is an important prognostic factor for vulvar cancer patients and previous statistical analysis showed, that the number and size of positive lymph nodes are correlated with the recurrence and survival rates (Oonk et al., 2010a, Woelber et al., 2012). In general, the 5-year survival rate of vulvar cancer patients with negative LNM is 86%, compared to 57% when regional metastasis is detected and 17.4% for patients with distant metastasis (Board, 2016, Howlader N, 2017). To date, there is no reliable LNM prediction method available for vulvar cancer patients other than invasive lymph node biopsy, and surgical removal of inguinal lymph nodes with additional radiotherapy is mandatory for almost all HPV-independent vulvar cancer patients diagnosed from FIGO stage IB to IV(Origoni et al., 1992, Board, 2016, Pecorelli, 2009). However, the risk of LNM for vulvar cancer patients from stage IB and II is significantly overestimated; with only 36% patients presenting with LNM, meaning that two-thirds of vulvar cancer patients receive unnecessary treatments (AIHW, 2016, Howlader N, 2017).

To classify VSCC patients with and without LNM, research has focused on histopathological parameters and other clinical factors, including tumour grade, patient's age at the time of diagnosis, the presence of precursor lesions, such as dVIN and LS, depth of stromal invasion, and lymphovascular and perineural invasion (Board, 2016). It has been suggested that a large tumour

size with deep stroma invasion correlated with higher metastatic risk. In addition, moderate and poorly differentiated tumours are more often associated with LNM than well-differentiated primary tumours (Davis et al., 1986). Currently, the most accurate way to determine LNM is the review of material from a lymph node biopsy by a trained pathologist (Oonk et al., 2006). Recently, the identification of potential molecular predictors, such as p53 and $p14^{ARF}$, has been proposed as a non-invasive and sensitive vulvar cancer patient LNM classification approach in addition to the lymph node biopsy result (Lerma et al., 1999, Knopp et al., 2006). Previous immunochemistry studies showed increased levels of tumour proliferation marker Ki67 in HPV-dependent vulvar cancer samples with positive lymph nodes (Gordinier et al., 1997, Liegl and Regauer, 2006, Vanin et al., 2002). Another protein biomarker, squamous cell carcinoma unique antigen (SCC-Ag), was reported with higher abundance in primary tumours of patients with distant metastasis (Hefler et al., 2005). However, those protein predictors mentioned above are still unable to discriminate VSCCs with adjacent LNM from non-metastatic HPV-independent tumour samples. The previous study has assembled a set of VSCC samples and using a novel mass spectrometric approach, aim to identify biomarkers which could be used to provide a more robust vulvar cancer patient classification system.

Matrix-assisted laser desorption/Ionization mass spectrometry imaging (MALDI-MSI) is an *in-situ* proteomics profiling technique that allows spatial quantification of hundreds of tryptic peptides and even direct identification from non-homogenized tissue samples. In addition, because the spatial distribution of potential tissue biomarkers is maintained, it can be compared to tissue structures observed using generic H&E stain and its annotation by pathologists. A recent study reported an 88% accuracy in classifying endometrial cancer patients with and without LNM using MALDI-MSI (Mittal et al.), indicating this approach is feasible. In this chapter, two vulvar cancer tissue microarrays (TMAs) containing tissue cores from 16 LNM positive and 12 LNM negative primary tumour samples were analyzed using MALDI-MSI. Proteins with potential discriminative ability between LNM positive and LNM negative samples in MALDI-MSI results were revealed by Receiver Operating Characteristic (ROC) analysis and Canonical Corresponding Analysis and Linear Discrimination Analysis (CCA-LDA) classification (Winderbaum et al., 2016). Identified protein predictor candidates were validated by immunohistochemistry (IHC) on consecutive TMA

sections and further confirmed by *nano*-LC-ESI-MS/MS based label-free quantification (LFQ) analysis of six LNM positive and six LNM negative patient samples.

4.2 MATERIALS AND METHODS

4.2.1 Sample Cohort and Tissue microarray construction: 28 FFPE tissue blocks of vulvar squamous cell carcinoma were collected and used to construct two TMAs (SA Pathology). The TMA annotations and histopathological details are summarized in Chapter 2.2.1.1. In this experiments, a total of 28 HPV-independent VSCC patients, 16 patients with LNM positive and 12 patients with LNM negative were incorporated in two TMA blocks, named TMAI and TMAII, sample arrangements within the TMAs are shown in Chapter 2.2.1.2.

4.2.2 MALDI-MSI Sample Preparation and Acquisition: MALDI-MSI was performed on two TMAs in duplicate runs as previously described in Chapter 2.2.1.9. In brief, TMAs were sectioned and mounted on ITO slides. Tissues were deparaffinized, then subjected to heat-induced citric acid antigen retrieval followed by digestion with trypsin gold. Peptide Internal calibrants and CHCA matrix were overlaid onto the tissue sections. An ultrafleXtreme MALDI-TOF/TOF MS system was used for MS and *in-situ* MS/MS data acquisition (Chapter 2.2.2.2).

4.2.3 MALDI-MSI Data Analysis: Acquired spectra were recalibrated using the internal calibrants and pre-processed using flexAnalysis (v4.0) and flexImaging software (v4.0) in order to produce peak lists. During the pre-processing the Snap Algorithm with a signal to noise ratio of 2, a baseline subtraction was performed using the TopHat method, and baseline smoothing was performed using the Gauss algorithm. VSCC specific regions of interest (ROI) were selected and exported into flexImaging (v4.0). Peaklists from all ROI were combined and density-based clustering (DBSCAN*, with an epsilon of 0.02 and a minimum density of 500 peaks) was used to cluster peaks into peak groups based on their *m/z* values. Only peakgroups containing at least 1000 peaks were considered for further analysis. The median log intensity between LNM positive and negative was calculated for each peak group. The raw data was further analyzed using the SCiLS lab software where the processing steps of baseline removal and normalization were carried out. Ion intensity maps and ROC curves for the peptides of interest were generated using the SCiLS lab software (Chapter 2.2.3.3).

4.2.4 CCA-LDA Classification: A detailed description of the canonical correspondence analysis with Linear discriminant analysis (CCA-LDA) method has been published previously (Winderbaum et al., 2016). In brief, arbitrary peak binning with a width of 0.25 Da were used to discretize *m/z* domains in order to group peaks from MALDI-MS. The intensity or SN levels of peaks in each of these defined *m/z* bin classifier were transformed and averaged across the annotated ROIs for each sample. These averages were assembled into a new dataframe with columns representing each patient for classification and rows corresponding to bins. Corresponding bins were then ranked using the CCA based method for their ability to distinguish between LNM positive and negative samples. A CCA dimension reduced submatrix consisting of the top-ranked rows was then analyzed using linear discriminant analysis (LDA) in order to predict lymph node status of the 28 patients from duplicated experiments. Selective bins were then ranked using the developed LDA based method for their ability to distinguish between primary carcinomas with and without lymph nodes metastasis. A majority rule was used to combine the data and top rank bins were selected as candidates for further validation.

4.2.5 Identification of *m/z* values by *in-situ* MALDI MS/MS and *nano*-LC-ESI-MS/MS: Aim to achieve 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 (v2.3.2) as previously described (Chapter 2.2.2.4) . For matching back to peptide sequences obtained by data-dependent acquisition, *nano*-LC-ESI-MS/MS, ROIs were collected from tissue sections, subjected to HIAR, and digested with trypsin. The *nano*-LC-ESI-MS/MS was performed using an Ultimate 3000 RSLC system coupled to an Impact II QTOF mass spectrometer via an Advance CaptiveSpray source. Peptide separation was performed using a 75µm C18 column over 180 min. MS scans were acquired in the mass range of 300 to 2200 *m/z* in a data-dependent fashion using Bruker's Shotgun Instant method (Chapter 2.2.2.5). Acquired spectra were subjected to Compass DataAnalysis for OTOF (v1.7). Processed spectra were then submitted to MaxQuant (v1.6.0.1) for identification using the integrated peptide search engine Andromeda. Search parameters were as follows; SwissProt Homo sapiens database search, trypsin/P with up to 2 missed cleavages fixed modification of carbamidomethylation of cysteine, variable modification of oxidation of methionine, N-terminal acetylation, and glutamine to pyroglutamate as variable modifications.; 40ppm MS/MS mass tolerance and 1% False Discovery Rate (FDR). The peptide identification

cut-off was set with MaxQaunt PEP value less than 0.01. The protein LFQ intensities of data-dependent data acquisition LC-MS/MS were automatically calculated when the MaxQuant LFQ tool was selected. In general, the LFQ intensities were normalized based on extract-ion-current (XIC) among different comparing runs using the internal MaxLFQ with minimum LFQ count ration of 2 (Chapter 2.2.3.5). Matching between the MALDI-MSI and *nano*-LC-ESI-MS/MS was done by comparing the experimental *m/z* values of the *nano*-LC-ESI-MS/MS sequenced peptides to the *m/z* values from the compiled peak bins in R (Chapter 2.2.4.5).

4.2.7 Immunohistochemistry: For the analysis of NDRG1, ANXA8, Col1, VIME, K2C5/6, and K2C1/K1C10 by IHC, two TMAs were analyzed as previously described in Chapter 2.2.2.9. Briefly, TMA sections were de-waxed, rehydrated and subjected to antigen retrieval, then incubated overnight at 4 °C with antibodies followed by incubation with secondary antibodies, streptavidin-HRP, and Diaminobenzidine. Slides were further stain with hematoxylin and scanned using a NanoZommer. The analysis was carried out using the freely available IHC Profiler within ImageJ, for each TMA sample core, three random representative images at 40x magnification were analyzed in Chapter 2.2.3.8.

4.2.8 Data-dependent acquisition based Label-free quantification: All tryptic peptide samples prepared before were diluted to 1µg/µL and 5µL solution was used for the data-dependent acquisition in Impact II™ QTOF mass spectrometer based on the method described above (Chapter 2.2.3.5). Processed data was exported to MaxQuant for LFQ intensities calculation and results were compared with the t-tests and interpreted as volcano plot in Perseus (Chapter 2.2.3.6). LFQ intensities of proteins with significant different abundances were plotted into heatmaps based on R scripts in Chapter 2.2.4.7.

4.3 RESULTS

4.3.1 Patient cohort

VSCC tissue blocks from 28 HPV-independent vulvar cancer patients (16 LNM positive and 12 LNM negative) were annotated by a trained pathologist, and used to establish two tissue microarrays (TMAs), named TMA I and TMA II, for this study. For each patient FFPE block, two 1.5mm sample cores contributed to the TMAs and were randomly within the TMAs, with benign epithelium and necrotic tissue added as additional controls. The layouts of two TMAs are shown in ***Figure 4.1***.

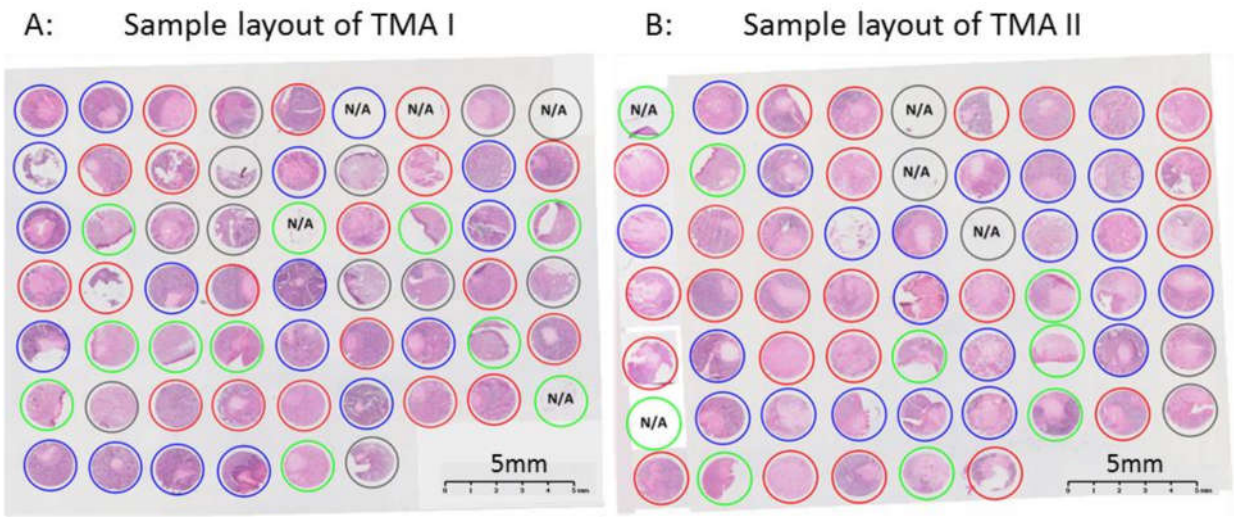


Figure 4.1: Sample layouts of VSCC tumour cores in two TMAs. Patient lab number, cohort number and tumour grades were listed in ***table 4.1***. The benign cores are highlighted in green circles, the primary tumour cores with negative LNM were highlighted in blue and positive LNM are in red circles. The grey circles are necrotic tumour cells or missing cores. Tissues cores lost in the sectioning were marked with N/A.

The cohort included 16 patients (stage III; median age 75) diagnosed with positive LNM and 12 negative LNM patients (stage IB; median age 64). Patient clinical and demographic information gathered in this vulvar cancer LNM study is listed in ***Table 4.1***.

Table 4.1: Demographic and clinical information of study cohort patients (n=28) with LNM positive (n=16) and LNM negative (n=14) vulvar squamous cell carcinoma.

Lab No.[a]	Tissue Type[b]	Age at diagnosis	VIN[c]	Lichen sclerosus[d]	FIGO Stage	Grade	Max diameter (mm)	Depth invasion (mm)	Lympho-vascular Space invasion (LVSI)	Perineural invasion	LNM Negative	LNM Positive
1	LNM -	37	1	0	1B	2	12	6	0	0	1	0
2	LNM -	47	1	0	1B	2	8	5	0	0	1	0
3	LNM -	48	1	Unknown	1B	1	32	6	0	0	1	0
4	LNM -	59	1	0	1B	2	27	4.4	0	0	1	0
5	LNM -	62	0	1	1B	2	35	14	0	0	1	0
6	LNM -	66	1	0	1B	2	40	7	0	1	1	0
7	LNM -	67	1	0	1B	2	20	6	0	0	1	0
8	LNM -	68	0	1	1B	1	25	6	0	0	1	0
9	LNM -	70	1	0	1B	2	35	5	0	0	1	0
10	LNM -	72	0	1	1B	2	50	11	0	0	1	0
11	LNM -	78	0	1	1B	1	12	2	0	0	1	0
I2	LNM -	48	1	Unknown	1B	1	32	6	0	0	1	0
13	LNM +	37	1	?	3	3	18	5	1	1	0	1
14	LNM +	50	0	0	3	2	65	11	0	0	0	1
15	LNM +	61	1	0	3	2	35	7	1	0	0	1
16	LNM +	63	0	1	3	3	17	2.5	0	0	0	1
17	LNM +	64	1	Unknown	3	2	42	5	0	1	0	1
18	LNM +	64	0	1	3	3	57	x	1	1	0	1
19	LNM +	66	1	0	3	2	20	8	1	0	0	1
20	LNM +	75	1	1	3	3	14	11.5	0	0	0	1
21	LNM +	76	0	0	3	2	30	10	0	1	0	1
22	LNM +	79	1	Unknown	3	2	65	9	1	0	0	1
23	LNM +	81	1	1	3	1	15	4	0	0	0	1
24	LNM +	81	0	1	3	2	20	8	0	1	0	1
25	LNM +	84	0	1	3	2	30	10	1	0	0	1
26	LNM +	87	0	1	3	3	90	4	0	0	0	1
27	LNM +	79	1	Unknown	3	2	65	9	1	0	0	1
28	LNM +	75	1	1	3	3	14	11.5	0	0	0	1
Control	Benign	77	1	1	1B	2	40	4	0	1	NA	NA
Control	Benign	76	1	0	3	3	45	10	1	1	NA	NA
Control	Benign	81	1	0	3	2	_	5.3	0	0	NA	NA
Control	Necrotic	79	1	Unknown	3	2	65	9	1	0	0	1
Control	Necrotic	37	1	?	3	3	18	5	1	1	0	1

a. Lab number of cohort patients in MALDI-MSI and LC-MS analysis.

b. There are 4 tissues types in this study, namely benign (green), primary tumours with negative (blue) and positive LNM (red), and necrotic tumour (black).

c. Some patients (1) reported with precursor lesion differentiated vulvar intraepithelial neoplasia (VIN)

d. Lichen sclerosis is found some cohort patients (1 and unknown)

4.3.2 MALDI-MSI

MALDI-MSI was carried out on TMA sections containing sample cores from four tissue types: benign epithelium, necrotic tumours and VSCCs diagnosed with positive and negative LNM from 28 patients mentioned above. MALDI-MSI was performed on two TMAs (TMAI and TMAII) in duplicate with 40000 to 50000 spectra collected from each TMA run (700 to 1100 spectra per sample core). Peakgroups, representing peaks from sample peptides across analysed TMA sections with *m/z* values less than 0.02 Da difference, were grouped using density-based clustering of detected peaks in R (Rodriguez and Laio, 2014), DBSCAN with a radius of neighborhood (eps) of 0.02 and a minimum of 500 points, with stringent criterion of a minimum of 1000 peaks per group. This criterion was initially put in place to ensure that peak groups included minimum numbers of peptides and could be reproducibly detected across a majority of spectra. With this criterion in place, 756 peakgroups were detected with 1000 to 30000 peaks in each group and widths of significant peakgroups ranged from ±0.027 to ±0.385 Da. Within each peakgroup, the abundance weighted mean (AWM) was calculated, representing the apex of each peakgroup and possible *m/z* value of the most abundant peptide within the group. Kruskal-Wallis test was performed on peaks from all four tissue types, then the Wilcoxon rank test was calculated to reveal peaks with significant intensity difference between LNM positive and negative samples. Significant rates of hypothesis tests were revised by Benjamini–Hochberg test based False Discovery Rate (FDR) control and resulted in 230 significant peakgroups in the Wilcoxon rank test. The detail of significant peakgroups is listed in ***Supplementary file S4.1***, including group number, minimum and maximum *m/z* values, mass gap, number of peaks, AWM, *p*-values of Kruskal-Wallis (KW) across four tissue types (Benign, necrotic, LNM positive and negative), Wilcoxon rank test (WX) between LNM positive and LNM negative samples, Benjamini–Hochberg (F) tests, and Signal to Noise (S/N) difference between LNM positive and negative samples. All peakgroups were sent to the SCiLS lab software to calculate the discriminative AUC score and results are shown *in* ***Supplementary file S 4.2*****.** Peakgroups of the highest 10 and lowest 10 AUC cores were selected and are shown in ***Table 4.2*** (AUC of 1 represents samples that were perfectly differentiated; an area of 0.5 represents no difference).

Table 4.2: MALDI-MSI peak groups ranked heuristically by the S/N or log intensity discriminative power between VSCC with positive and negative LNM. Peakgroups of the highest 10 and lowest 10 AUC scores are shown in this table.

Peak Group No.[a]	*m/z min*	*m/z max*	*mass gap*	*number of peaks*	*AWM*	*KW test p-value*	*WX test p-value*	*F-test p-value*[b]	*SN diff*[d]	*AUC score*[e]
140	874.332	874.496	0.164	19532	874.420	8.21E-103	6.35E-84	1.89E-100	12.55	0.641
144	875.374	875.521	0.147	14493	875.454	1.76E-112	6.63E-09	5.85E-184	2.1	0.640
604	1246.61	1246.698	0.088	7624	1246.661	1.44E-50	2.83E-48	3.41E-44	7.5	0.640
605	1247.63	1247.697	0.067	2888	1247.664	2.25E-06	2.65E-07	2.88E-06	1.3	0.639
146	876.338	876.482	0.144	17006	876.421	4.15E-27	7.74E-21	2.16E-05	4.25	0.637
603	1244.607	1244.671	0.064	3382	1244.641	5.69E-11	3.80E-11	1.81E-12	7.4	0.634
34	822.383	822.476	0.093	3853	822.432	7.91E-19	3.92E-21	7.07E-11	6	0.633
602	1243.595	1243.665	0.07	2432	1243.632	6.45E-17	4.01E-16	2.42E-18	6.8	0.632
516	1130.586	1130.669	0.083	2759	1130.639	2.56E-126	2.04E-12	5.44E-115	8.5	0.632
417	1062.486	1062.579	0.093	7187	1062.534	1.50E-62	1.95E-54	8.36E-27	19.4	0.630
					… …					
971	2426.162	2426.207	0.045	1289	2426.186	3.68E-13	4.45E-14	6.07E-15	2	0.335
395	1043.398	1043.649	0.251	82741	1043.544	4.28E-58	5.14E-46	5.95E-127	11.5	0.334
964	2278.059	2278.215	0.156	14031	2278.154	2.48E-19	6.22E-08	5.29E-20	3.1	0.333
967	2406.113	2406.200	0.087	4872	2406.165	1.38E-08	7.29E-07	1.84E-15	0.3	0.331
711	1392.678	1392.738	0.06	2544	1392.710	1.28E-45	1.45E-40	3.78E-21	26.55	0.331
374	1026.437	1026.620	0.183	53062	1026.520	3.20E-313	5.64E-41	1.80E-176	19.65	0.328
849	1667.682	1667.851	0.169	31758	1667.775	4.30E-158	7.31E-10	4.52E-81	11.1	0.327
775	1519.878	1519.905	0.027	1121	1519.887	2.23E-32	5.83E-13	0.844491	245.8	0.325
794	1549.64	1549.776	0.136	12287	1549.692	1.04E-46	5.92E-26	2.99E-58	18.05	0.314
386	1037.473	1037.635	0.162	57903	1037.565	8.40E-308	2.86E-57	6.68E-209	22.9	0.310

a. Peakgroups number from DBSCAN peakgrouping of comprehensive peaklists for TMA MALDI-IMS data.

b. The Wilcoxon signed rank test *p*-values within each peak group

c. *p*-value of Benjamini–Hochberg test of false discovery rate (FDR) control

d. SN difference was based on median SN change between LNM positive and negative peaks within each peak group

e. AUC score of each peak group was calculated from TIC normalized intensities of metastasis and non-metastasis patients.

Analysis of single peakgroups resulted in Receiver Operator Characteristic (ROC) curves with low calculated values, hence two top-ranked peakgroups were combined and tested in the SCiLS lab software (***Figure 4.2***). After testing several *m/z* values combinations, the regression model constructed by the peakgroups with the AWM of 874.420 and 1335.721 achieved only a very small

increase (AUC =0.679) compared to single predictive values (AUC=0.641 and AUC=0.625) in the ROC analysis (***Figure 4.2***).

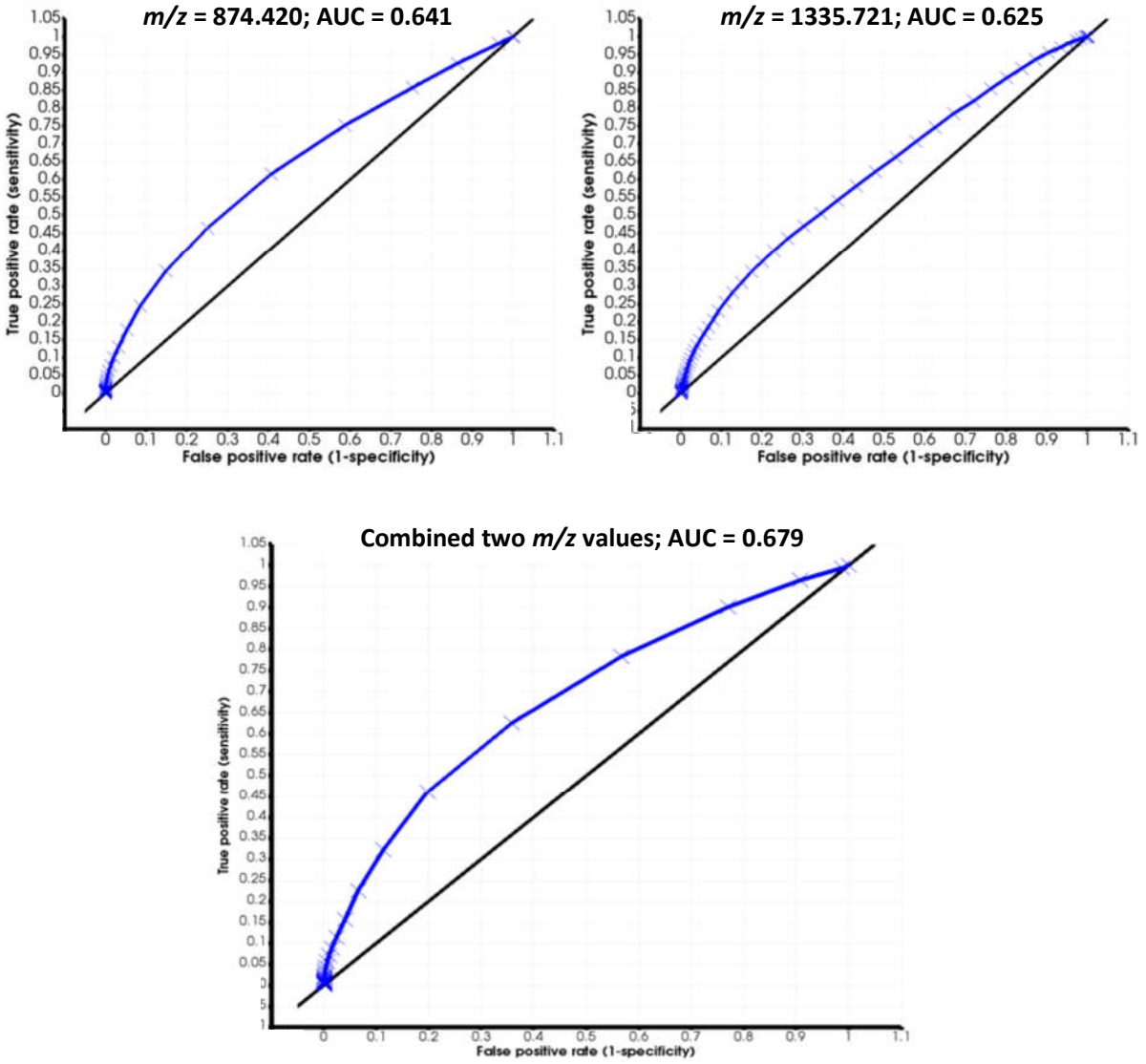


Figure 4.2: Receiver operating characteristic (ROC) curves for prediction of LNM condition by two *m/z* bins. Top left: ROC curve of *m/z* 874.420 between LNM positive and LNM negative samples, AUC= 0.641; Top right: ROC curve of *m/z* 1335.721 between LNM positive and LNM negative samples, AUC= 0.625; Bottom: Combined two peaks that achieved AUC = 0.679.

4.3.3 CCA-LDA classification

As AUC scores of individual *m/z* peaks could not achieve a good discriminative outcome (AUC<0.3 or AUC>0.7) between the LNM positive and LNM negative samples from the TMA cohort, Linear discriminative analysis (LDA) classification was used to classify VSCC patient samples based on the combination of multiple *m/z* peaks. Here, the LDA classification without any dimension reduction was performed on all spectra (TopHat baseline removal, TIC normalization, medium denoise and peak picking) using the internal classification tool of SCiLS Lab software (Chapter 2.2.3.3). Processed peaks were re-ranked according to statistical power in determining LNM positive and negative patients in LDA classification and MALDI-MSI spectra classification performance was evaluated by leave-one-out cross-validation (training data with n-1 patients and tested on remaining cohort samples) (Braga-Neto and Dougherty, 2004). Using the direct LDA method, MALDI-MSI results could classify the LNM status of TMA patients with an accuracy of 80.0% in leave-one-out cross-validation (***Supplementary file S4.2***).

Although this study had a relatively low number of cohort samples (n=28) in this vulvar cancer LNM study, the classification power of MALDI-MSI results could be improved by using canonical corresponding analysis (CCA), and then reduced the data dimensions before LDA classification. To achieve CCA-LDA based cohort sample classification, raw spectra were processed to peaklist in flexanalysis (v4.0), then arbitrarily grouped into 0.25 Da bin components and CCA-LDA classification were performed on annotated LNM positive and LNM negative patient samples using the published manuscript (Winderbaum et al., 2016). A Principal Component Analysis (PCA) based LDA classification was also performed on the vulvar cancer TMA data as a comparison method. The results from this analysis showed a 0 rate of misclassification could be achieved for the CCA-LDA when the best 14 bins were selected (a 0 rate of misclassification could also be achieved using 20 ~ 23 bins) while the highest misclassification of 11 samples occurred for CCA-LDA analysis when only one *m/z* bin was used. The misclassification rate of CCA-LDA results using different data types and numbers of bins is shown in the ***Supplementary file S4.8***.

To achieve the lowest misclassification outcome, at least 11 to 23 *m/z* bins were required as classifiers when their binary or log-transformed intensities were used in the CCA-LDA. In total, there were 20 lists of CCA-LDA bins could reach the 0 rate of misclassification (28 of 28 correct in leave-one-out cross-validation) in determining LNM conditions of the TMA samples. Details of different parameters, namely, Data-Type, Normalization, Empty-bin, Smooth, Annotation, Calibration, Spectra/Core, Spectra/Bin, and Shift, used in data dimensional reduction and classification of LNM patients are listed in ***Supplementary file S4.3*** and top 30 *m/z* bins were re-ordered according to their power in CCA-LDA classification (***Supplementary file S4.4***).

In this thesis, a list of the top ranked CCA-LDA classifiers constituted with 14 log-intensity type data *m/z* bins were selected for peptide identification and validation, those bins were centred at 1011.583, 1186.583, 847.333, 1117.583, 1391.833, 1547.833, 868.583, 1028.583, 1316.833, 1539.833, 999.583, 1043.583, 1335.833 and 841.333, ordered by classification power in CCA-LDA. Other classification bins not included in this list and their associated data types are listed ***Supplementary file S4.4***. The abundance weighted means (AWM) of candidate bins within the 0.25Da mass range were calculated using the same method mentioned above (Zhang et al., 2016a), and *p*-values of the Wilcoxon rank-sum test, AUCs of ROC analysis, and rank of bins in vulvar cancer patient sample CCA-LDA classification outcomes are shown in ***Table 4.3*****.** The prediction model based on two selective bins were also tested in the SCiLS lab software, which shown AUC values ranging from 0.446 to 0.573. Peptide identification of the remaining 14 *m/z* bins, as well as other bins from selective CCA-LDA lists from classification results, were compromised by matching back AWM values of *m/z* bins to protonated masses of identified peptides in *nano*-LC-ESI-MS/MS runs.

Table 4.3: Top-ranked *m/z* bins from selective CCA-LDA classification list of VSCC primary tumour TMAs. The *p*-values and significance in Wilcoxon rank sum test were calculated based on abundant weight means within the bin range, AUC scores and classification in metastasis and non-metastasis classification were also given.

Bin centre (*m/z*)	Bin margin	AWM within bin	AUC[a]	p-value of Wilcoxon rank sum test	Ranks of classifiers
1011.583	± 0.125Da	1011.651	0.56	0.0041	1
1186.583	± 0.125Da	1186.552	0.581	6.98e-08	2
847.333	± 0.125Da	847.371	0.47	1.21e-34	3
1117.583	± 0.125Da	1117.584	0.581	1.83e-31	4
1391.833	± 0.125Da	1391.741	0.43	3.40e-24	5
1547.833	± 0.125Da	1547.822	0.505	1.03e-24	6
868.583	± 0.125Da	868.444	0.586	6.68e-94	8
1028.583	± 0.125Da	1028.555	0.481	6.95e-65	10
1316.833	± 0.125Da	1316.733	0.543	0.0022	11
1539.833	± 0.125Da	1539.774	0.587	5.34e-08	12
999.583	± 0.125Da	999.481	0.556	5.81e-44	13
1043.583	± 0.125Da	1043.556	0.548	5.14e-46	14
1335.833	± 0.125Da	1335.721	0.623	1.12e-30	15
841.333	± 0.125Da	841.447	0.43	0.0074	16

a. AUC scores of bins were calculated from ROC analysis on peak intensities from metastasis and non-metastasis patient samples.

b. Classification power was ranked by loading power as components in CCA-LDA classification.

4.3.4 Identification of discriminative *m/z* values

To be able to identify all peptides observed during the MALDI-MSI run, an HPV-independent VSCC patient peptide library was established. This was achieved by analyzing a tryptic tissue digest by *nano*-LC-ESI-MS/MS proteomic profiling of FFPE tissue sections from the TMA patient cohorts. A total of 12 VSCC primary tumour samples, consisting of six LNM positive and six LNM negative vulvar cancer patient samples, were subject to protein extraction and trypsin digestion using filter-aided sample preparation (FASP) (Wisniewski et al., 2009). Digested peptides were analyzed on *nano*-LC-ESI-MS/MS and MS/MS spectra were searched against the UniProt human protein database in MaxQuant software. The peptide library of identified peptides is listed in ***Supplementary S4.5***. In summary, 3464 peptides were identified from VSCC patient samples past the 1% FDR threshold. 14 candidate *m/z* values from the CCA-LDA list and the bin containing the top discriminate *m/z* values were matched back to protonated masses of identified peptides using in-house R script (Chapter 2.2.4.5) and results are shown in ***Table 4.4*.**

Table 4.4: Identification of discriminative *m/z* values from MALDI-MSI results by matching a peptide library created from sample *nano*-LC-ESI-MS/MS and *in-situ* MS/MS proteomic profiling.

Protein			**Tandem MS/MS**					**MALDI-MSI**	
Protein Name	Change	Protein ID	Sequence	Unique Proteins	Measured Mass [M]	LC-MS/MS MaxQuant PEP	*In-situ* MS/MS Mascot Score	AWM [M+H]	mass error (ppm)
Annexin A8 (ANXA8)	Up	P13928	LIVALMYPPYR	Yes	1334.7421	0.001359		**1335.721#**	-15.81
ATP synthase subunit alpha (ATPA)	Down	P25705	TSIAIDTIINQK	Yes	1315.7300	1.86e-17		**1316.733#**	2.28
Collagen alpha-1(I) chain (CO1A1)	Up	P02452	GEAGPQGPR	Yes	867.4199	0.000323	25	**868.444*#**	27.78
			GPAGPQGPR	No	835.4404		39	**836.417**	-28.13
Collagen alpha-1(XII) chain (COCA1)	Down	Q99715	QGSGTTASR	Yes	846.3832	0.001826		**847.371*#**	-14.41
Collagen alpha-2(I) chain (CO1A2)	up	P08123	GPSGPQGIR	Yes	867.4563	1.54e-22		**868.444*#**	-14.18
			GPAGPQGPR	No	835.4404		39	**836.417**	-28.13
			GEAGAAGPAGPAGPR	Yes	1234.6148		31	**1235.630**	12.47
Collagen alpha-3(VI) chain (CO6A3)	Up	P12111	GDSIDQCALIQSIK	Yes	1546.7661	0.000736		**1547.822#**	36.14
Copine-3 (CPNE3)	Down	O75131	QAIVNASR	Yes	840.4454	0.005907		**841.447#**	1.90
Fatty acid-binding protein (FABP5)	Down	Q01469	ATVQQLEGR	Yes	1042.5407	1.9e-277	18	**1043.556*#**	14.68
Ig gamma-3 chain C region (IGHG3)	Down	P01860	SCDTPPPCPR	Yes	1185.4907	2.32e-11		**1186.552#**	51.71
Keratin, type I cytoskeletal 17 (K1C17)	up	Q04695	TIVEEVQDGK	Yes	1116.5663	3.13e-71		**1117.584#**	15.85
			EQVHQTTR	Yes	997.4587		29	**998.493**	34.79
Keratin, type II cytoskeletal 1 (K2C1)	up	P04264	QLDSIVGER	No	998.5033	3.4e-43		**999.481#**	-22.33
			FVSTTYSGVTR	Yes	1216.6121		41	**1217.633**	17.01
			SLVNLGGSKSISISVAR	Yes	1687.1253		69	**1687.955**	-100.00
Keratin, type II cytoskeletal 1b (K2C1B)	up	Q7Z794	VDTLTGEVNFLK	Yes	1334.7082	0.002528		**1335.721#**	9.59
Keratin, type II cytoskeletal 6B (K2C6B)	Down	P04259	QLDNIVGER	Yes	1042.5407	0.000815		**1043.556*#**	14.68
			LALDVEIATYRK	No	1390.7820	0.0004		**1391.741#**	-29.48
			SGFSSVSVSR	No	1011.4335		43	**1012.501**	67.13
			SGFSSISVSR	No	1025.4807		40	**1026.520**	38.42
			GSGGLGGACGGAGFGSR	No	1423.6910		82	**1424.656**	-24.79
			ASTSTTIRSHSSSR	No	1518.7560		31	**1519.756**	-0.07

			RGFSANSARLPGVSR	No	1573.8988		34	**1574.880**	-12.26
			ISIGGGSCAISGGYGSR	No	1597.7967		69	**1598.775**	-13.77
			SRGSGGLGGACGGAGFGSR	No	1666.8176		85	**1667.775**	-25.86
			AIGGGLSSVGGGSSTIKYTTTSSSSR	No	2417.1849		36	**2418.185**	0.12
L-lactate dehydrogenase A chain (LDHA)	Up	P00338	DYNVTANSK	Yes	1010.4669	0.003802		**1011.651#**	182.19
Prelamin-A/C;Lamin-A/C (LMNA)	Down	P02545	EGDLIAAQAR	Yes	1042.5407	0.001154		**1043.556*#**	14.68
N-myc downstream regulated gene 1 (NDRG1)	Up	Q92597	SHTSEGTR	Yes	873.3941	0.002981		**874.420#**	29.65
Protein S100-A7 (S10A7)	Down	P31151	SNTQAER	Yes	846.3832	1.18e-34		**847.371*#**	-14.41
Vimentin (VIME)	Down	P08670	SVSSSSYRR	Yes	1027.5069		22	**1028.555*#**	46.81
			ILLAELEQLKGQGK	Yes	1538.9032	0.002923		**1539.774#**	-83.96

* Peptides identified in both *nano*-LC-MS/MS and *in-situ* MS/MS sequencing

Peptide mass matched within CCA-LDA classification bins

In the selective list of potentially discriminative *m/z* values, 14 were successfully matched with peptides from the peptide library within 0.125 Da bin margin, except for *m/z* bins of 1028.583. However, the *m/z* bin of 1011.583 matched L-lactate dehydrogenase A chain which demonstrated a very large matching error over 0.125 Da between the AWM and the measured mass and hence was removed from the candidate lists. Three *m/z* bins were matched with more than one peptide sequence identified by LC-MS/MS within bin range and two *m/z* bins matched peptides shared by multiple proteins. To further confirm multiple matched and unmatched *m/z* values, *in-situ* MS/MS was performed on patient tissue sections as a direct protein identification method from their original spatial location complementary to the *in-solution* LC-MS/MS identification result described above. As a result, *m/z* bins 847.333, 868.583 and 1028.583 were identified as peptides from histidine-tRNA ligase, cytoplasmic, collagen alpha-1 or 2 chains, keratin, type II cytoskeletal 1, and vimentin from direct *in-situ* MS/MS sequencing of tissue sections. Tables of *in-situ* MS/MS identification results are listed in ***Supplementary S4.6***. However, *m/z* bin of 847.333 matched different proteins (S100A7 and collagen alpha-1(XII) chain), so it was rejected in this bin matching result. Peptides sequences from bin 868.583 were supported by both in-solution LC-MS/MS and *in-situ* MALDI-MS/MS results as confident matches. Ion intensity maps of identified protein candidates with their discriminative ROC curve from the TMA cohort are shown in the following figures.

Among the selective CCA-LDA classification bin list, *m/z* bin 1335.833, AWM at 1335.721 showed the highest AUC score of 0.625 between LNM positive and LNM negative samples (***Figure 4.3***). This *m/z* bin matched peptides from two proteins, keratin, type II cytoskeletal 1b of VDTLTGEVNFLK and Annexin A8 of LIVALMYPPYR, but keratin, type II cytoskeletal 1b was identified only in a few samples of the cohort, while expression of Annexin A8 was identified as a unique protein in more than half of patient samples in the TMA cohort. This peptide within 1335.833 bin displayed a higher intensity signal in most LNM positive sample cores, highlighted by red colour in ***Figure 4.3***, compared with median or lower peptide intensity reported among LNM negative patients. Results suggest Annexin A8 protein may be upregulated during vulvar cancer metastasis progress.

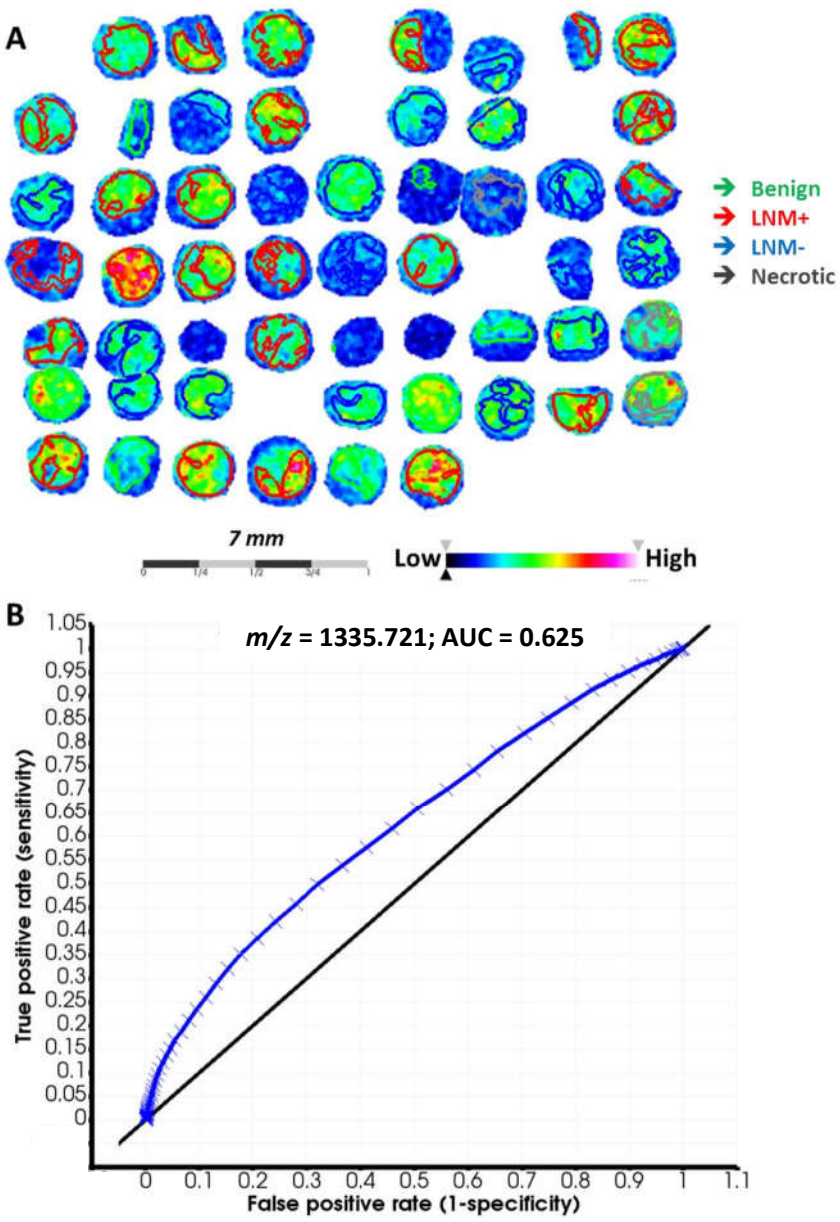


Figure 4.3: Representative MALDI-MSI images for *m/z* 1335.721. (A) Overview of one TMA slide, MALDI-MSI image of (intensity range from blue (lowest) to red (highest)). The samples belonging to different groups are indicated by different coloured circles: benign (green), LNM positive (red), LNM negative (blue) and necrotic tumour /not included in the study (grey). Tumour regions within samples were annotated by a pathologist (B) ROC curve of selective peptide *m/z* value showed an AUC of 0.625 between LNM positive and LNM negative vulvar cancer patient samples in the TMA cohort.

The *m/z* value of 868.444 with AUC score of 0.593 in the candidate list was matched back to Collagen type I (COL-1) subunit alpha 1 and 2 peptides, GPAGPQGPR and GEQGAPGPR, in both LC-MS/MS and *in-situ* MS/MS peptide identification. Although this CCA-LDA classification was shared by two possible peptide sequences, both came from subunits of the same protein, called collagen type I. The bin of collagen type I demonstrated higher ion intensity levels in LNM positive and necrotic sample cores compared with its intensity level in most of LNM negative samples (***Figure 4.4***).

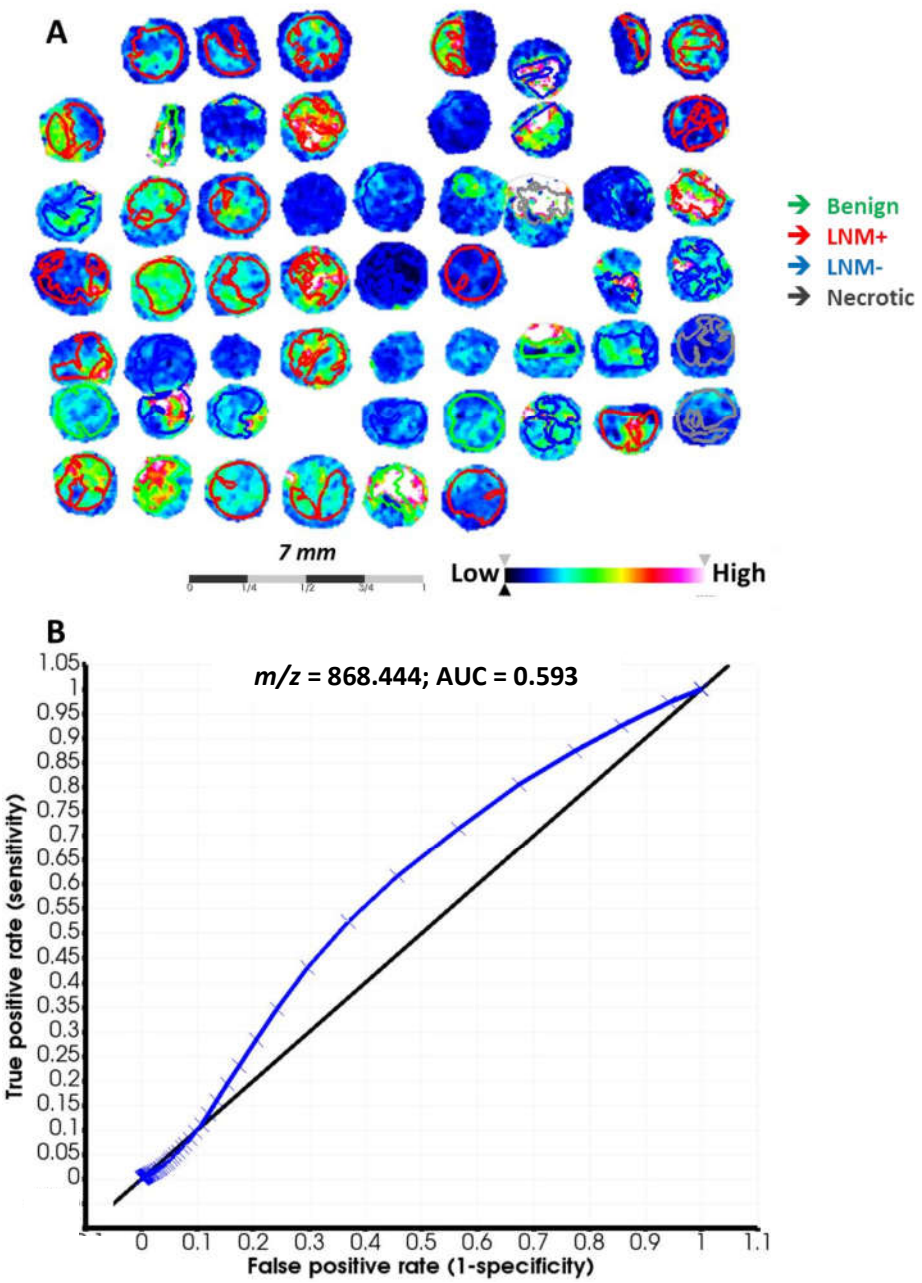


Figure 4.4: Representative MALDI-MSI images for peptide GEQGAPGPR as well as GEQGAPGPR, both matched the peptide bin of 868.583Da, AWM at 868.444. (A) Overview of one TMA slide, MALDI-MSI image of (intensity range from blue (lowest) to red (highest)). Samples belonging to different groups are indicated by different coloured circles: benign (green), LNM positive (red), LNM negative (blue) and necrotic/not included in the study (grey). Tumour regions within samples were annotated by a pathologist. B: ROC curve of identified Collagen type I peptide bin suggested higher abundance in LNM positive patient samples.

Although *m/z* bin of 1028.583 (AWM of 1028.555) did not match any peptide identified in the peptide library from patient samples in-solution peptide profiling, direct *in-situ* MS/MS on tissue sections suggested the dominant peptide within this bin could be the vimentin peptide SVSSSSYRR, whilst another *m/z* bin 1539.833 (AWM of 1539.774) was also identified as the vimentin peptide, ILLAELEQLKGQGK. The peptide ion intensity maps for these two peptides are shown in ***Figure 4.5***. These two vimentin tryptic peptide peaks were able to be detected in most of LNM positive and negative tumour sample cores, while other identified vimentin peptides did not match peaks from the MALDI-MSI results. ROC curves of those two peptides indicated relatively higher signal levels presenting in LNM negative patient samples.

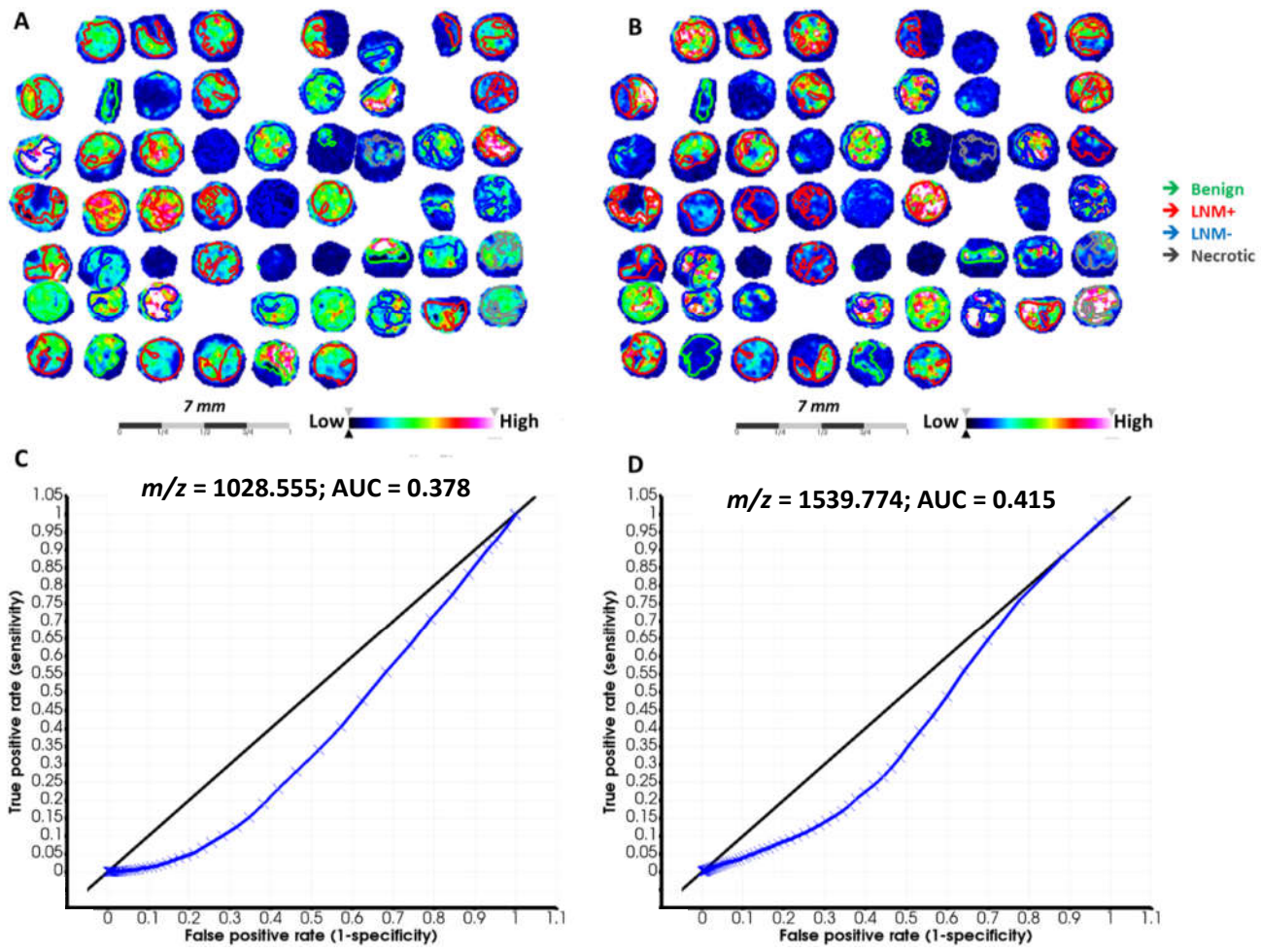


Figure 4.5: Representative MALDI-MSI images for *m/z* 1028.555 identified as peptide SVSSSSYRR (A). Peptide ILLAELEQLKGQGK matched *m/z* 1539.774 (B). A and C: Overview of one TMA slide, MALDI-MSI image of (intensity range from blue (lowest) to red (highest)). Samples belonging to different groups are indicated by different coloured circles: benign (green), LNM positive (red), LNM negative (blue) and necrotic/not included in the study (grey). Tumour regions within samples were annotated by a pathologist. C and D: ROC curve of identified vimentin peptide bin of peptide intensities.

The peptide TIVEEVQDGK from cytokeratin 17 (AWM of 1117.584) was the only confident identification (peptide PEP score <0.01) that matched *m/z* 1117.583 bin range. Although detection of cytokeratin 17 may be caused by potential contamination, many other cytokeratin 17 peptides were also identified in LC-MS/MS and *in-situ* MS (***Supplementary S4.6***) experiments. This *m/z* bin was highlighted in a majority of LNM positive sample cores and was difficult to detect in LNM negative patient samples (***Figure 4.6***). Results indicate that the expression level of cytokeratin 17 may be associated with lymph node status of vulvar cancer patients.

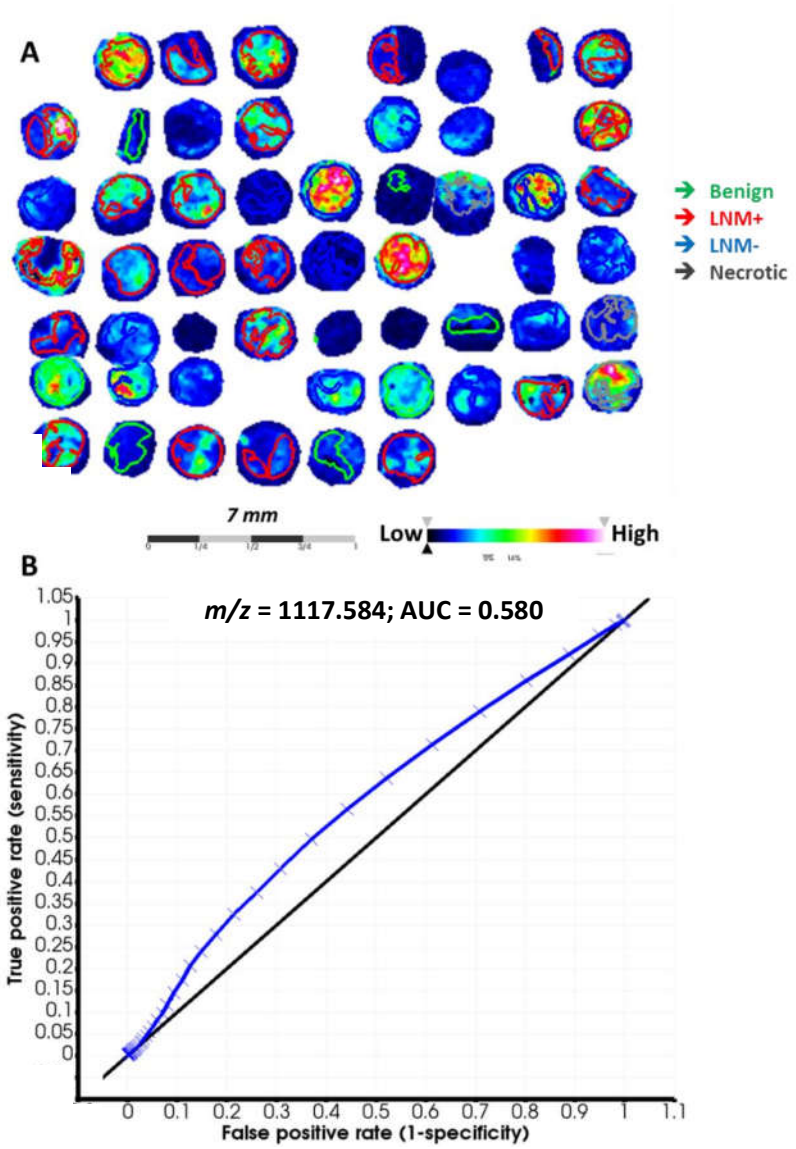


Figure 4.6: Representative MALDI-MSI images for cytokeratin 17 peptide TIVEEVQDGK *m/z* of AWM 1117.584. A: Overview of one TMA slide, MALDI-MSI image of (intensity range from blue (lowest) to red (highest)). Samples belonging to different groups are indicated by different coloured circles: benign (green), LNM positive (red), LNM negative (blue) and necrotic/not included in the study (grey). Tumour regions within samples were annotated by a pathologist. B: ROC curve of identified cytokeratin 17 peptide bin of peptide intensities.

The selective CCA-LDA bin centred at *m/z* 999.583 (AWM 999.481) was matched to the protonated mass of the peptide QLDSIVGER from cytokeratin 1, in the peptide library created by

nano-LC-ESI-MS/MS results. Expression of cytokeratin 1 was mainly observed among VSCC and necrotic tumour cores, rather than benign tumours (***Figure 4.7***). Some LNM positive samples were highlighted by strong signal levels of this peptide bin, where the remainder of positive node samples had very low-intensity levels.

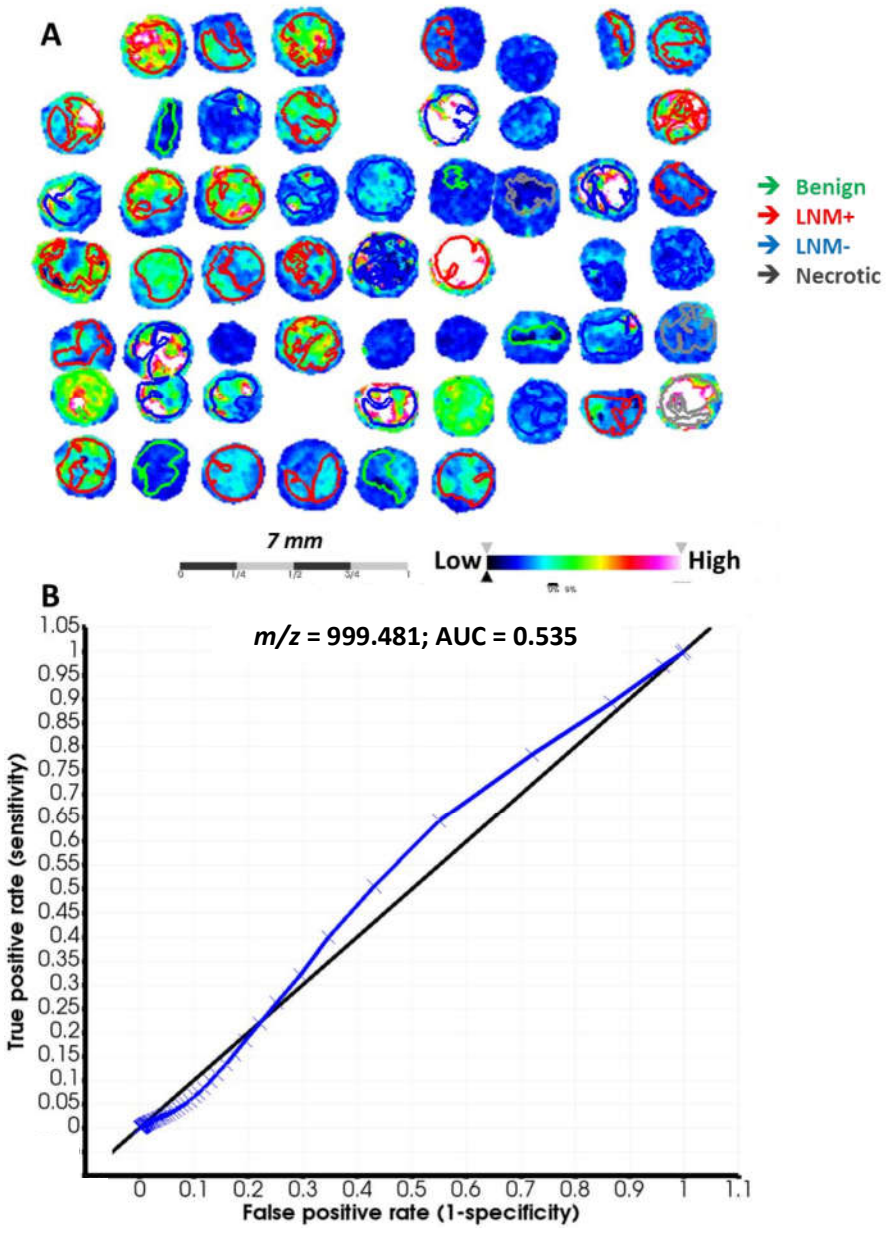


Figure 4.7: Representative MALDI-MSI images for cytokeratin 1 peptide QLDSIVGER *m/z* of AWM 999.481. A: Overview of one TMA slide, MALDI-MSI image of (intensity range from blue (lowest) to red (highest)). Samples

belonging to different groups are indicated by different coloured circles: benign (greiveen), LNM positive (red), LNM negative (blue) and necrotic/not included in the study (grey). Tumour regions within samples were annotated by a pathologist. B: ROC curve of identified cytokeratin 1 peptide bin of peptide intensities.

The classification bin centred at 1186.583 (AWM of 1186.552) was matched with peptide, SCDTPPPCPR, of Ig gamma-3 chain C in the library, which could be widely found in MALDI-MSI results of VSCC TMAs. However, the sum of the peptide AUC score between LNM positive and LNM negative patients was 0.399 in the TMA cohort, suggesting that a relatively stronger signal may present in LNM negative sample cores (***Figure 4.8***).

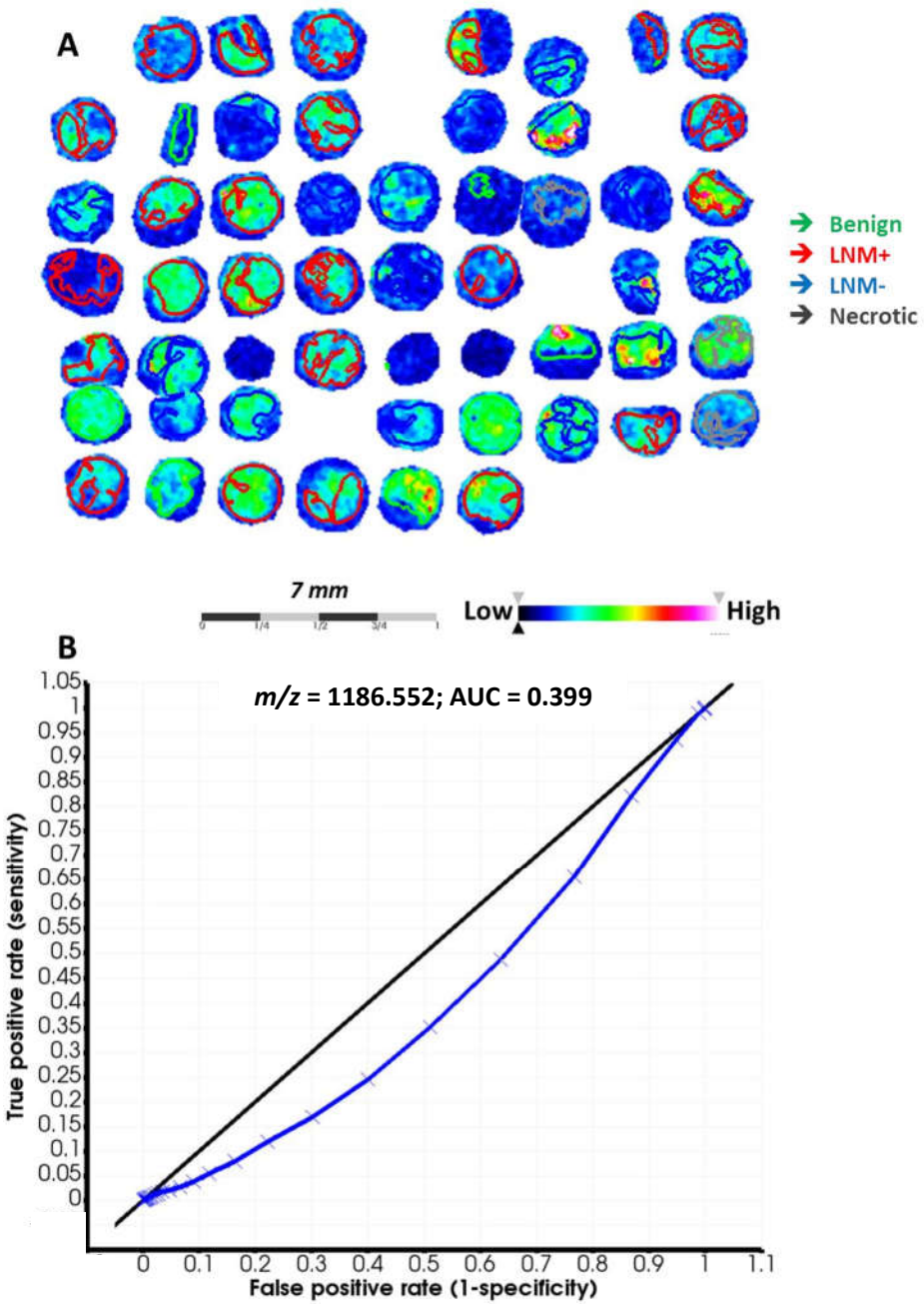


Figure 4.8: Representative MALDI-MSI images for Ig gamma-3 chain C region peptide SCDTPPPCPR *m/z* of AWM 1186.552. A: Overview of one TMA slide, MALDI-MSI image of (intensity range from blue (lowest) to red (highest)). Samples belonging to different groups are indicated by different coloured circles: benign (green), LNM positive (red), LNM negative (blue) and necrotic/not included in the study (grey). Tumour regions within samples were annotated by a pathologist. B: ROC curve of identified peptide bin of peptide intensities.

The top-ranked classification bin 1391.833 ± 0.125 Da and other large numbers of peptides were identified as peptides from cytokeratin 6 proteins in both peptide library and *in-situ* MS/MS identification. All *m/z* values that identified as cytokeratin 6 protein peptides displayed a relatively higher intensity on LNM negative samples compared with other TMA cores from LNM positive patient samples (***Figure 4.9***).

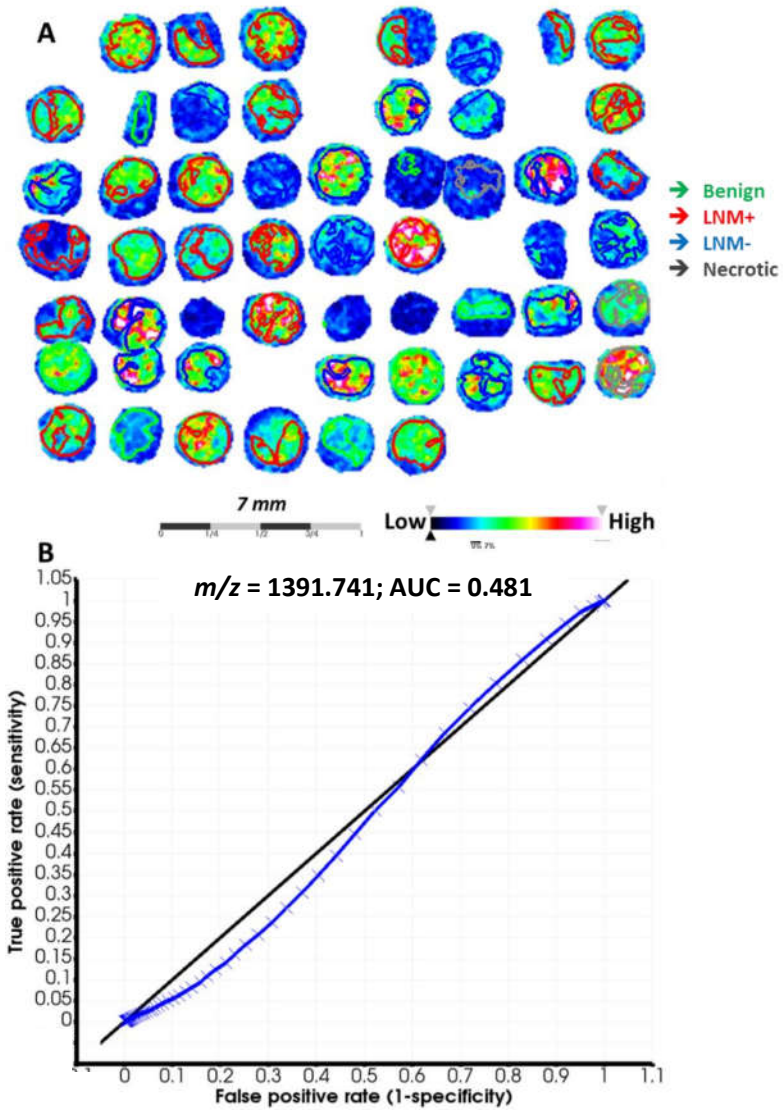


Figure 4.9: Peakgroup AWM 1391.741 from bin 1391.833 was identified as of LALDVEIATYRK of cytokeratin 6 with overall AUC= 0.481. A: Overview of one TMA slide, MALDI-MSI image of (intensity range from blue (lowest) to red (highest)). Samples belonging to different groups are indicated by different coloured circles: benign (green), LNM positive (red), LNM negative (blue) and necrotic/not included in the study (grey). Tumour regions within samples were annotated by a pathologist. B: ROC curve of identified cytokeratin peptide bin of peptide intensities.

Along with the best discriminative *m/z* bin 1335.833 from the selective classifier list, another *m/z* bin of 874.333 matched NDRG1 peptide SHTSEGTR (AWM=874.420) in binary CCA-LDA classification and demonstrated a fair discriminative power between VSCC from LNM positive and negative patient samples. The AUC score between two patient sample groups reached 0.641, which was higher than other MALDI-MSI classification bins (***Figure 4.10***).

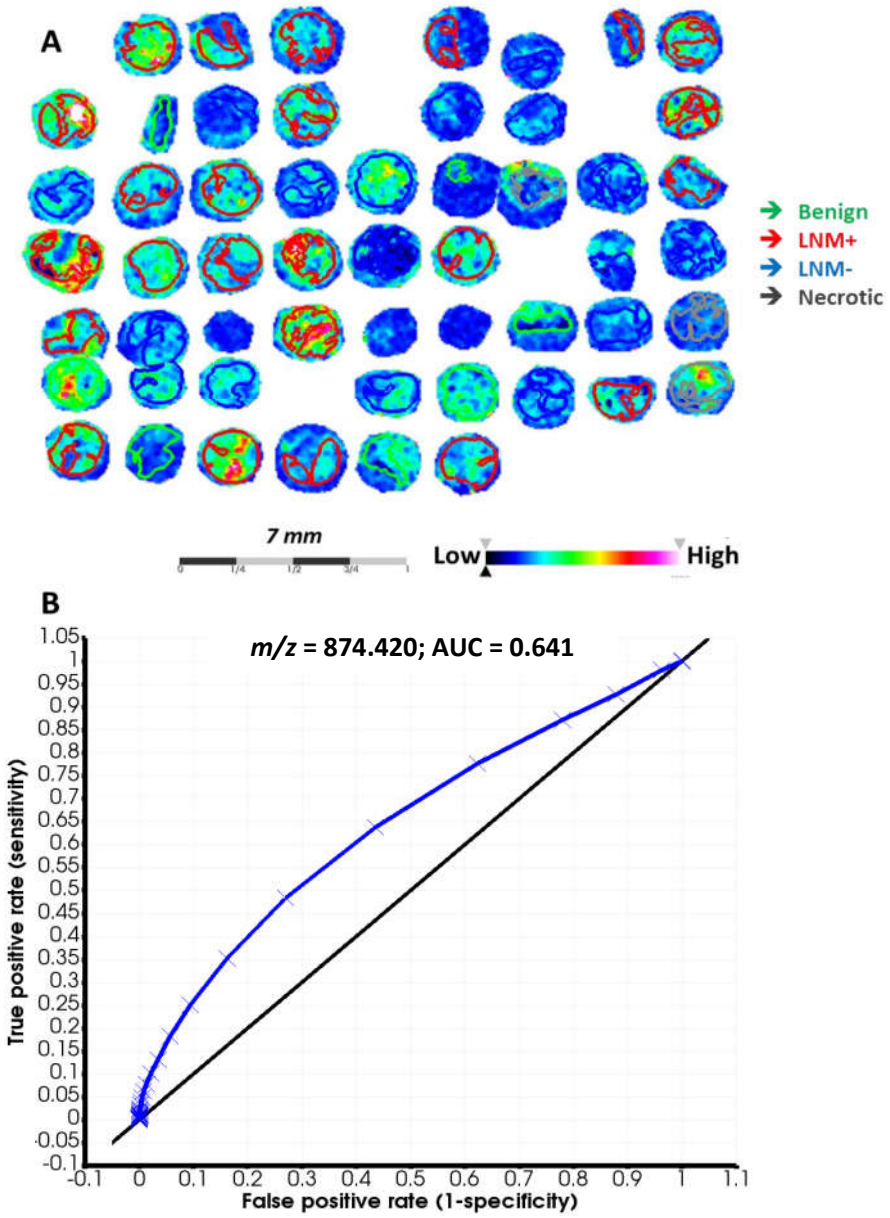


Figure 4.10: Representative MALDI-MSI images for NDRG1 peptide SHTSEGTR *m/z* 874.420. (A) Overview of one TMA slide, MALDI-MSI image of (intensity range from blue (lowest) to red (highest)). The samples belonging to different groups are indicated by different coloured circles: benign (green), LNM positive (red), LNM negative (blue) and necrotic/not included in the study (grey). Tumour regions within samples were annotated by a pathologist. B: ROC curve of identified NDRG1 protein peptide bin with associated boxplot of peptide intensities.

The bin of 841.333 and 1043.583 displayed very low ion intensities in MALDI-MSI results of vulvar cancer TMAs, while bin of *m/z* 1316.833 (AUC=0.543) matched peptide from ATP synthase subunit alpha and bin of *m/z* 1547.833(AUC= 0.444) matched collagen alpha-3(VI) chain in very small number of patient samples. Therefore, those classification bins were removed from the candidate list as they were less representative of cohort patient samples.

In summary, 14 *m/z* statistically significant bin classifiers from selective CCA-LDA classification results were matched with peptides identified in *nano*-LC-MS/MS and *in-situ* MS/MS experiments. Among identified proteins, *m/z* 874.420 matched protein NDRG1 and *m/z* 1335.583 matched Annexin A8 with higher AUC scores for LNM positive than LNM negative patient samples. For other identified upregulating proteins, cytokeratin 1 and collagen type I, expression was supported by tandem MS/MS identification. Another keratin protein, cytokeratin matching bin *m/z* 1391.833, was down-regulated in LNM positive VSCC patient samples together with vimentin peptides. To validate LNM VSCC proteomic patterns identified in MALDI-MSI, immunohistochemistry (IHC) stain of six candidate proteins was performed to validate expression change and spatial distribution on the tissue.

4.3.5 Immunohistochemistry (IHC) validation of candidate proteins identified in MALDI-MSI

The different abundance of K2C5/K2C6 between LNM positive and negative samples as identified in MALDI-MSI classification experiments were further investigated by IHC on consecutive TMA sections. Stronger positive stain of K2C5/K2C6 was mainly observed in tumour regions of LNM negative samples rather than LNM positive VSCC tumour samples. IHC analysis of K2C5/K2C6 revealed higher percentage in the level of total positive staining, especially high positive and positive staining, in LNM negative samples (34.57 ± 6.985%) compared with LNM positive samples (14.41 ± 1.684%) with statistical significance ($p=0.0230$). The results of IHC analysis and K2C5/K2C6 staining from two representative LNM positive and negative TMA samples are provided in ***Figure 4.11***.

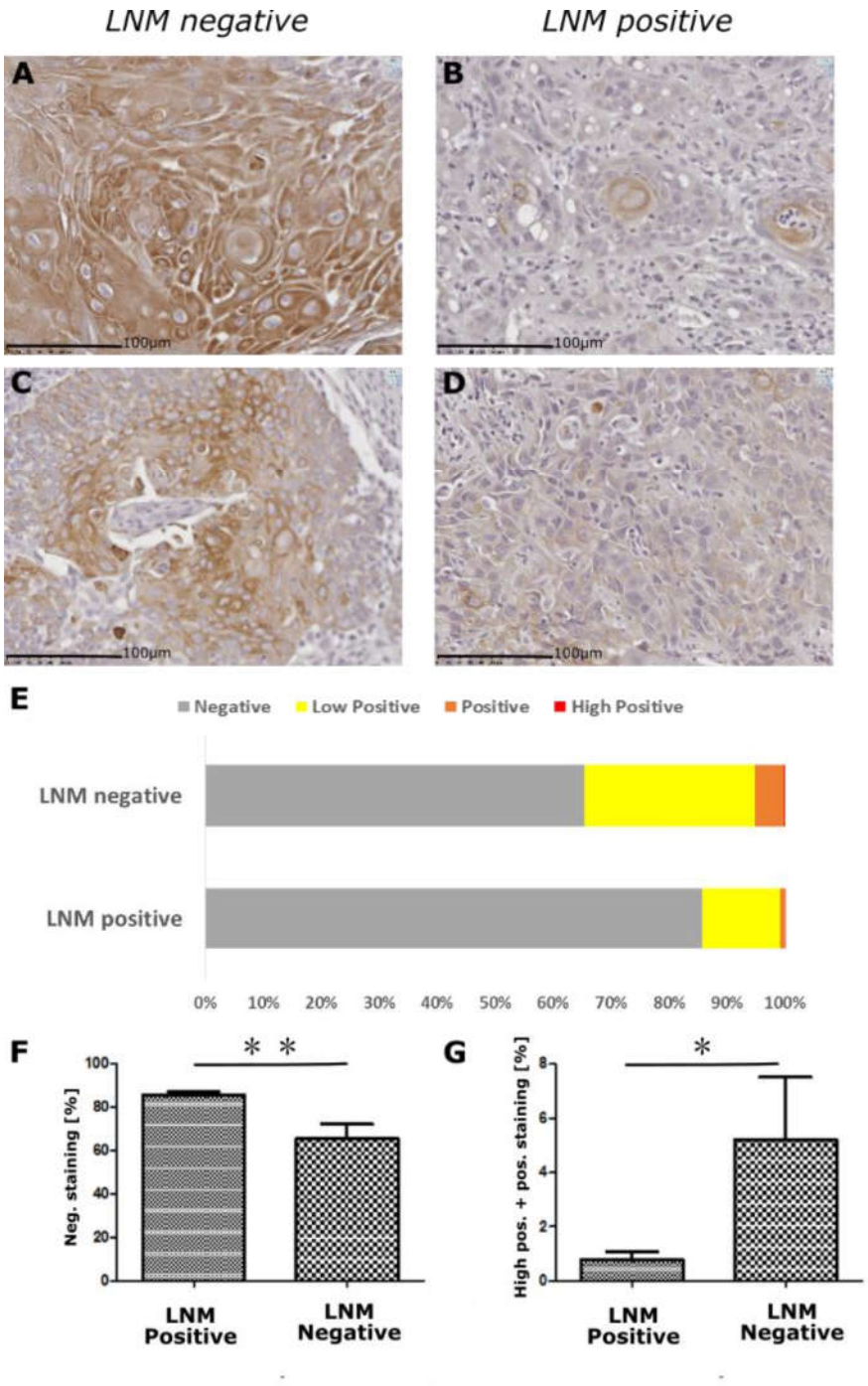


Figure 4.11: IHC analysis of K2C5/K2C6 on TMA sections containing samples of LNM positive and negative VSCC samples, the nucleus is stained purple in the figure above, at 40× magnification. (A) K2C5/K2C6 staining of LNM negative VSCC sample; (B) LNM positive sample; (C) K2C5/K2C6 staining of different LNM negative VSCC sample; (D) LNM negative from different sample; (E) Quantitative analysis of K2C5/K2C6 staining across the TMA section was performed using IHC profiler-Image J. For each tissue section, three representative photo-

micrographic images at 40× magnification were used and each image was assigned a staining score of negative, low positive, positive and high positive; (F) The per cent of negative staining was significantly different in the t-test between LNM positive and negative VSCC samples (p=0.0115); (G) The percent of positive and high positive staining was significantly different between LNM positive and negative VSCC samples ($p = 0.0447$).

The different abundance of K2C1/K1C10 between LNM positive and negative samples as revealed in MALDI-MSI was further investigated by IHC on consecutive TMA sections. A very weak high positive and positive staining of K2C1/K1C10 was observed in the cytoplasm of both LNM positive and negative VSCC tumour samples, regardless a large percentage of low positive staining was detected. IHC analysis of K2C1/K1C10 revealed differences in the percentage of total positive (1-negative) staining between LNM positive samples (10.98 ± 2.068%) and LNM negative samples (4.730 ± 1.153%) in statistical tests ($p=0.0166$). Results for IHC analysis and K2C1/K1C10 staining from two representative LNM positive and negative TMA samples are provided in ***Figure 4.12***. However, the suspected K2C1/K1C10 peptide matched m/z 1335.721 in MALDI-MSI could not be validated by K2C1/K1C10 antibody stains in IHC on consecutive VSCC TMA sections, which suggests the ion signal of m/z 1335.721 may come from other proteins.

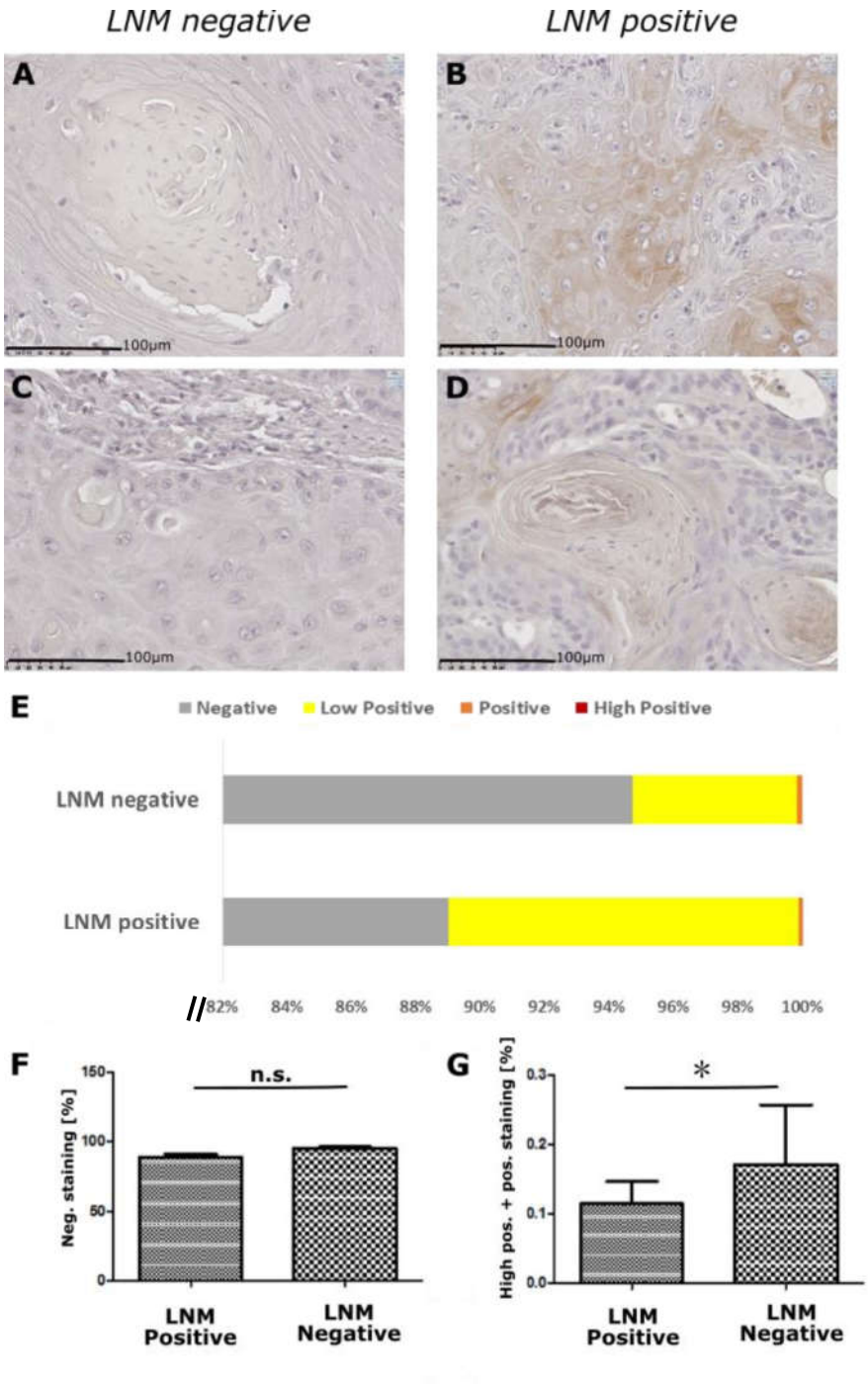


Figure 4.12: IHC analysis of K2C1/K1C10 on TMA sections containing samples of LNM positive and negative VSCC samples, all at 40× magnification. (A) K2C1/K1C10 staining of LNM negative VSCC sample; (B) LNM positive sample; (C) K2C1/K1C10 staining of from different LNM negative VSCC sample; (D) LNM negative from different sample; (E) Quantitative analysis of NDRG1 staining across the TMA section was performed using IHC profiler-Image J. For each tissue section, three representative photo-micrographic images at 40×

magnification were used and each image was assigned a staining score of negative, low positive, positive and high positive; (F) The per cent of negative staining was significantly different between LNM positive and negative VSCC samples (p= 0.0172); (G) The per cent of positive and high positive staining was absolutely not significant between LNM positive and negative VSCC samples ($p=0.5503$).

The different abundance of vimentin between LNM positive and negative samples as displayed in MALDI-MSI was further investigated by IHC on consecutive TMA sections. The stronger positive stain of vimentin was widely observed in cytoplasm of stromal regions than VSCCs in both LNM negative and positive VSCC tumour samples. IHC analysis of vimentin revealed a higher overall intensity in the level of total positive staining in LNM negative samples (29.83 ± 6.725%) compared with LNM positive samples (11.21 ± 1.515%) with statistical significance ($p=0.0270$). Results of IHC analysis and vimentin staining from two representative LNM positive and negative TMA samples are provided in ***Figure 4.13***.

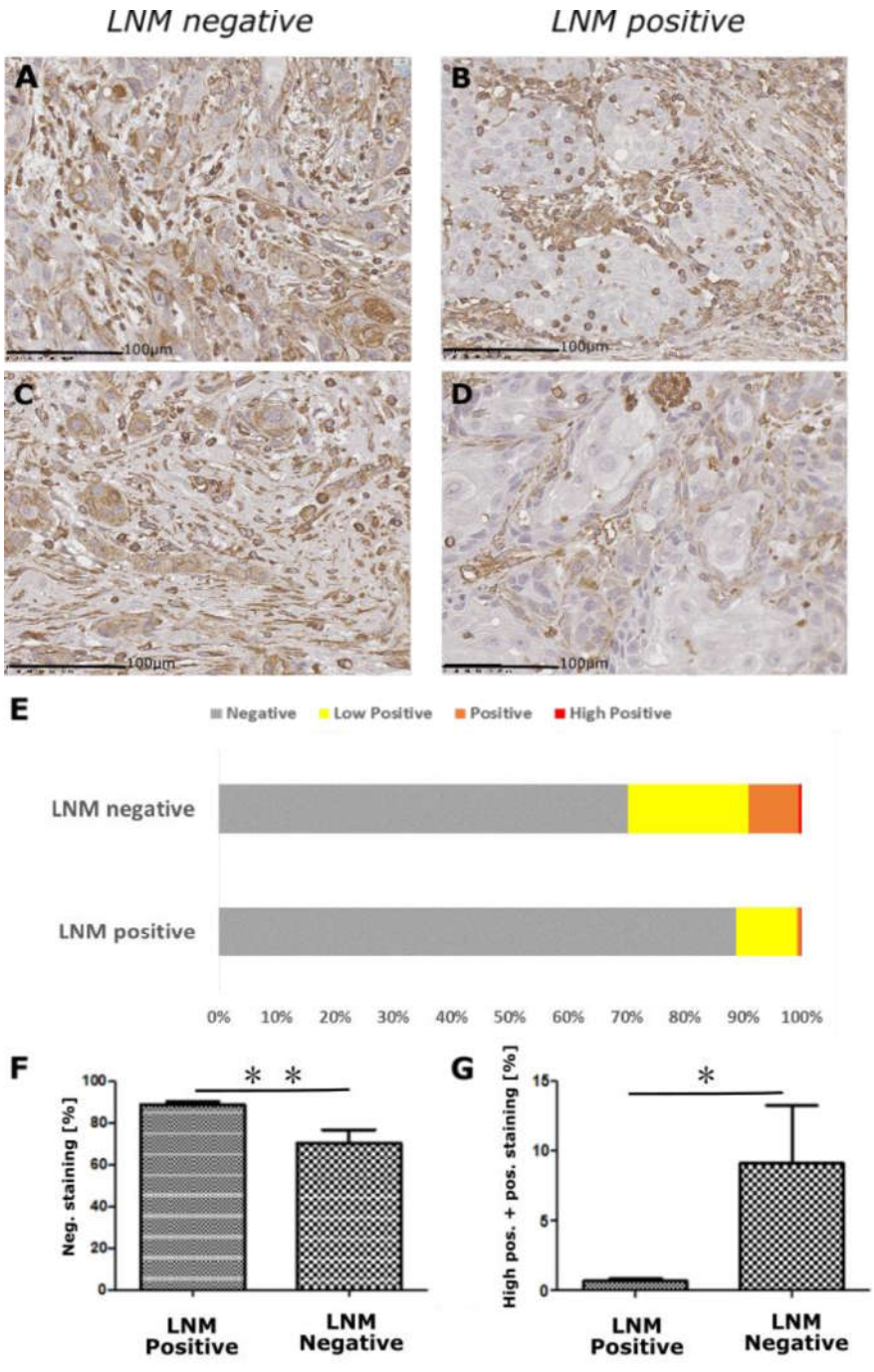


Figure 4.13: IHC analysis of vimentin (VIME) on TMA sections containing samples of LNM positive and negative VSCC samples, at 40× magnification. (A) VIME staining of LNM negative VSCC sample; (B) LNM positive sample; (C) VIME staining of different LNM negative VSCC sample; (D) LNM negative from different sample; (E) Quantitative analysis of VIME staining across the TMA section was performed using IHC profiler-Image J. For each tissue section, three representative photo-micrographic images at 40× magnification were used

and each image was assigned a staining score of negative, low positive, positive, and high positive; (F) The per cent of negative staining was significantly different between LNM positive and negative VSCC samples (p=0.0135); (G) The percent of positive and high positive staining was significantly different between LNM positive and negative VSCC samples ($p = 0.0362$).

The different abundance of Collagen type 1 between LNM positive and negative samples as revealed by MALDI-MSI was further investigated by IHC on consecutive TMA sections. Positive IHC stain of Collagen type 1 was predominantly found in the extracellular matrix of both LNM positive and negative VSCC tumour samples. Weak positive stain of collagen type 1 was also observed in cytoplasm of LNM positive tumours. IHC analysis of collagen type 1 revealed difference in the level of total positive staining between LNM positive samples (19.61 ± 3.312%) and LNM negative samples (7.002 ± 1.138%) with statistical significance ($p=0.0014$). Results for IHC analysis and Collagen type 1 staining from two representative LNM positive and negative TMA samples are provided in ***Figure 4.14***, which showed slightly stronger positive staining levels in LNM positive samples.

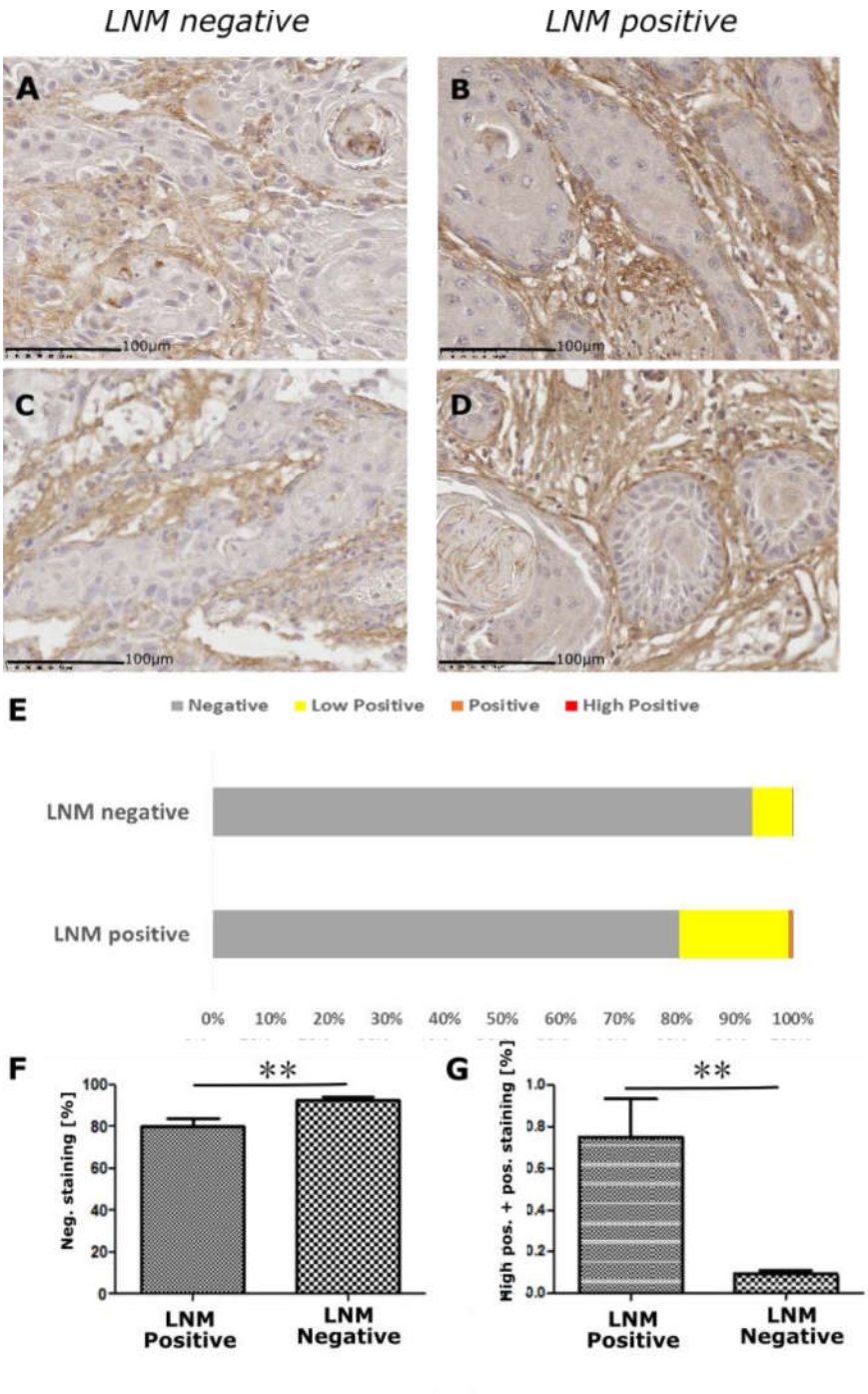


Figure 4.14: IHC analysis of Collagen type 1 (COL-1) on TMA sections containing samples of LNM positive and negative VSCC samples, at 40× magnification. (A) COL-1 staining of LNM negative VSCC sample; (B) LNM positive sample; (C) COL-1 staining of from different LNM negative VSCC sample; (D) LNM negative from different sample; (E) Quantitative analysis of COL-1 staining across the TMA section was performed using IHC profiler-Image J. For each tissue section, three representative photo-micrographic images at 40× magnification

were used and each image was assigned a staining score of negative, low positive, positive, and high positive; (F) The per cent of negative staining was significantly different between LNM positive and negative VSCC samples (p=0.0014); (G) The percent of positive and high positive staining was significantly different between LNM positive and negative VSCC samples ($p=0.0018$).

The different abundance of NDRG1 between LNM positive and negative samples as revealed in MALDI-MSI was further investigated by IHC on consecutive TMA sections. Stronger positive stain of NDRG1 was mainly observed in cytoplasm of VSCC tumour from LNM positive samples rather than LNM negative VSCC tumour samples. IHC analysis of NDRG1 revealed a significant difference ($p=0.0007$) in the level of total positive staining between LNM positive samples (31.55 ± 2.852%) and LNM negative samples (18.91 ± 2.096%). Results of IHC analysis and NDRG1 staining from two representative LNM positive and negative TMA samples are provided in ***Figure 4.15***, which revealed stronger high positive and positive stain levels in LNM positive samples than staining levels from LNM negative sample.

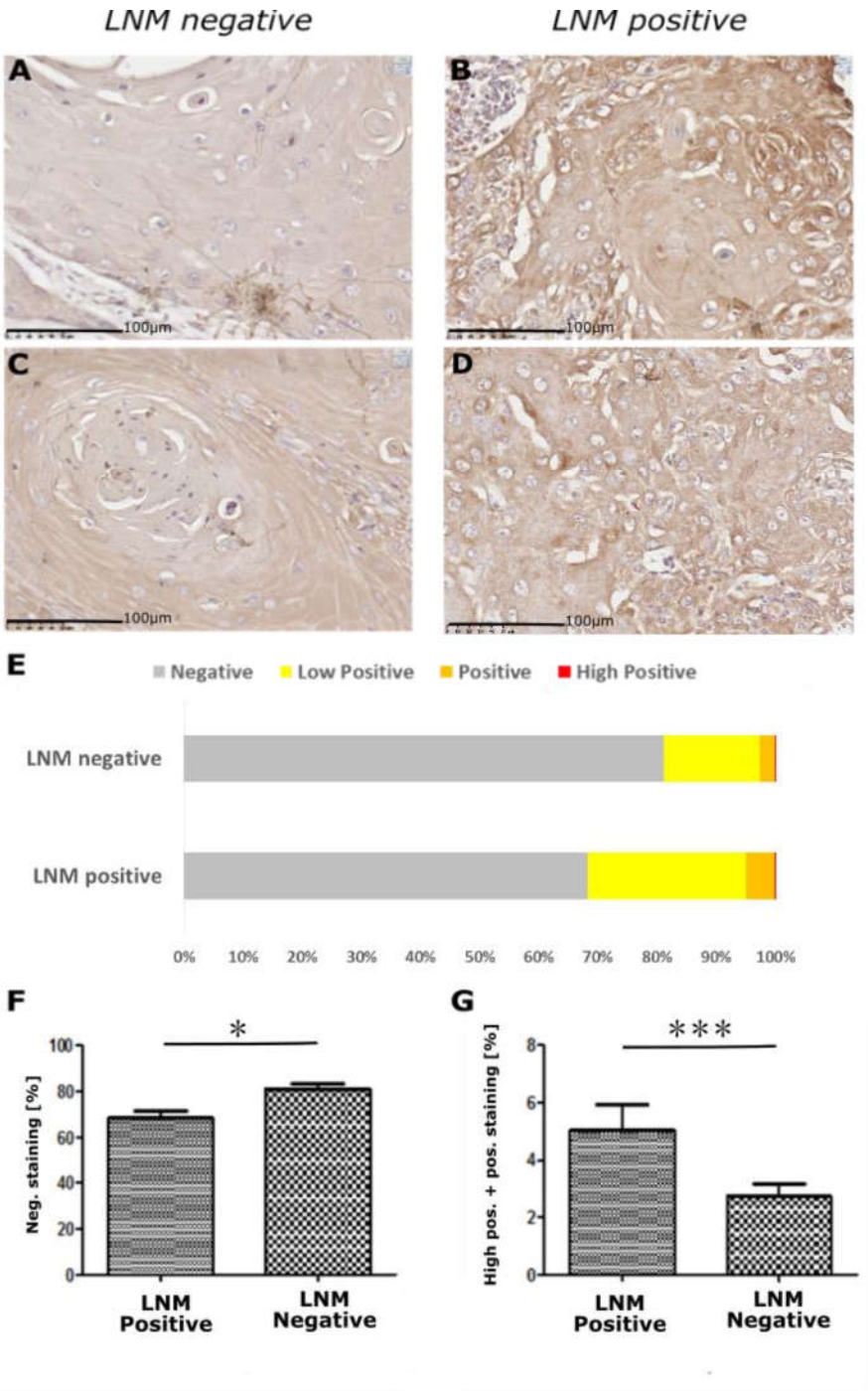


Figure 4.15: IHC analysis of NDRG1 on TMA sections containing samples of LNM positive and negative VSCC samples, at 40× magnification. (A) NDRG1 staining of LNM negative VSCC sample; (B) LNM positive sample; (C) NDRG1 staining of from different LNM negative VSCC sample; (D) LNM negative from different sample; (E) Quantitative analysis of NDRG1 staining across the TMA section was performed using IHC profiler-Image J. For each tissue section, three representative photo-micrographic images at 40× magnification were used and each

image was assigned a staining score of negative, low positive, positive, and high positive; (F) The per cent of negative staining was significantly different between LNM positive and negative VSCC samples (p=0.0122); (G) The percent of positive and high positive staining was significantly different between LNM positive and negative VSCC samples ($p=0.0004$).

The different abundance of ANXA8 between LNM positive and negative samples as discovered in MALDI-MSI was further investigated by IHC on consecutive TMA sections. Stronger positive stain of ANXA8 was observed in cytoplasm of VSCC tumours from LNM positive samples than in LNM negative VSCC tumour samples. IHC analysis of ANXA8 revealed a statistically significant ($p<0.0001$) high level of total positive staining in LNM positive samples (32.50 ± 2.413%) compared with LNM negative samples (20.55 ± 1.341%). Results of IHC analysis and ANXA8 staining from two representative LNM positive and negative TMA samples are provided in ***Figure 4.16,*** which showed higher positive stain in LNM positive patients than LNM negative patients.

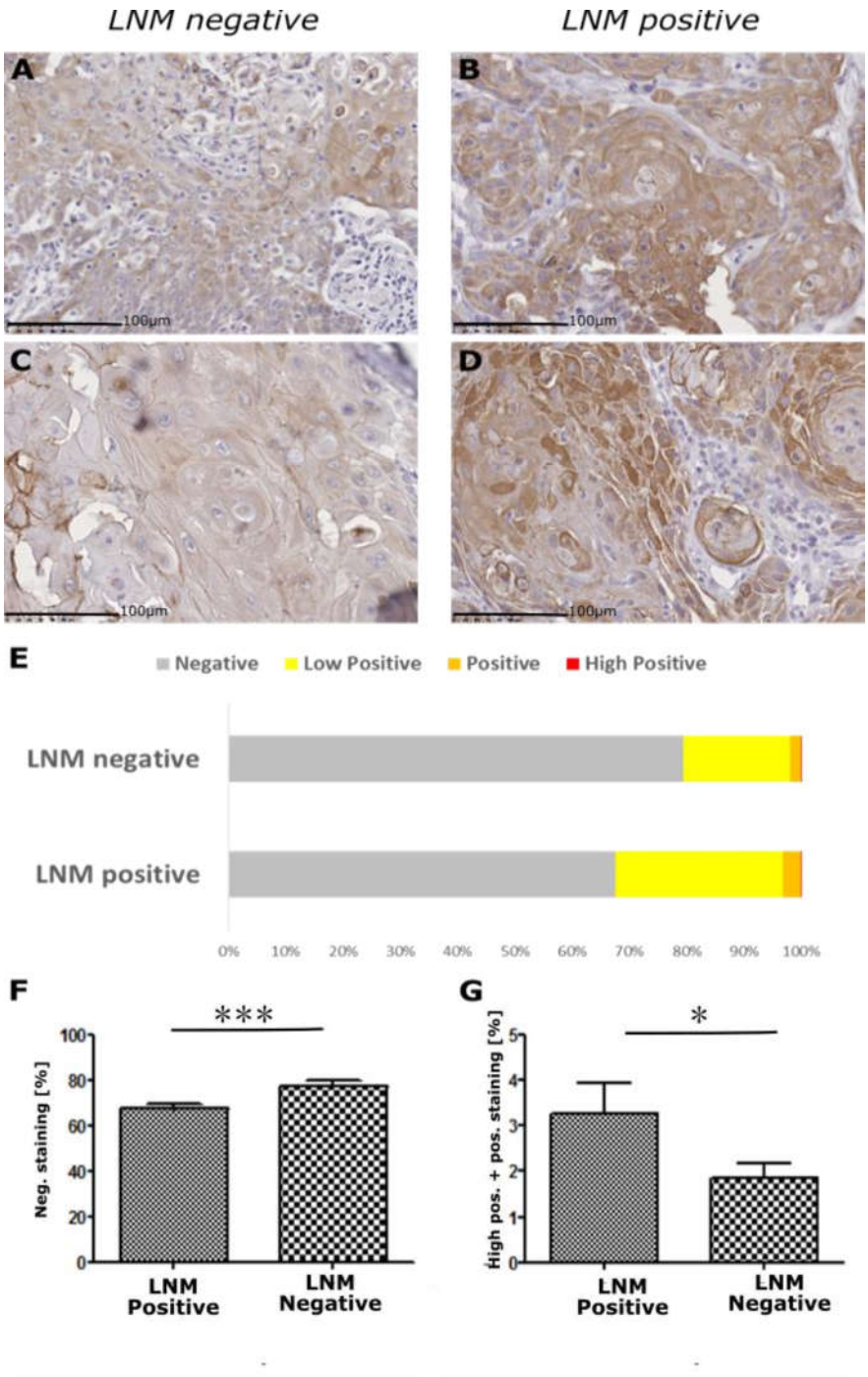


Figure 4.16: IHC analysis of ANXA8 on TMA sections containing samples of LNM positive and negative VSCC samples, at 40× magnification. (A) ANXA8 staining of LNM negative VSCC sample; (B) LNM positive sample; (C) ANXA8 staining of from different LNM negative VSCC sample; (D) LNM negative from different sample; (E) Quantitative analysis of ANXA8 staining across the TMA section was performed using IHC profiler-Image J. For each tissue section, three representative photo-micrographic images at 40× magnification were used and each

image was assigned a staining score of negative, low positive, positive and high positive; (F) The per cent of negative staining was significantly different between LNM positive and negative VSCC samples (*p=0.00034*); (G) The percent of positive and high positive staining was significantly different between LNM positive and negative VSCC samples (*p=0.0324*).

In summary, increases of NDRG1, ANXA8, Collagen type I and decrease of cytokeratin 6 and vimentin on LNM positive VSCC patient samples were highlighted by positive stain scores obtained by ImageJ analysis. Comparison of the cytokeratin 6, vimentin, collagen type 1, and NDRG1 and ANXA8 peptides, detected by MALDI-MSI as *m*/*z* 1391.741, 1028.555, 868.444, 874.420 and 1335.721 respectively, with their IHC stains on consecutive VSCC TMA sections is shown in ***Figure 4.17***. Their protein abundance differences between LNM positive and negative sample cores reflected in IHC analysis demonstrated similar changes with MALDI-MSI results. Further label-free quantitation (LFQ) using a high mass resolution qTOF instrument was performed on 12 VSCC patient samples as the second validation methods for LNM protein predictors discovered in MALDI-MSI experiments.

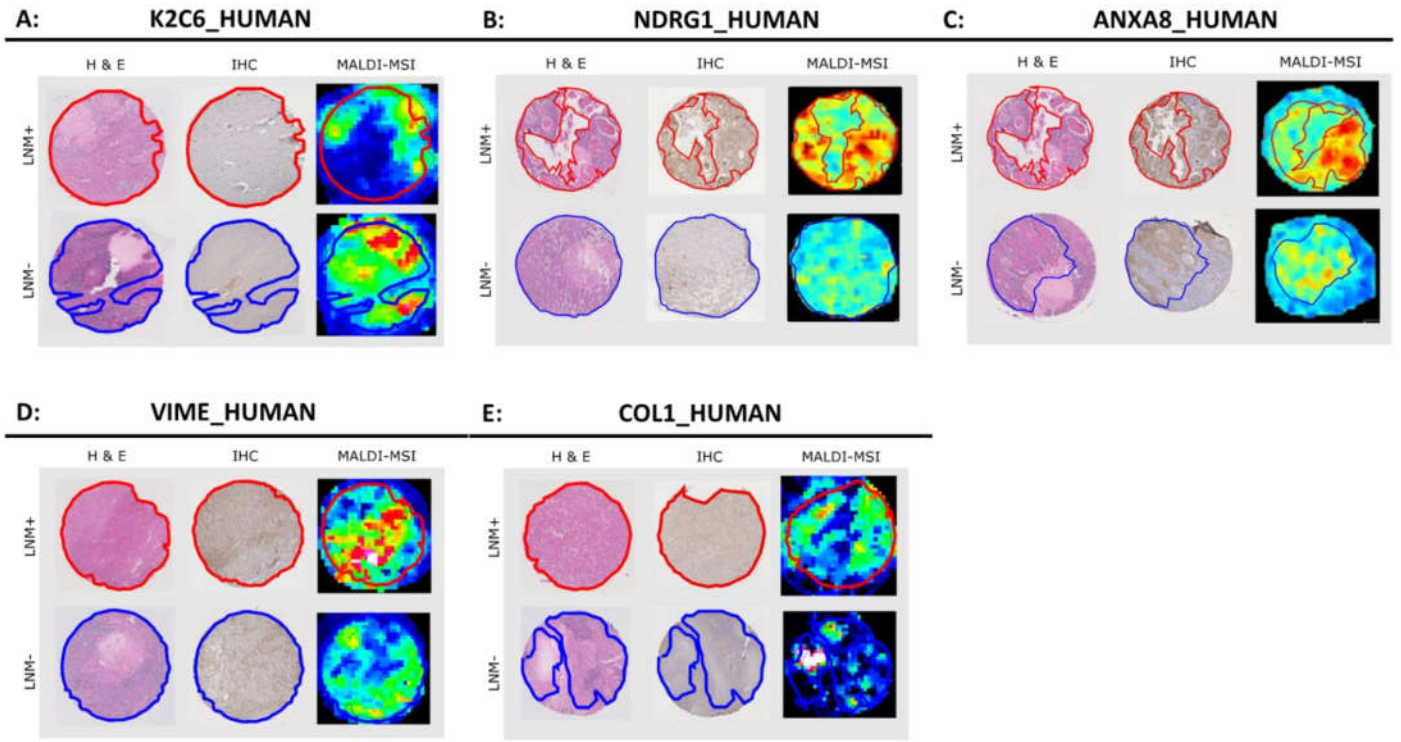


Figure 4.17: Comparing the results of H&E staining, IHC staining and MALDI-MSI.(A) Comparison of expression pattern observed for K2C6 peptides LALDVEIATYRK, in *m/z* 1391.741, in LNM negative (blue) and positive (red) TMA cores using MALDI-MSI ion intensity map (right), H&E staining (left), and K2C6 IHC staining (middle) on consecutively cut VSCC TMA sections; (B) Comparison of expression pattern observed for NDRG1 peptide SHTSEGTR, in *m/z* 874.420, in LNM negative (blue) and positive (red) TMA cores using MALDI-MSI ion intensity map (right), H&E staining (left), and NDRG1 IHC staining (middle) on consecutively cut VSCC TMA sections; (C) Comparison of expression pattern observed for ANXA8 peptide LIVALMYPPYR, in *m/z* 1335.721, in LNM negative (blue) and positive (red) TMA cores using MALDI-MSI ion intensity map (right), H&E staining (left), and IHC staining (middle) on consecutively cut VSCC TMA sections; (D) Comparison of expression pattern observed for VIME peptide SVSSSYRR, in *m/z* 1025.555, in LNM negative (blue) and positive (red) TMA cores using MALDI-MSI ion intensity map (right), H&E staining (left), and IHC staining (middle) on consecutively cut VSCC TMA sections; (E) Comparison of expression pattern observed for COL1 peptide GPSGPQGIR, in *m/z* 868.583, in LNM negative (blue) and positive (red) TMA cores using MALDI-MSI ion intensity map (right), H&E staining (left), and IHC staining (middle) on consecutively cut VSCC TMA sections.

4.3.6 Label-Free Quantification of Data-dependent data acquisition

Label-free quantification (LFQ) of vulvar cancer LNM associated proteins was based on *nano*-LC-ESI-MS/MS data-dependent acquisition (DDA) from six LNM positive and six LNM negative patient tumour samples from the TMA cohort. In brief, 1.5 µm diameter and 6 µm thick FFPE VSCC sections from those 12 patients were collected using Laser Capture Microdissections (LCM) and subject to trypsin digest based on filter-aided sample preparation (FASP) protocol described in Chapter 2.2.1. 10 (Wisniewski et al., 2009). Eluted peptide solution was loaded equally (1µg/µL) onto LC-MS/MS for DDA analysis. Total ion count chromatograms of 12 samples were conducted to show reproducibility between samples (***Figure 4.18***).

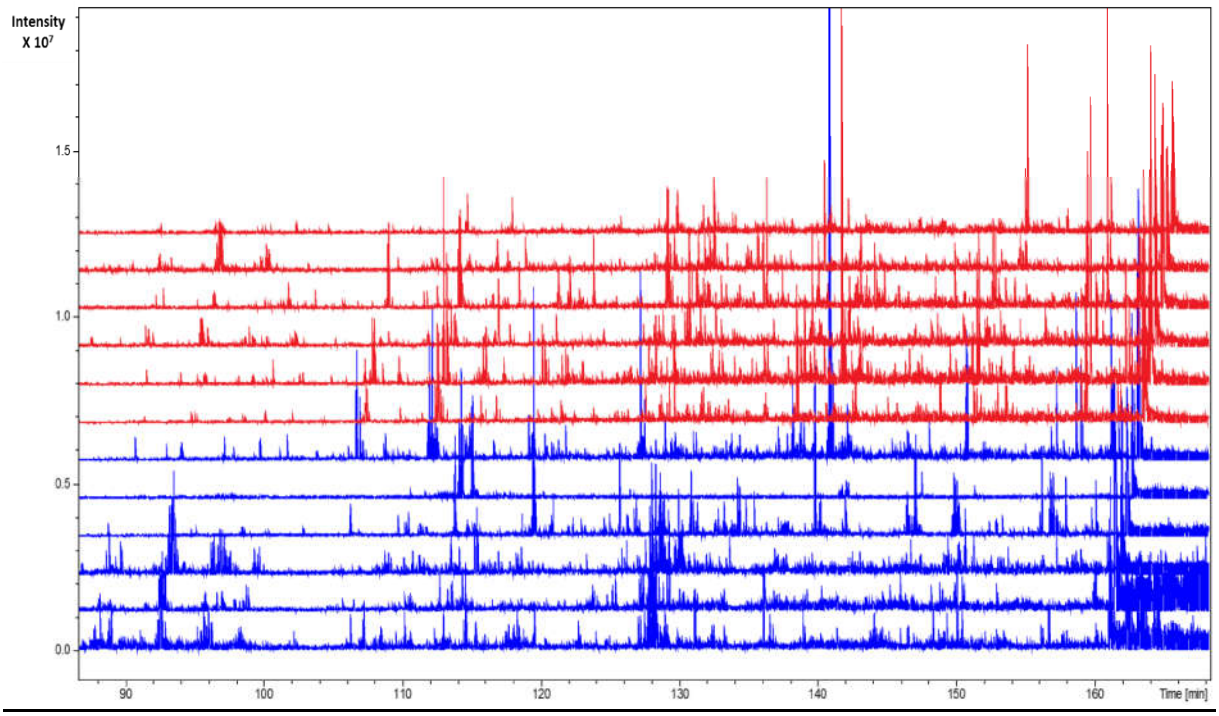


Figure 4.18: Total ion count chromatograms of *nano*-LC-MS/MS DDA analysis of six LNM positive (red) and six LNM negative (blue) digested peptide samples.

Details of the parameter offsets for the protein identifications are described in Chapter 2, section 2.2.3.5. Protein intensities used for LFQ were normalized based on extract-ion-current (XIC) among different runs using the MaxLFQ tool in MaxQuant software (Cox et al., 2014b). Vulvar

cancer protein identification and quantitation results are summarized in protein files (***Supplementary S 4.7***). LFQ data of six LNM positive and six negative samples were processed using Perseus software platform which is able to analyze highly multivariate quantitative protein abundance data generated using mass spectrometry. After filtering of proteins by identification and decoy matching, a total of 1002 proteins were identified from those 12 patient samples and used to compare LFQ intensities between patients with and without lymph node metastasis. Protein LFQ intensities were transferred into $\log_2$ form and proteins detected in two or fewer patients were removed from subsequent analysis. Thus, this study analyzed 392 proteins present in more than three patients from either the LNM positive or negative patient group. To highlight proteins that expressed differently between six LNM positive and six negative patients, the student t-test was performed on protein LFQ intensities with a *p*-value threshold of 0.01. As a result, 15 proteins had a stringent *p*-value less than 0.01 (***Table 4.5***) and results of the volcano plot are shown in ***Figure 4.19***.

Table 4.5: Summary of significantly different ($p<0.01$) proteins in t-test results of LFQ analysis between six LNM positive and six LNM negative patients

LFQ intensity 1 LNM+	LFQ intensity 2 LNM+	LFQ intensity 3 LNM+	LFQ intensity 4 LNM+	LFQ intensity 5 LNM+	LFQ intensity 6 LNM+	LFQ intensity 1 LNM-	LFQ intensity 2 LNM-	LFQ intensity 3 LNM-	LFQ intensity 4 LNM-	LFQ intensity 5 LNM-	LFQ intensity 6 LNM-	Unique peptides	Unique sequence coverage [%]	N: Score	-Log Student's T-test p-value M_T	Majority protein IDs
21.11	20.48	21.11	21.07	20.27	20.30	19.77	20.21	19.29	19.63	20.12	19.61	10.00	36.00	323.31	2.83	sp|P00558|PGK1_HUMAN
15.02	16.19	15.09	15.77	15.68	15.65	18.64	17.61	16.78	17.89	18.47	16.12	3.00	5.20	7.18	3.01	sp|P05023|AT1A1_HUMAN sp|P13637|AT1A3_HUMAN sp|P50993|AT1A2_HUMAN
19.47	19.37	18.52	18.49	19.63	19.41	16.70	19.20	15.96	16.24	16.32	15.41	3.00	25.50	172.58	2.83	sp|P46782|RS5_HUMAN
19.57	19.43	18.81	18.43	18.90	19.11	16.02	19.08	16.00	16.32	15.31	16.42	6.00	19.10	89.94	2.93	sp|Q08380|LG3BP_HUMAN
18.10	19.51	18.29	19.45	19.09	19.09	16.32	18.29	15.19	16.40	16.60	15.98	3.00	4.40	138.82	3.33	sp|Q13751|LAMB3_HUMAN
19.88	18.13	21.57	20.16	19.58	19.63	15.89	18.33	15.60	17.64	17.08	17.71	4.00	17.80	46.60	2.87	sp|Q92597|NDRG1_HUMAN
19.65	20.11	20.51	19.75	20.83	20.56	18.42	19.55	19.54	16.64	17.57	19.32	19.00	15.10	250.47	2.08	sp|O75369|FLNB_HUMAN
21.69	21.75	20.95	21.47	21.76	21.99	20.58	21.50	19.00	18.85	20.67	20.35	8.00	19.10	323.31	2.05	sp|P08238|HS90B_HUMAN
16.61	16.65	16.69	17.43	16.89	17.15	16.15	16.56	15.85	16.11	16.44	16.43	2.00	14.40	8.72	2.45	sp|P62826|RAN_HUMAN
20.54	19.74	20.69	19.09	19.78	19.99	17.12	19.88	19.00	17.58	15.63	15.84	2.00	7.80	131.42	2.15	sp|P68366|TBA4A_HUMAN
17.68	18.07	17.62	18.11	16.97	17.93	17.22	15.21	16.54	16.16	16.31	16.80	3.00	13.00	5.64	2.70	sp|Q16181|SEPT7_HUMAN
20.04	19.88	20.14	20.20	20.83	20.70	19.07	19.86	18.56	16.64	19.24	18.92	11.00	35.50	287.28	2.11	sp|Q16658|FSCN1_HUMAN
17.97	19.50	18.00	19.21	19.41	19.38	16.73	18.19	16.41	17.48	15.72	18.63	8.00	4.60	41.40	2.00	sp|Q16787|LAMA3_HUMAN
21.38	19.48	19.84	20.13	18.64	18.53	16.34	19.59	15.92	15.77	16.14	16.13	9.00	38.50	323.31	2.68	sp|Q5VT79|AXA81_HUMAN sp|P13928|ANXA8_HUMAN

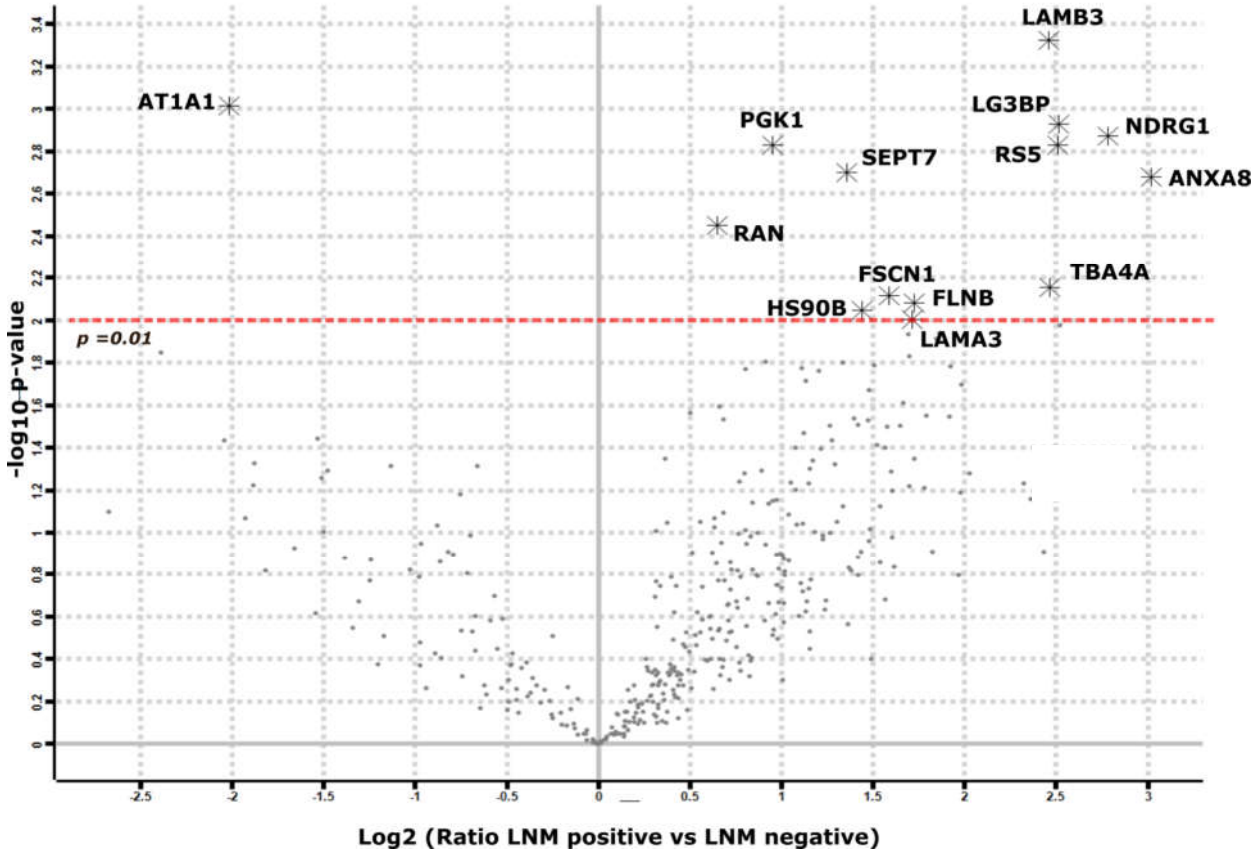


Figure 4.19: Volcano plot of student t-test results of protein expression difference in LFQ analysis between six LNM positive and six LNM negative vulvar cancer patients. X-axis: intensity differences (log_2 transformed) between metastasis patients group (right side) and non-metastasis group (left side). Y-axis: *p*-values ($-log_{10}$ transformed) from student t-test results. Proteins that were statistically significant in t-test ($p<0.01$) of LNM positive and LNM negative samples are highlighted by stars

Protein log_2 LFQ intensity profiles across all 12 patients with clustering analysis are shown in a heat map in ***Figure 4.20***. This figure shows wide protein abundance differences in samples. Clustering results using LFQ intensities of significantly changed proteins grouped 11 of 12 VSCC patients into the correct LNM classes, while one LNM negative patient was grouped into the LNM positive class probably due to high protein abundance, especially ANXA8 with over 8-fold of changes between two comparison classes. In order to improve the accuracy of LNM patient clustering using significant change in proteins, this study applied 5% FDR filter of permutation-based test on t-test results. As a result, intensities of five unique proteins, NDRG1, LAMB3, LG3BP, RS5 and PGK1, showed higher expression levels among patient samples from the LNM positive group with more than 1.5-fold increase on group normalised LFQ intensities (6.88, 5.50, 5.72, 5.69, and 1.94) compared with that of the LNM negative

group. The protein AT1A1/2/3 was more abundant among LNM negative patient samples with four times (log_2 intensity change of 4.05) change compared to LNM positive patient samples in the analysis.

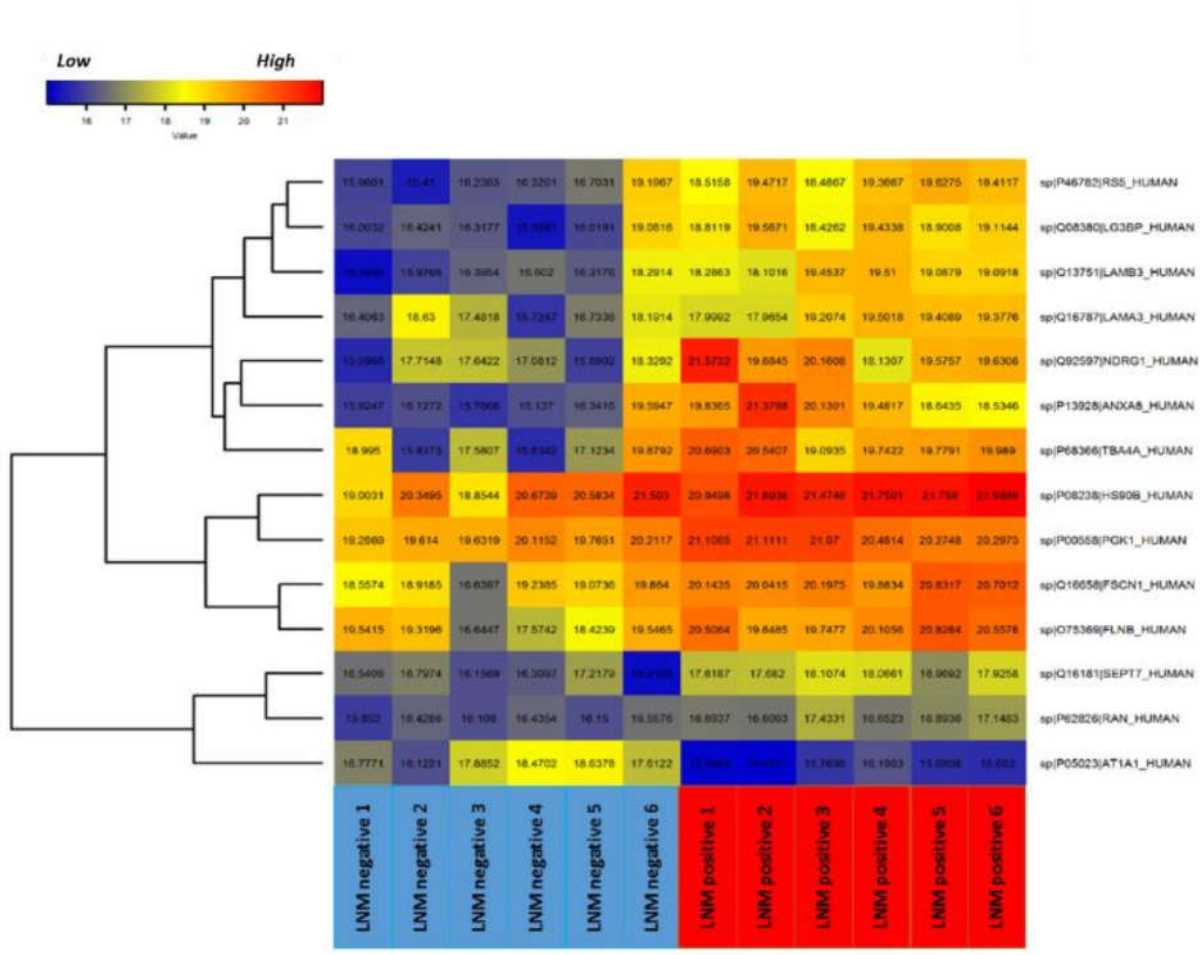


Figure 4.20: Clustering heat map of LFQ candidate protein intensities among six LNM positive patients and six LNM negative patients. The scale bar indicating changes in level is listed above.

In summary, NDRG1 protein and Annexin A8 were significantly upregulated in six LNM positive patient samples compared with six LNM negative patient samples. In particular, NDRG1 showed more than 6-fold increase in LFQ intensity in the LNM positive patient group. Furthermore, NDRG1 protein and Annexin A8 have also highlighted in MALDI-MSI and IHC results as having higher abundance among LNM positive patient samples compared with negative patient samples in the cohort. Other protein classifiers in the MALDI-MSI classification did not show significant change among samples from 12 VSCC patients in LFQ analysis.

4.4 DISCUSSION

In this study, two VSCC TMAs constituted by 16 LNM positive and 12 negative HPV-independent patients were used to perform MS analysis in an attempt to identify protein fingerprints associated with lymph node status, which could then be used for LNM prediction. The classification of MALDI-MSI results has revealed a list of peptide peaks which were able to discriminate LNM positive and negative patient samples of TMAs using their intensities. Those peaks are identified as NDRG1, ANXA8 and other cytoskeleton proteins tryptic peptides and could achieve fair AUC scores in determining lymph node status of vulvar cancer patients in ROC curve analysis. Further IHC staining and DDA based LFQ experiments also found higher protein levels of NDRG1 and ANXA8 in LNM positive patients compared with LNM negative ones. These results suggest those two proteins may play important roles in HPV-independent vulvar cancer metastasis and could be used as LNM indicators for vulvar cancer prognosis.

Using CCA-LDA method to classify VSCC TMA data has achieved 28 out of 28 accuracies in determining LNM status of patients. This results could complement current patient classification system based on pathohistological evidence and improve the diagnostic accuracy of lymph node biopsy. Without lymph node biopsy, the histologically defined LNM status of vulvar cancer patients remains a key challenge because tumour stages and grades of VSCCs are hardly achieved an interobserver agreement using tumour morphological criteria alone (Oonk et al., 2010a). Whilst the clinical course and prognostic management of HPV-independent VSCCs are highly variable in different areas (Woelber et al., 2012). Genomics and transcriptomics profiles have been considered in previous vulvar cancer patient classification, whereas poorly documented genetic profiles cannot support an accuracy genetic model in addition to current vulvar cancer LNM prediction. On the contrary, many proteomics studies have highlighted MALDI-MSI as an ideal technique in cancer patient classification and it has shown unique advantages in FFPE tissue sample analysis and closely related with cancer tissue morphology (Kriegsmann et al., 2016, Aichler and Walch, 2015, Winderbaum et al., 2016). Previously, MALDI-MSI was mainly employed to reveal spatial distribution of selective analytes, such as proteins or glycans, rather than selected as a promising diagnostic tool in clinical practice, because analytical protocols used for MALDI-MSI data processing and classification have not achieved any agreements yet. In this study, available MALDI-MSI data

classification methods have been compared in classifying LNM positive and negative patients from TMAs (***Supplementary files S4.8***). These results suggest the CCA-LDA could largely reduce MALDI-MSI data complexity and overcome the problem of overloading peak classifiers in the small cohort classification. Moreover, unlike the widely picked PCA analysis, CCA-LDA does not show any variable projection problem and largely preserved intact peptide peak masses for the following peptide identification (Winderbaum et al., 2016).

In this study, CCA-LDA has achieved 28 of 28 corrections in vulvar cancer patient classification. However, the performance of CCA-LDA might be overestimated due to several experimental variations. A recent study suggested TMA sampling strategies with fewer than three cores per patient may not accurately represent heterogeneous tumour protein expression (Rubin et al., 2002). Unfortunately, patient samples used in the TMAs are only able to produce two sampling cores per block because of the smaller tumour region within each tissue block. Additionally, the detection of peptide signals also yielded to many technical difficulties, including ionization efficiency, in-source decay, post-translated modification, and mass resolution of MALDI instruments, which caused a poor discriminative ability towards low abundant peptides. Aim to improve the reliability of patient classification using MALDI-MSI data, this study also tried to incorporate necessary histopathological information, such as tumour grade and lymphovascular invasion, into the patient classification model, which could help to adjust most experimental variations in future analysis.

The different abundance of N-myc down-stream regulated protein 1 (NDRG1) was revealed in MALDI-MSI results (AUC=0.641) of LNM positive and LNM negative VSCC samples. As a promising non-cytoskeletal LNM protein, the high abundance of NDRG1 in LNM positive VSCCs has been validated by both IHC staining and LFQ analysis on consecutive sections. The expression change of NDRG1 has been reported as a cancer metastatic indicator in ovarian, breast, uterine, cervix, and other gynecological cancers (Zhao et al., 2011, Chen et al., 2008, Nagai et al., 2011). Previous genetic study has shown the expression of NDRG1 is directly triggered by p53 and PTEN via an activated PI3K signalling pathway in response to DNA damage (Fang et al., 2014). However, the NDRG1 expression has also been found to be upregulated by p21 expression independent from p53 regulation pathways in prostate cancer (Kovacevic et al., 2011). Apart from being regulated by tumour-suppressor pathways, a recent study showed NDRG1 expression may be induced by hypoxia stimuli via HIF-1α-dependent pathway (Salnikow et al., 2000). Interestingly, tumour hypoxia-related proteins, such as PGK1

and AT1A, were also found in the LFQ results of LNM positive VSCC samples together with NDRG1. There is increasing evidence that the role of NDRG1 in cancer metastasis is highly dependent on tumour origins (Ellen et al., 2007, Kovacevic and Richardson, 2006). For example, NDRG1 is able to reduce cell-cell/cell-matrix adhesion and suppress metastasis in colorectal and prostate cancers (Kovacevic and Richardson, 2006). In oral squamous cell carcinoma, NDGR1 was reported in actin-remodelling regulation and EMT progression by blocking ROCK1 and pMLC2 pathways (Lee et al., 2014). However, NDRG1 was found overexpressed in esophageal squamous cell carcinomas patients with poor 5-year survival outcomes and worked as EMT promoter rather than inhibitor by mediating the Wnt pathway (Ai et al., 2016). While knocking down NDRG1 in gastric cancer cell line showed an enhancement of E-cadherin expression and downregulation of vimentin and Snail, inhibiting NDRG1 expression in animal xenograft model showed negative effects on TLE2 and β-catenin expression (Ureshino et al., 2012). Furthermore, the up-regulation of NDRG1 may also able to increase the tumour resistance to chemotherapeutic drugs in neuroblastoma cell line (Jia et al., 2015). Therefore, the protein NDRG1 may not only be used as a potential vulvar cancer LNM predicator but also a potential therapeutic target for future vulvar cancer treatments.

Annexin A8 (ANXA8) is another potential protein marker identified in this study, to predict the LNM of vulvar cancer. A significant high positive stain level of ANXA8 in LNM positive samples was detected using IHC analysis and an estimated over eight-fold increase in protein intensity difference between LNM positive and LNM negative vulvar cancer samples in LFQ analysis. In the RNA-Seq analysis of maxillary carcinoma, the up-regulation of ANXA8 has been reported in TP53 mutated tumour cells when compared to wild-type p53 samples (Kudo et al., 2017). Further studies found the ANXA8 expression is up-regulated by p63 which is the downregulation of myc and cytokeratin 5 gene expression in epithelial ovarian cancer (Trink et al., 2007). Moreover, previous studies have provided evidence that the level of ANXA8 together with cytokeratin 4 and cytokeratin 15 levels can be used to determine cervical cancer patients with and without metastasis (Nakabayashi and Hiyama, 2011). A similar result was also reported in qRT-PCR results of oral squamous cell carcinoma samples where cytokeratin 6 and ANXA8 showed significant expression differences between LNM positive and LNM negative patients (Oka et al., 2016). In a recent gene microarray study, ANXA8 and NDRG1 were able to discriminate cervical squamous cell carcinoma patients from other types of cervical cancer (Chao et al., 2006). Here, this study also examined the LNM predicting the power of ANXA8 and NDRG1 in vulvar cancer TMAs. Using their combined tryptic peptide

intensities resulted in an AUC of 0.679 in the ROC analysis, which was better than prediction results using NDRG1 (AUC=0.641) and ANXA8 (AUC=0.625) individually, indicating their potential prognostic values in the clinical management of HPV-independent vulvar cancer.

Most tryptic peptides selected in the CCA-LDA classification list are identified as cytoskeletal proteins, such as cytokeratins or β-actin, while cytoskeletal network remodelling has been increasingly accepted as a hallmark in cancer invasion and metastasis (Kedrin et al., 2007). For example, the actin-remodelling has been reported as a subsequent outcome of epithelial-mesenchymal transition (EMT) changes in VSCCs which has been recently detected as a key biological process driving the metastatic spread of HPV-independent vulvar cancer (Wicki et al., 2006, Zannoni et al., 2011). In the MALDI-MSI analysis, the identification of the peptide DLTDYLMK (present in *m/z* 999.458 bin) as β-actin and showed higher intensity levels in LNM positive samples when compared with LNM negative ones. Therefore, β-actin has a potential to be vulvar cancer LNM prediction marker. Previously, cytokeratin 5 has been highlighted as a protein marker for HPV-independent VSCCs in the published manuscript (Zhang et al., 2016b). Here, the decrease levels of intermedia filament proteins, cytokeratin 5/6, has been shown in late-stage vulvar cancer patients when compared to early-stage patients. These results suggest that a loss of cytokeratin 5/6 in vulvar epithelium might be correlated with the progression of HPV-independent vulvar cancer. Moreover, another cytokeratins, namely cytokeratin 17, has been identified with high peptide ion intensity levels in LNM positive patients. As a promoter of cell proliferation, cytokeratin 17 has been shown clinically significant role in cancer invasion, metastasis, and drug resistance (Goyal et al., 2016). Recent studies found the cytokeratin 17 acting as a basal cell differentiation regulator in the epithelium via binding to 14-3-3σ and activating the Akt/mTOR pathway in HPV-dependent tumorigenesis (Hobbs et al., 2016). The extracellular matrix protein, collagen type 1, is widely expressed in mammal skin and involved in many important biological progresses of embryonic skeletal development and regulating cell-substrate adhesion (Salasznyk et al., 2004). Here, collagen type 1 peptide matched with *m/z* bin of 868.583 was highlighted in the stroma of LNM positive sample cores. Functions of collagen type 1 during vulvar cancer progression have not been fully studied as yet, but previous experiments suggested that collagen accelerate the downregulation of E-cadherin by inducing its negative regulators (Hosper et al., 2013). However, the LNM prediction model using cytokeratin proteins achieved a very poor outcome in determining lymph node status of vulvar cancer patients, other proteins to improve this proposed vulvar cancer LNM prediction model.

4.5 CONCLUSION

In summary, MALDI-MSI has demonstrated advantages for protein identification with untainted spatial information from vulvar cancer tissue microarrays and VSCC patient classification. In this chapter, the description of the use of the *m/z* bin list from MALDI-MSI data to classify LNM positive and LNM negative patient samples and achieved a 28 out of 28 accuracy in leave-one-out validation. Importantly, those *m/z* bins matched cytokeratin 5/6, vimentin, collagen type I, NDRG1 and ANXA8 from the peptide library and abundance and spatial distribution was validated by IHC. Furthermore, LFQ analysis based on DDA confirmed higher protein abundance of NDRG1 and ANXA8 in the LNM positive samples and further highlighted other proteins as potential vulvar cancer metastasis biomarkers. These proteins are potential biomarkers to predict LNM in vulvar cancer and their functions in vulvar cancer development will be studied in future experiments.

This page intentionally left blank

CHAPTER 5: PROTEOMICS ANALYSIS OF INDIGENOUS VULVAR CANCER USING *NANO*-LC-ESI-MS/MS

5.1 Introduction

Vulvar cancer and its precursor lesion VIN have remarkably high prevalence rates within Indigenous Australian communities. For non-Indigenous Australians, vulvar cancer is a rare gynecological cancer with an incidence rate of 2 to 3 per 100,000 with most patients diagnosed over 60 years of age or post-menopause (Board, 2016). However, within the Indigenous community, vulvar cancer incidence rates have increased by 31% since 1998, in contrast with a 7% decrease in the non-Indigenous population. Furthermore, the risk of vulvar cancer and its precursor lesion VIN is significantly higher for Indigenous women from remote areas (3.9%) compared with Indigenous women in major cities (1.3%) (Condon et al., 2009a). Strikingly, most Indigenous vulvar cancer patients are located in Northern Territory (NT) or rural areas of other states and almost all of them are diagnosed before the menopause (McGrath et al., 2015). For example, the incidence ratio of vulvar cancer between Indigenous women from NT and remaining population was as high as 12.2 in the age group younger than 65, especially within the East Arnhem Land region where incidence rates are 31.1 per 100,000 among young Indigenous women (Condon et al., 2005). Additionally, the incidence rate of high-grade VIN is also high within the East Arnhem Land community at 34.7 per 100,000. Despite the high incidence rates for young Indigenous women within the community, no Indigenous women above the age of 50 and no non-Indigenous women have been diagnosed with VIN in East Arnhem Land (Condon et al., 2009b). The problem of Indigenous vulvar cancer has been documented and requires urgent attention due to the current lack of understanding of the disease development within the Indigenous Australian community.

Studies aiming to explore the molecular cause of the high incidence rates of vulvar cancer within the Indigenous Australian population have performed genome and proteome analyzes comparing differences against other non-Indigenous patients. Importantly, HPV testing of 2152 Australian women did not show any obvious difference in the infective rates of high-risk HPV genotypes, including HPV16 and 18, between Indigenous and non-Indigenous women (Garland et al., 2011). To date, no virus-induced genetic mutation has been identified that is unique to Indigenous vulvar cancer patients, therefore, other risk factors, such as immune disorders or imbalanced nutrition, might be the higher incidence rates among Indigenous women (McWhirter et al., 2014). The genome-wide screening of type 2 diabetes-associated genes in Indigenous Australians suggested that Indigenous population may harbour unique genetic changes related to insulin metabolism pathways on chromosome 2q (Busfield et al.,

2002). The mitochondrial genome has also shown different genetic patterns among the Indigenous population compared with the non-Indigenous population (van Holst Pellekaan et al., 2006). Furthermore, DNA sequencing of the Mannose-binding lectin (MBL) gene and it's promoters demonstrate that Indigenous Australians have unique genetic mutations which interfere with the structure and expression of MBL, and deficiency of MBL proteins has been shown increase the risk of infections, autoimmune diseases and cancer (Turner et al., 2000, Dommett et al., 2006, Peterslund et al., 2001). However, such genetic variations are still unable to explain the high prevalence rates of vulvar cancer among young Indigenous women.

Recently, a proteomic study on serum collected from a cohort of Indigenous Australians reported two serum inflammatory proteins, C-reactive protein (CRP) and soluble E-selectins, to be significantly increased in abundance (Rowley et al., 2003). Additionally, the geographical location of Indigenous communities with higher CRP levels, East Arnhem Land, for example, are also reported with high vulvar cancer incidence rates (Condon et al., 2003). Previously, chronic inflammation was believed to be the major cause of HPV-independent vulvar cancer as it correlates with the prognostic outcomes of patients (Taniguchi and Karin, 2014, Hefler-Frischmuth et al., 2010, Bodelon et al., 2014). Interestingly, whole genome sequencing of Indigenous vulvar cancer samples has revealed several unique single nucleotide polymorphism (SNP) patterns in eight cytokine genes that play important roles in the inflammatory immune response (Cox et al., 2014a). Comparing the proteomes of vulvar cancer tissues collected from Indigenous Australian to those collected from the remaining community may partially elucidate the molecular cause of the significant differences observed in incidence rates.

Proteome profiling by label-free quantification (LFQ) *nano*-LC-ESI-MS/MS has become a popular approach for the relative quantification and comparison of complex protein samples. Compared with methods that require protein or peptide labels, LFQ is a reliable, high-throughput, and low-cost alternative, capable of relatively quantifying a large number of proteins in biological samples, and has been used to elucidate variations and disease-related pathways in lung, colorectal, prostate, ovarian, breast and cervical cancers (Zhu et al., 2009, Alluri et al., 2016, Ramberg et al., 2017, Liu et al., 2013). In the previous chapter, LFQ analysis has been performed on non-Indigenous VSCC samples to validate the high levels of NDRG1and ANXA8 detected in lymph node metastasis (LNM) positive vulvar cancer patients. Here, *nano*-LC-ESI-MS/MS is used to compare vulvar cancer tissues from Indigenous and

non-Indigenous patients with the aim of providing an insight into the molecular pathways contributing to the increased prevalence of vulvar cancer amongst the Indigenous population.

To achieve this aim, the data-dependent acquisition (DDA) based LFQ was employed and revealed significant differences between Indigenous (n=4) and non-Indigenous (n=4) HPV-independent vulvar cancer tissues. The same analysis was also used to compare LNM positive (n=2) and negative (n=2) vulvar squamous cell carcinoma (VSCC) collected from Indigenous patients. Proteins found to have significantly different abundance levels were further validated by MALDI-MSI and IHC stain on consecutive tissue sections from the sample patients, and were subject to Reactome Pathway Analysis to identify signalling pathways related to Indigenous vulvar cancer. The proteins highlighted in this study may help to explain the unique progression pattern and epidemiological characteristics of vulvar cancer in Indigenous patients.

5.2 Materials and Methods

5.2.1 Sample cohort and pathological analysis: FFPE tissue blocks from Indigenous vulvar cancer patients (n=4) were retrieved from the archive of SA Pathology and another four non-Indigenous sample blocks from patient cohort were chosen as the comparison group. All patient samples were diagnosed as HPV-independent type vulvar cancer and regions of interests (ROIs) on tissues were carefully examined by experienced pathologists. All tissue annotations and histopathological details were shown in ***Supplementary S2.1***.

5.2.2 Peptide Identification and Label-free quantification by *nano*-LC-ESI-MS/MS: Tumor areas (ROIs) from eight vulvar cancer patients were collected using laser capture microdissection (LCM) for trypsin digestion and identification by *nano*-LC-ESI-MS/MS as described in Chapter 2.2.1.9. In brief, FFPE sample blocks were sectioned at 6µm thickness, water bath mounted onto PEN membrane slides, and deparaffinized in xylene, ethanol and rehydrated in NH_4HCO_3. Patient samples were dissected and transferred into 10 mM citric acid buffer (pH = 6) for heat-induced antigen retrieval, and then digested with 100 ng of trypsin gold overnight at 37 °C. Sample preparation followed Filter Aided Sample Preparation (FASP) protocol described in Chapter 2.2.1.10. All tryptic peptide samples prepared as above were diluted to 2µg/µL with 1µL solution was injected for the data-dependent acquisition.

The *nano*-LC-ESI-MS/MS was performed using an Ultimate 3000 RSLC system coupled to an Impact II QTOF mass spectrometer via an Advance CaptiveSpray source (Chapter 2.2.2.6). In brief, peptide separation was performed using a 75µm C18 column over 180 min. MS scans were acquired in the mass range of 300 to 2200 *m/z* in a data-dependent fashion using Bruker's Shotgun Instant method. Acquired spectra were subjected to Compass DataAnalysis for OTOF (v1.7). Raw spectra were then submitted to MaxQuant (v1.6.0.1) for identification. All raw files were searched against the UniProt human-reviewed-validated sequence database (generated from version 2017_06, human taxonomy, 20196 entries), trypsin/P with up to 2 missed cleavages fixed modification of carbamidomethylation of cysteine, variable modification of oxidation of methionine, N-terminal acetylation, and glutamine to pyroglutamate as variable modifications.; 40ppm MS/MS mass tolerance and 1% False Discovery Rate (FDR). The peptide identification cut-off was set with MaxQaunt PEP value less than 0.01. Proteins identified by unique peptides are grouped in one protein group and

quantified together in the MaxQuant, and proteins of the group are used to represent this identified protein group in the following study. The LFQ intensities of each identified proteins were automatically calculated when the MaxQuant LFQ tool was selected during data analysis. In brief, the LFQ intensities were normalized based on extract-ion-current (XIC) among different comparing runs using the internal MaxLFQ with minimum LFQ count ration of 2, processed LFQ data was further sent into Perseus (v1.6.0.2) for statistical tests and volcano plotting (Chapter 2.2.3.5).

5.2.3 Pathway analysis: Selected protein candidates were copied into the Pathway Browser of the Reactome Pathway Analysis Brower software (https://reactome.org/PathwayBrowser/) and analyzed using the Human proteome database. The signalling networks containing the proteins of interests are discussed in section 5.3.5.

5.2.4 MALDI-MSI data acquisition and analysis: Consecutive FFPE tissue sections from eight HPV-independent vulvar cancer patients were sectioned at 6μm thickness and mounted onto ITO slides. Sample slides were deparaffinized and subjected to heat-induced antigen retrieval into 10 mM citric acid buffer (pH = 6), and sprayed Trypsin Gold, Internal Calibrants and CHCA matrix onto sample slides using ImagePrep station (described in Chapter 2.2.3.3). Data acquisition was achieved in an ultrafleXtreme MALDI-TOF/TOF MS system and details of MALDI-MSI experiments was shown in Chapter 2.2.2.1. Acquired spectra were submitted to the SCiLS lab for the peptide ion intensity map and ROC curve creations (Chapter 2.2.3.3)

5.2.5 Immunohistochemistry: For the analysis of NDRG1, RAP1A, and POSTN by IHC, tissue collected from Indigenous and non-Indigenous VSCC patients were analyzed as previously described in Chapter 2.2.2.9. Briefly, tissue sections were de-waxed, rehydrated and subjected to antigen retrieval, then incubated overnight at 4 °C with antibodies followed by incubation with secondary antibodies, streptavidin-HRP, and Diaminobenzidine. Slides were further stain with hematoxylin and scanned using a NanoZommer scanner. An analysis was carried out using the IHC Profiler and antibodies' positive staining for Indigenous and non-Indigenous samples were compared across patients using GraphPad Prism 6 v008 (Chapter 2.2.3.8).

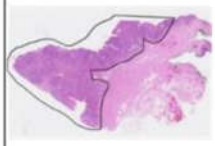
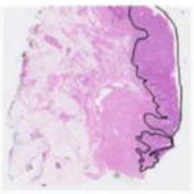
5.3 Results

5.3.1 Patient Samples

In total, tissue samples from four HPV-independent vulvar squamous cell carcinoma (VSCC) collected from Indigenous patients were used for proteomic characterization and validation. The patient cohort was constructed with a small number of samples due to the rareness of vulvar cancer in South Australian regions and the challenges of collecting samples from Indigenous Australians due to cultural issues (Anderson et al.). Another four VSCC tissues from non-Indigenous patients with parallel histopathological patterns were included in the study for the comparison. In total, eight formalin-fixed-paraffin-embedded (FFPE) sample blocks were sectioned at 6μm mounted on ITO, PEM membranes and Superfrost glass slides. Hematoxylin and eosin (H&E) staining was performed on tissue sections for morphological analysis. Tumour regions were identified by an experienced pathologist. Histopathological information is provided in ***Table 5.1***.

Table 5.1: Clinical information for the Indigenous (n=4) and non-Indigenous (n=4) vulvar squamous cell carcinoma tissues (tumour areas circled in black).

Indigenous Patients

ID	p09-502329	p10-525537	p06-513210	p08-528562
Age	58	52	39	25
HPV	No	No	No	No
FIGO stage	IA	IB	III	III
Grade	3	2	1	2
LNs	*neg*	*neg*	*pos*	*pos*
Annotation				

Non-Indigenous Patients

ID	p08-509939	p13-539105	p11-531450	p13-6644191
Age	67	37	66	75
HPV	No	No	No	No
FIGO stage	IB	IB	III	III
Grade	2	2	3	3
LNs	*neg*	*neg*	*pos*	*pos*
Annotation				

Analysis of the clinical information revealed some unique pathological and morphological characteristics within the VSCC Indigenous cohort. The diagnostic age of the Indigenous patients was significantly younger (median at 39) than that of non-Indigenous patients (median at 66), which is also younger than the national average diagnostic age of 68 for vulvar cancer (Condon et al., 2003). The Indigenous cohort also demonstrated unique morphological patterns as detailed in the report from the pathologists, summarized as follows:

Patient 1 (No. 09-503290): The patient sample was identified as FIGO stage IB HPV-independent squamous cell carcinoma of tumour grade 3. The malignant tumour growth

covered almost the entire vulvar skin surface and a large area of fat tissue was found under the vulvar skin. The pathological analysis also highlighted that widespread dVIN on other samples from this patient, but dVIN lesions were not present on the tissue block used for this study.

Patient 2 (No. 10-525537): Early stage (stage IB, Grade 2) HPV-independent Indigenous VSCC patient with LNM negative. Tissue histology was described as a central, hyperkeratotic, ulcerate like lesion surrounded by several pieces of fatty tissue, dVIN lesions were not found in the specimen sample of this patient.

Patient 3 (No.08-528562): The youngest HPV-independent and LNM positive Indigenous VSCC patient in the cohort diagnosed with later stage (Stage III, Grade 2) VSCC at 25. The vulvar tissue section showed hair-bearing skin ellipse and the skin surface was replaced by identical polypoid and invaded hyperkeratotic VSCCs with underlying fat tissues. The tumour region was surrounded by lymphoplasmacytic inflammation, precancerous dVIN lesions were not found in this patient.

Patient 4 (No.06-513210): HPV-independent Indigenous patient diagnosed with LNM positive VSCC (Stage III). Well-differentiated (Grade 1) invasive cancer was not limited to the vulva and was present in the nearby skin. The pre-cancerous dVIN lesion was found in the tumour margin and subcutaneous fat was located under the irregular skin ellipse.

H&E staining of the Indigenous tissue samples all showed identical keratinizing VSCC morphology, including the co-existence of precancerous dVIN lesions and hypo-keratinization structures at the tumour centre. The HPV-independent Indigenous VSCCs also displayed widely spread malignant cancer and underlying fatty tissue. Additionally, of the four HPV-independent Indigenous VSCC samples, two were identified with positive LNM at the young ages of 25 and 39. The higher incidence rates are likely related to the specific morphology observed in the Indigenous VSCC samples which is directly related to the rapid development of cancer.

5.3.2 Relative quantification of HPV-independent Indigenous and non-Indigenous VSCC tissues

Label-free quantification (LFQ) analysis using *nano*-LC-MS/MS was performed on a cohort of VSCC tissues from four Indigenous and four Non-Indigenous patients in duplicate runs. Briefly, tissues from the eight patients were prepared as described in this chapter section 5.2. A total of 6033 peptides attributed to 1406 unique proteins were detected and quantitated using the MaxQuant software (v1.6.0.1) (***Supplementary S5.2***). The distribution of the number of quantified proteins in each patient is presented in ***Figure 5.1*****,** which shows a normal distribution relationship between protein numbers and their $\log_2$ transformed LFQ intensities.

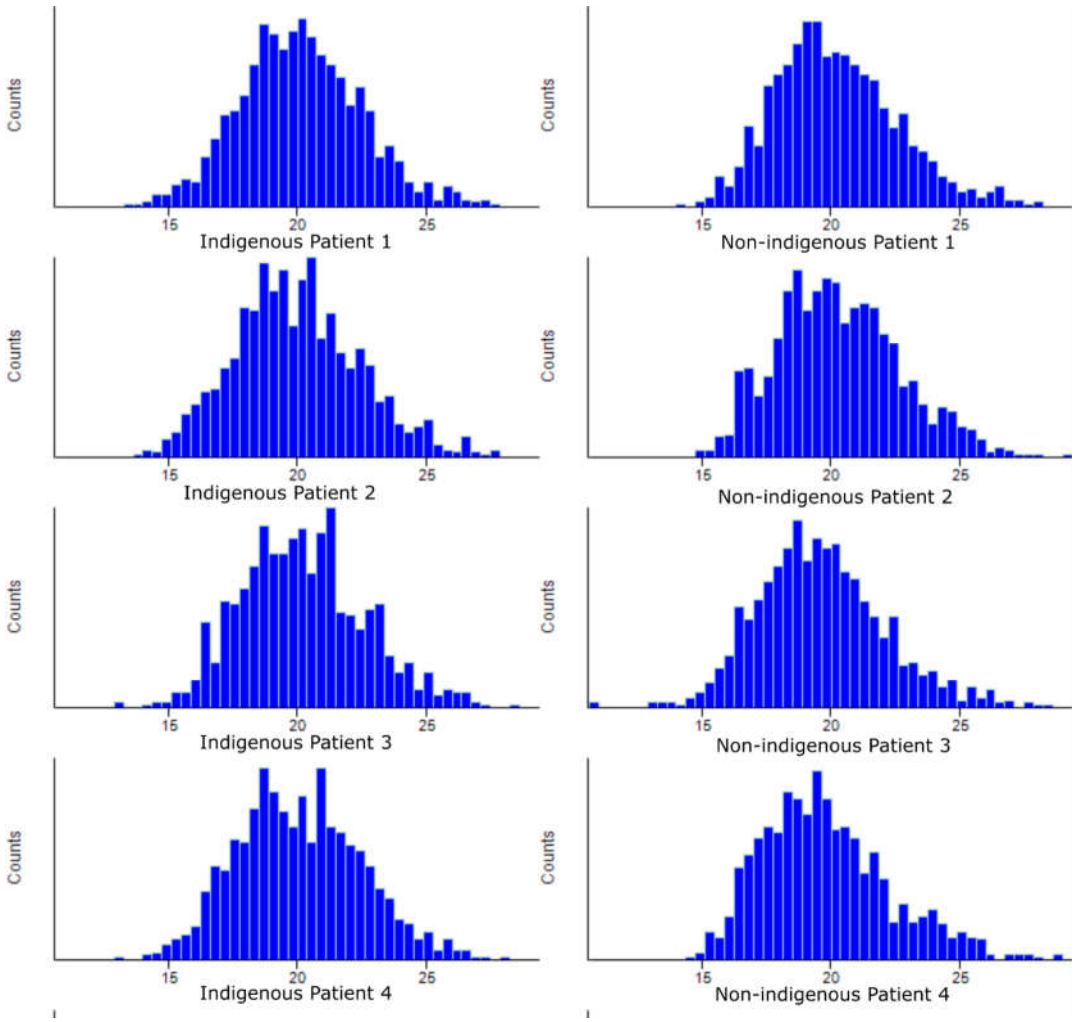


Figure 5.1: Peptide abundance distribution in four Indigenous patients (Left) and four non-Indigenous patients (Right) VSCC tissue samples. Y-axis: number of quantified proteins. X-axis: $\log_2$ transformed LFQ intensities for eight patient samples.

Further processing and interpretation of LFQ data of proteins were achieved in Perseus, according to Chapter 2 section 2.2.3.6. After removing proteins that were only detected in one patient of each cohort, there was 1195 proteins identification across the eight patients. The relative abundance levels of these proteins were then compared between the Indigenous and non-Indigenous cohorts using unpaired student t-test. Given the small sample number used in this study, protein candidates as determined by the t-test results ($p<0.05$) were further filtered by minimum 4-fold intensity difference between the Indigenous and non-Indigenous cohorts to eliminate potential false discoveries. Furthermore, the unpaired t-test was also performed within Indigenous cohort patients to compare protein changes in LNM positive and negative stages. All results were interpreted using a volcano plot of $\log_2$ protein change ratios and $-\log_{10}$ *p*-values. Protein with significantly different $\log_2$ LFQ intensities between compared cohorts is summarized in clustering heatmaps.

5.3.3 LFQ comparison of Indigenous and non-Indigenous vulvar cancer patients

The LFQ analysis of four Indigenous and four non-Indigenous VSCC tissues (including both LNM positive and LNM negative samples), revealed 164 of the identified 1195 proteins to have significantly different abundance levels (***Figure 5.2***). Details of these proteins are provided in ***Supplementary file S5.3***.

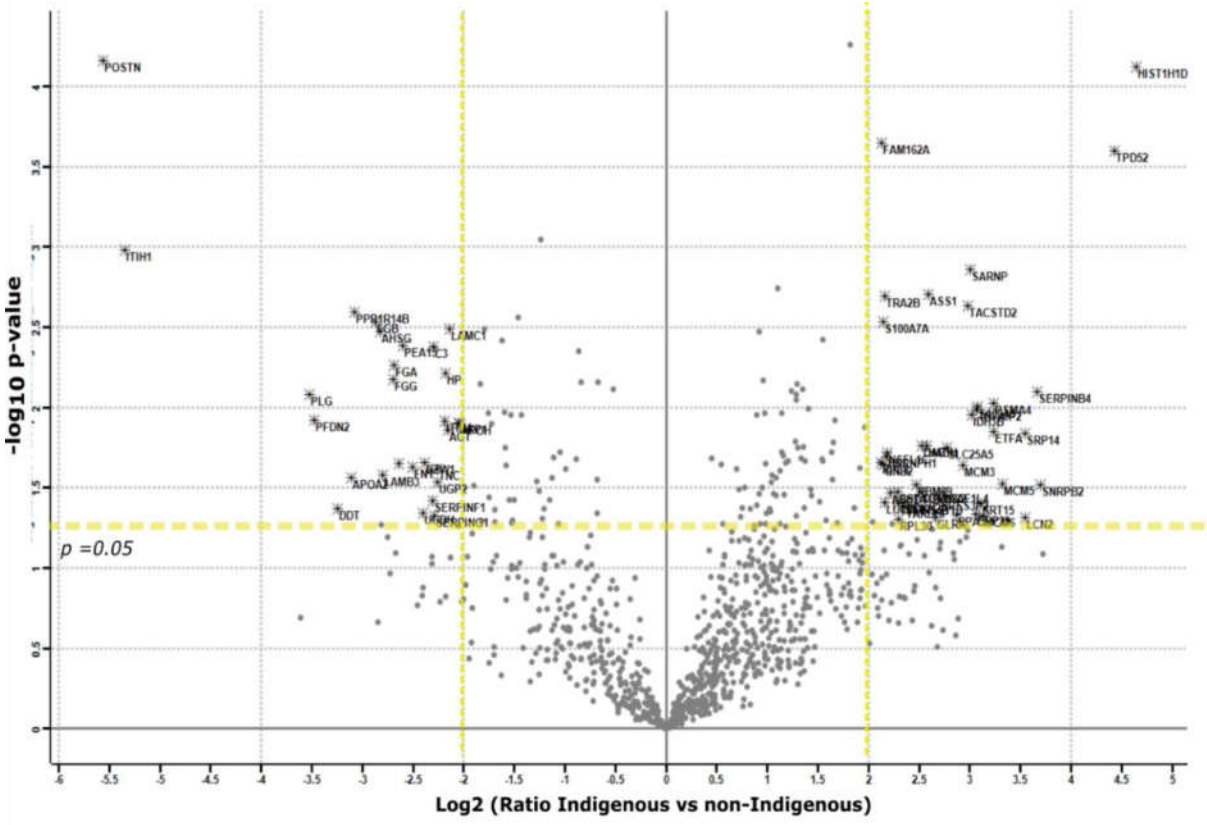


Figure 5.2: Volcano plot and student t-test results of protein abundance in the LFQ analysis between the four Indigenous and four non-Indigenous VSCC tissues. X-axis: intensity differences (log_2 transformed) between Indigenous patient group (right side) and non-Indigenous group (left side). Y-axis: *p*-values ($-log_{10}$ transformed) from student t-test results. Proteins passing the significant threshold ($p<0.05$) and showing more than 4-fold change are marked with stars and gene names.

After setting the criteria of a 4-fold change minimum with a significance level of $p \leq 0.05$ between the two cohorts, 70 proteins remained. When the LFQ intensities of the proteins were log_2 transformed, 42 of them, including Histone H1.3 (HIST1H1D), Tumor Protein D52 (TPD52), U2 small nuclear ribonucleoprotein B (SNRPB2), and Squamous cell carcinoma antigen 2 (SERPINB4), showed a higher than 2-fold log_2 LFQ intensity between the non-

Indigenous and Indigenous vulvar cancers, whilst Periostin (POSTN), Inter-alpha-trypsin inhibitor heavy chain H1 (ITIH1) and 26 other proteins displayed a decrease LFQ intensity in the Indigenous vulvar cancer compared with the non-Indigenous one (***Supplementary file S5.3***). The log_2 LFQ intensities of 70 differentially abundant proteins of interest across are shown in ***Figure 5.3***.

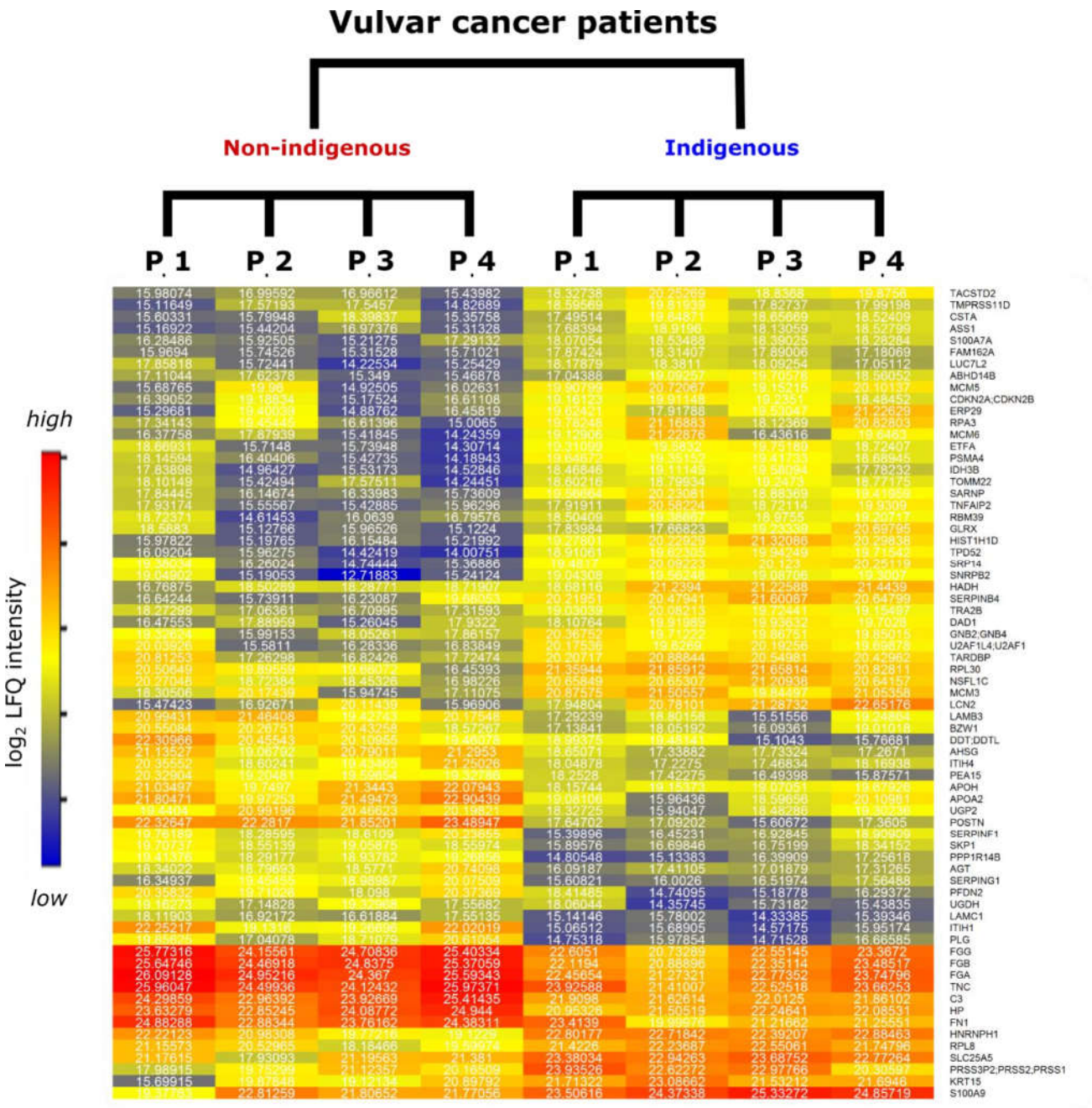


Figure 5.3: Clustering heatmap analysis of the proteins with a ≥4 fold abundance difference between the Indigenous and Non-Indigenous vulvar cancer tissues as determined by LFQ *nano*-LC- ESI-MS/MS. A scale bar

indicating changing log_2 level of intensity from low (blue, with the lowest intensity of 12) to high (red, with the highest intensity of 27) is shown to the left.

From this analysis, the protein periostin (POSTN) was observed at a significantly lower intensity level in all four of the Indigenous vulvar cancers compared to the non-Indigenous vulvar cancers. The log2 transformed LFQ intensities of POSTN was summed for Indigenous (mean 16.44 ± 1.290) and non-Indigenous (mean 22.49 ± 0.702), and graphed GraphPad Prism 6 (v008), revealing a large difference in the abundance of the protein between the two cohorts (***Figure 5.4***). These results suggest decreased POSTN levels may play a role in the development of vulvar cancer in the Indigenous patients.

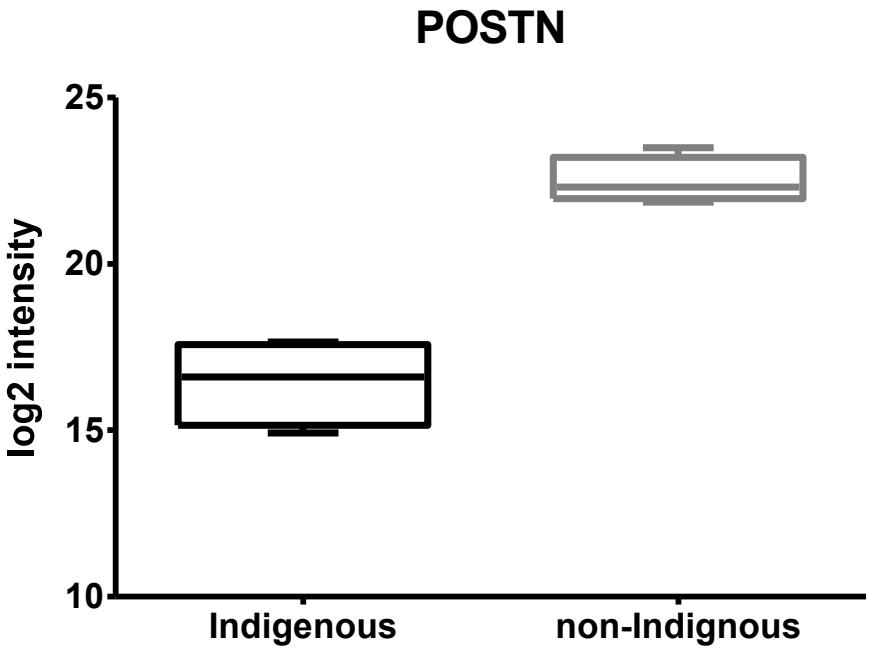


Figure 5.4. The protein POSTN was found to at significantly lower levels in vulvar cancer tissues collected from the Indigenous cohort (n=4) compared to the non-Indigenous cohort (n= 4).

Additionally, the LFQ analysis was also performed on the two LNM positive and two negative vulvar cancer patients within the Indigenous cohort, which revealed 54 of the identified 1195 proteins to have significantly different abundance levels (***Figure 5.5***). Details of these proteins are provided in ***Supplementary file S5.3***.

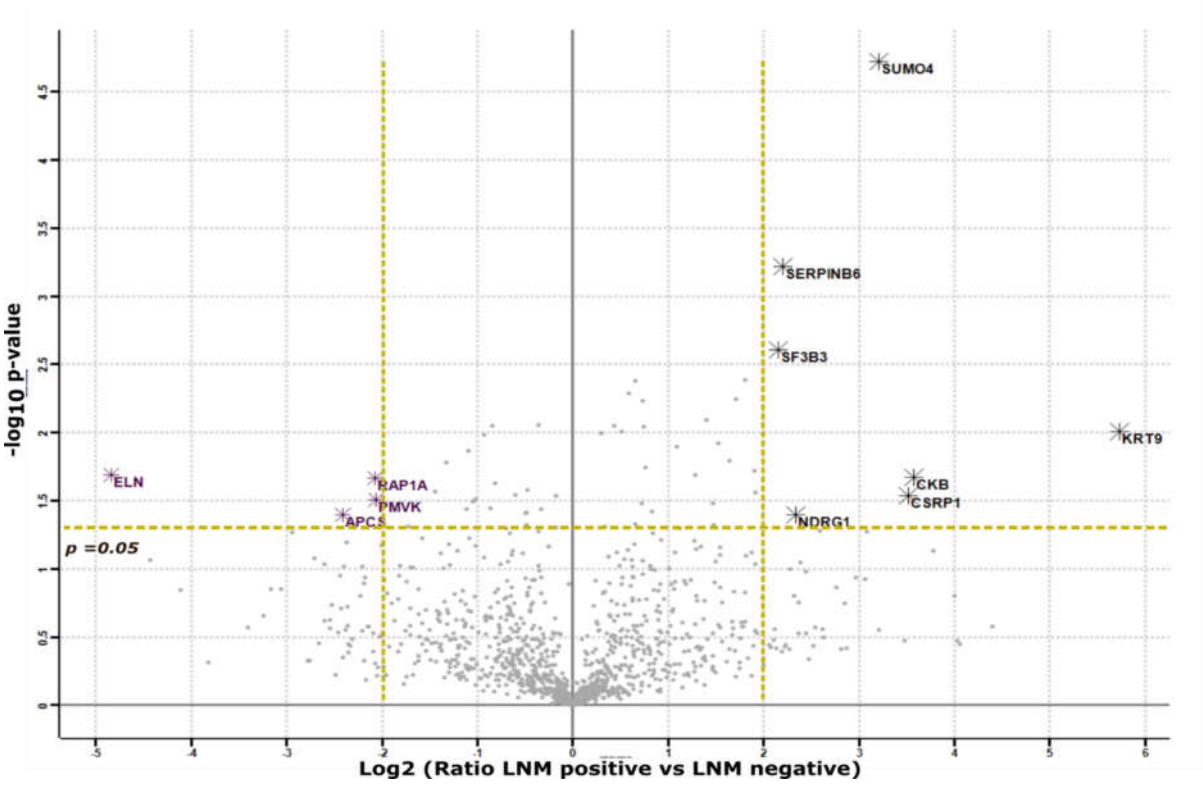


Figure 5.5: Volcano plot and student t-test results of protein abundance difference in the LFQ analysis between the two LNM positive and two negative Indigenous VSCC tissues. X-axis: intensity differences (log_2 transformed) between LNM positive Indigenous patients group (right side) and LNM negative Indigenous group (left side). Y-axis: *p*-values ($-log_{10}$ transformed) from student t-test results. Proteins passing the significance threshold ($p<0.05$) and more than 4-fold change are marked with stars and gene names.

After setting the criteria of a 4-fold change minimum with a significant level of $p \leq 0.05$ between LNM positive and negative Indigenous patients, in total 11 proteins remained, including seven proteins, namely cytokeratin9, Creatine kinase B-type, Cysteine and glycine-rich protein 1, Small ubiquitin-related modifier 4, Protein NDRG1, Serpin B6, and Splicing factor 3B subunit 3, showed a higher log_2 transformed LFQ intensities in two LNM positive cancers, contrast to four proteins, including Ras-related protein Rap-1A (RAP1A), Serum amyloid P-component,

Elastin and Phosphomevalonate kinase displayed a decrease LFQ intensities between the LNM positive and negative Indigenous vulvar cancers(***Supplementary file S5.3***). The log_2 LFQ intensities of 11 differentially abundant proteins of interest across patients are shown in ***Figure 5.6***.

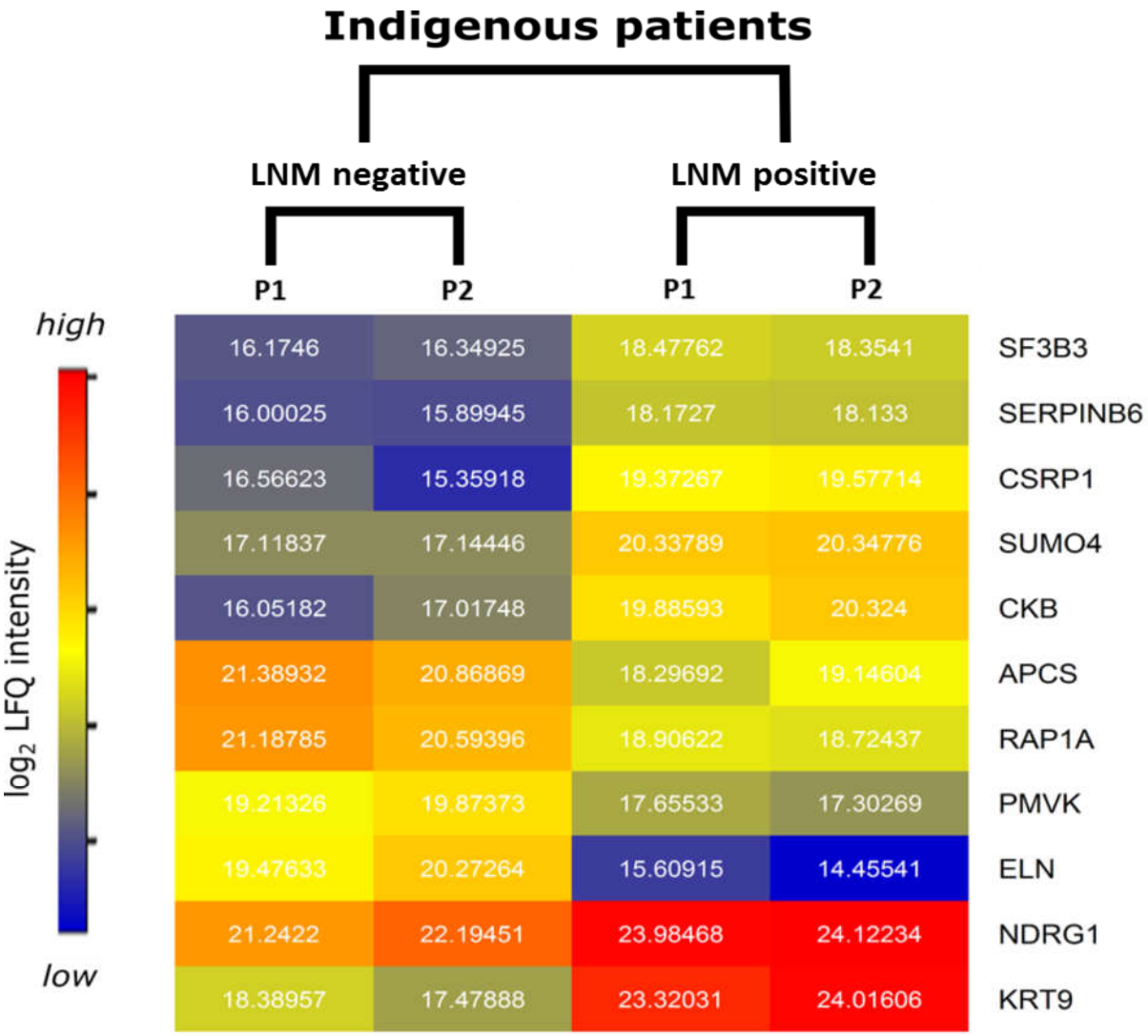


Figure 5.6: Clustering heatmap analysis of the proteins with a ≥ 4 fold abundance difference between the LNM positive and negative Indigenous vulvar cancer tissues as determined by LFQ *nano*-LC- ESI-MS/MS. A scale bar indicating the changing log_2 level of intensity from low (blue, with the lowest intensity of 14) to high (red, with the highest intensity of 25) is shown to the left.

With the aim to further explore the biological processes and signalling pathways that are potentially involved with the development of vulvar cancer in the Indigenous community, proteins of interests highlighted by the unpaired student t-test comparing the four Indigenous and four non-Indigenous VSCC samples were submitted to the Reactome pathway for analysis.

5.3.4 Biological implications of Indigenous vulvar cancer proteomic variations

To assess the biological pathways implicated by the proteome analysis of the Indigenous vulvar cancer tissues, a Reactome pathway analysis was performed on the 70 proteins of interests described above, of which 42 were increased in abundance and 28 were decreased in abundance in the Indigenous vulvar cancers.

The 42 proteins that were significantly increased in abundance were associated with the following functions: cell cycle (7); cellular responses to stimuli (8), developmental biology (6), disease (4), DNA repair (1), DNA replication (1), gene expression (6), hemostasis (1), immune system (7), metabolism (12), metabolism of proteins (8), metabolism of RNA (8), mitophagy (1), neuronal system (1), reproduction (1), signal transduction (3), transport of small molecules (3) and vesicle-mediated transport (1). A map of all biological processes and representative signalling pathways are shown in ***Supplementary S5.4***. On the other hand, the 28 proteins significantly decreased in the Indigenous vulvar cancers were associated with the following functions,: cell cycle (1), cell-cell communication (1), circadian clock (1), developmental biology (1), disease (5), DNA replication (1), extracellular matrix organization (8), gene expression (1), hemostasis (9), immune system (10), metabolism (6), metabolism of proteins (12), signal transduction (12), transport of small molecules (2) and vesicle-mediated transport (1). A map of revealed biological processes and representative signalling pathways are also displayed in ***Supplementary S5.4***.

Proteins relating to the immune system were further explored, given HPV-independent vulvar cancer is thought to arise from lichen sclerosis (LS), an inflammatory disorder. There were a considerable number of proteins highlighted that related to the immune system (***Figure 5.7***), with six of 17 immune-associated proteins (Protein S100A9, Signal recognition particle 14 kDa protein, Neutrophil gelatinase-associated lipocalin, Haptoglobin, Alpha-2-HS-glycoprotein and Complement C3) involved in neutrophil degranulation which has been identified as the hallmark of chronic inflammation (Doring, 1994). Many signalling pathways revealed in the Reactome analysis were closely associated with cellular metabolism, including both protein and RNA metabolism. In total, 20 proteins of interests were reported in protein metabolism related pathways, especially the signalling pathways of insulin-like growth factor transport and binding regulation (***Figure 5.8***). Fibrinogen alpha chain, Fibronectin and other seven proteins of interests were highlighted in the Reactome Pathway Analysis as potential insulin-like growth

factor pathways related proteins. Other metabolism regulation proteins, such as Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-2 and Angiotensinogen, were mainly found in the carbohydrate and vitamin metabolism. POSTN and seven other proteins found to be decreased in abundance in the Indigenous vulvar cancers are related to extracellular matrix organization, in particular, integrin cell surface interactions and the regulation of collagen binding. Another biological process that involved more than 10 differentially abundant proteins was the signal transduction, including signalling pathways of receptor tyrosine kinases, GPCR, WNT, Hedgehog, Integrin and MAPK signalling cascades.

In summary, proteins relating to the immune system, metabolism and extracellular matrix organizations, were found to be differentially abundant between the Indigenous and non-indigenous vulvar cancer cohorts, which could be related to high incidence rate and rapid development of vulvar cancer in Indigenous women.. Analysis of the spatial distribution of the proteins of interests was achieved using MALDI-MSI.

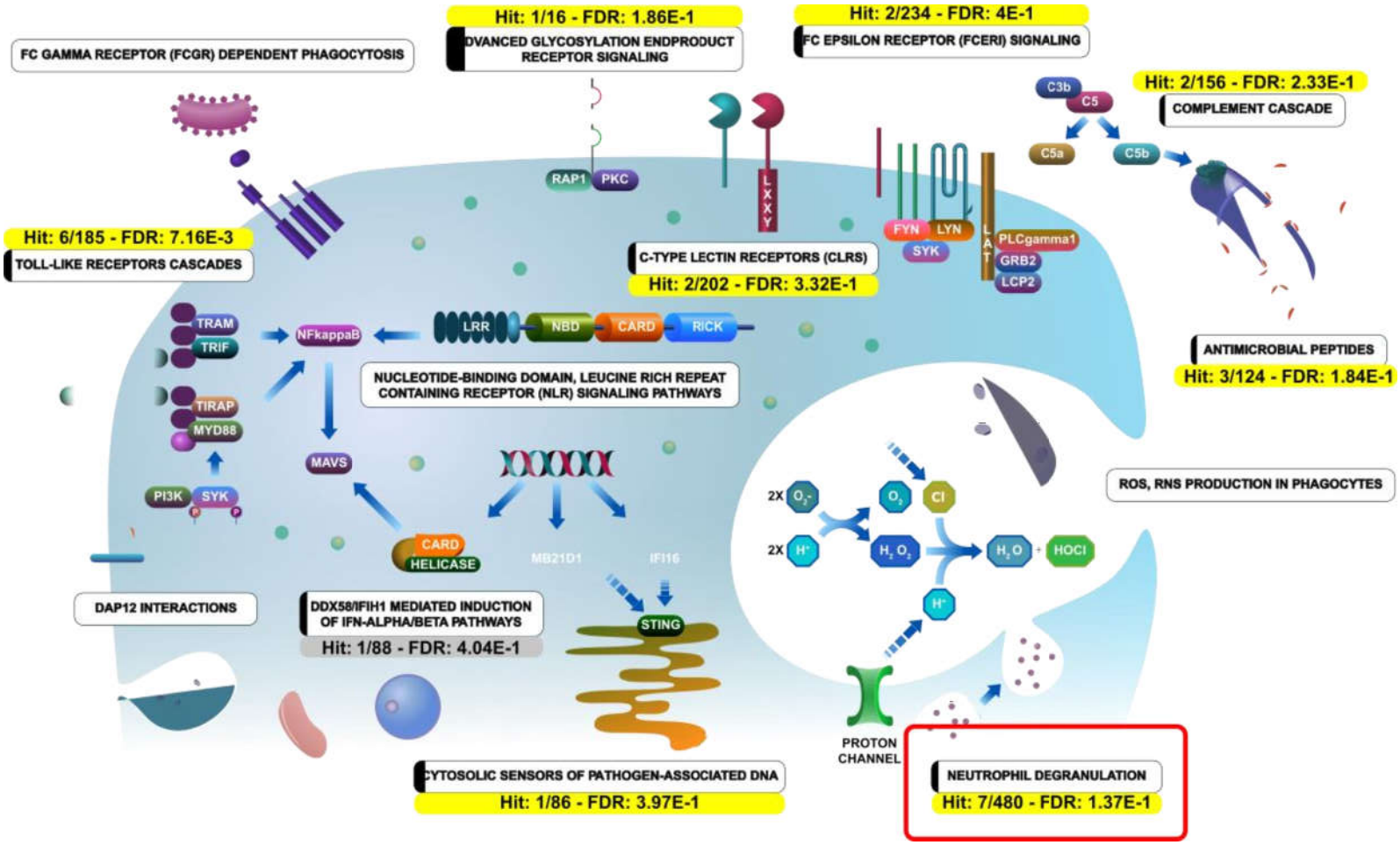


Figure 5.7: The innate immune system signalling pathways involved with the proteins of interest, found to be differentially abundance between Indigenous and non-Indigenous cancers in the Reactome Pathway Analysis. The number of proteins that hit this pathway in the database and the false discovery rate are shown in yellow rectangles, and the pathway involved with the highest number of protein hits are highlighted using a red circle.

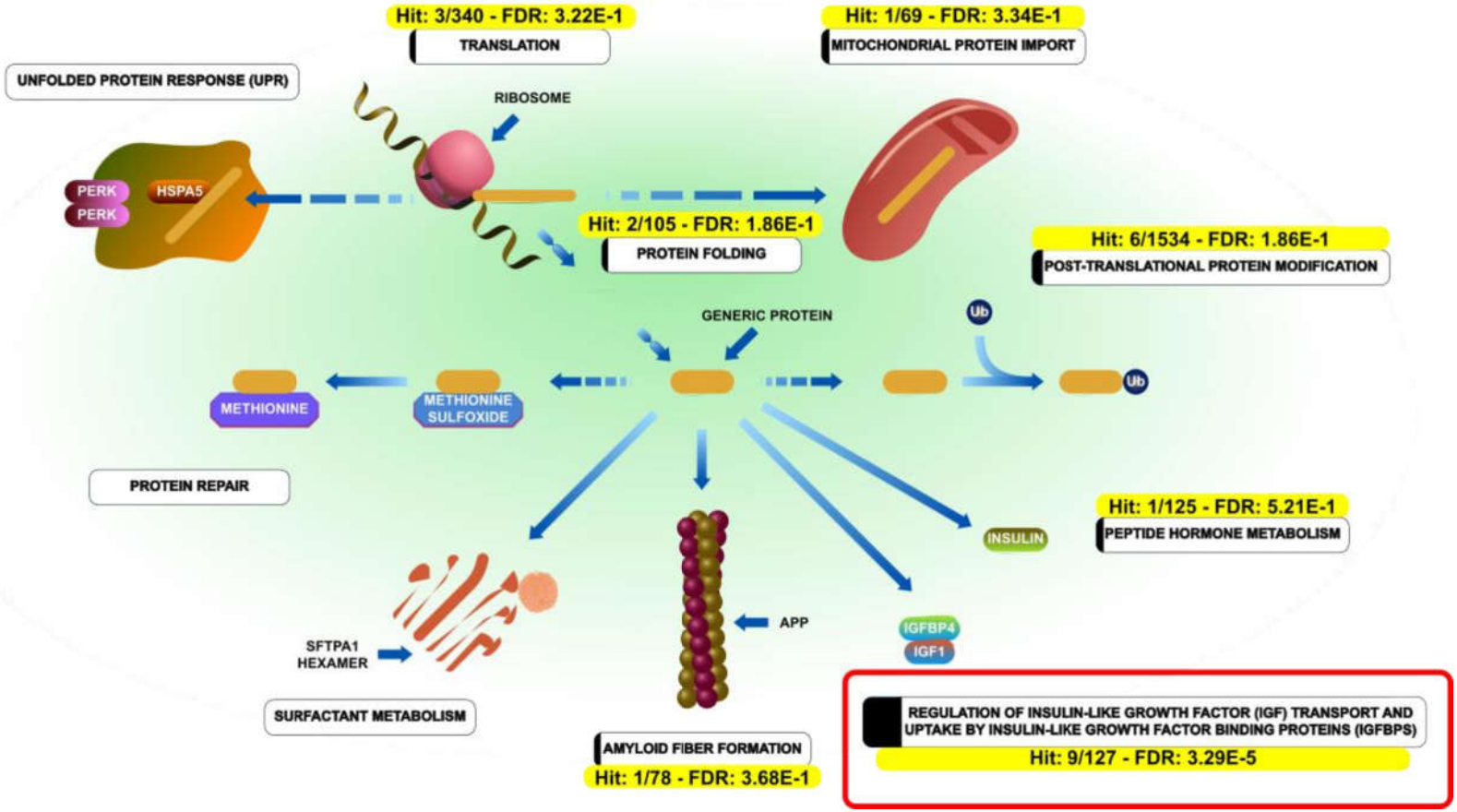


Figure 5.8: The metabolism signalling pathways involved with the proteins of interest, found to be differentially abundance between Indigenous and non-Indigenous cancers in the Reactome Pathway Analysis. The number of proteins that hit this pathway in the database and the false discovery rates are shown in yellow rectangles, and the pathway involved with the highest number of protein hits are highlighted using a red circle.

5.3.5 Validation of spatial distribution of proteins of interests in MALDI-MSI

The peptide identified from the trypsin digested Indigenous VSCC tissues by *nano*-LC-ESI-MS/MS analysis (section 5.2.2) were used to construct a peptide library. The acquired data were processed and searched using a Human SwissProt database in the MASCOT server according to the searching parameters described in Chapter 2 section 2.2.3.4. Peptide search results were exported and combined into a comprehensive file by an in-house R script (Chapter 2, section 2.2.4.5). In total, 6748 unique peptides from 1853 proteins were identified with a peptide expected score of less than 0.05. The identified peptides were used to construct an Indigenous vulvar cancer peptide library for MALDI-MSI peaks matching, as listed in ***Supplementary file S5.1***.

To explore the spatial distribution of the protein candidates from LFQ analysis described above, MALDI-MSI was performed on consecutive tissue sections of the same VSCC samples (section 5.2.1). The tumour tissue regions of interests (ROIs) were annotated by experienced pathologists. MALDI-MSI was performed on the Indigenous and non-Indigenous tumour sections in duplicate and resulted in the acquisition of 10000 to 20000 spectra per run. Sample peaklists were processed and exported in flexAnalysis (version 4.0, Bruker, Germany). Spectra processing was accomplished in the SCiLS Lab (2016b), including raw spectra alignment, total ion count (TIC) normalization and peak picking from the centre of the total aligned spectra. A Wilcoxon-rank test was performed on the peakgroups to analyze peak intensity differences between the cohorts. In total, 430 peaks were picked from the aligned spectra, of which 420 were found to be significantly different in the Wilcoxon-rank test ($p<0.001$) in SCiLS Lab. Peaks with significant intensity differences between the sample group were subject to Receiver Operating Characteristic (ROC) analysis and boxplots (***Supplementary file S5.5***).

In total, 230 of 420 peptide peaks of interest could be matched back to 95 candidate proteins, including 11 proteins found to be of interest from the LFQ analysis comparing Indigenous and non-Indigenous cancers, in the assembled peptide library (described in section 5.3.2). With the aim of examining the differential protein abundances between the Indigenous and non-Indigenous cancers in the MALDI-MSI results, ROCs were generated for the *m/z* values that matched to the tryptic peptides from the 11 proteins of interest described above (***Supplementary 5.6)***. Matched peptide peaks which were in the disagreement when comparing

the MALDI-MSI and LFQ analyses were removed from further analysis. Additionally, peptide peaks that lacked discriminative power across the patient cohort in the MALDI-MSI analyses (0.3<AUC <0.7) were also removed from the further analyses.

After filtering the data based on their MALDI-MSI AUC scores, four *m/z* values identified from four proteins of interest highlighted in the LFQ comparison between Indigenous and non-Indigenous VSCC remained. These *m/z* values matched to peptides from the protein Haptoglobin (HP), which is related to neutrophil degranulation, Fibrinogen alpha chain (FGA) and Fibronectin (FN1), which are involved in the insulin-like growth factor, and the extracellular matrix organizer, Periostin (POSTN). Additionally, two other *m/z* values matched to tryptic peptide from the protein NDRG1 and RAP1A, which were found to be differentially abundant between the LNM positive and negative Indigenous patients ***(Figure 5.6)***. Details of those identified peptides mentioned above are displayed in ***Table 5.2.***

Table 5.2: Identification of discriminative *m/z* values from the MALDI-MSI analyses by matching a peptide library created from *nano*-LC-ESI-MS/MS profiling and label-free quantification.

Protein		**Label-free quantification *nano*-LC-ESI-MS/MS**				**MALDI-MSI**			
Protein Name	Protein ID	Sequence	Unique Proteins	Measured Mass [M]	Student's T-test Difference Indigenous vs non-Indigenous	[M+H]	mass error (Da)	AUC scores	
								Indigenous (n=4) vs non-Indigenous (n=4)	Indigenous LNM positive (n=2) vs negative (n=2)
Fibrinogen alpha chain (FGA)	P02671	QHLPLIK	Yes	847.528	-2.688	848.52	-0.008	0.307	0.409
Fibronectin (FN1)	P02751	VFAVSHGR	Yes	871.4664	-2.506	872.44	-0.0264	0.267	0.436
Haptoglobin (HP)	P00738	VGYVSGWGR	Yes	979.5239	-2.182	980.47	-0.0539	0.281	0.339
Periostin (POSTN)	Q15063	NGAIHIFR	Yes	927.4926	-5.561	928.48	-0.0126	0.271	0.413
Protein NDRG1 (NDRG1)	Q92597	SHTSEGAHLDITPNSG AAGNSAGPK	Yes	2375.1184	2.335	2376.11	-0.0084	0.779	0.713
Ras-related protein Rap-1A (RAP1A)	P62834	LVVLGSGGVGK	Yes	984.5968	-2.076	985.60	0.0032	0.39	0.319

From the MALDI-MSI peaklist, three *m/z* values were matched to intact masses of tryptic peptide sequences from POSTN, which showed differential protein abundance levels between the Indigenous and non-Indigenous VSCC tissues. One MALDI-MSI *m/z* value matched the POSTN peptide NGAIHIFR (*m/z* 927.48) in the peptide library. A MALDI-MSI annotated ion intensity map for *m/z* 928.48, which matches to the POSTN peptide, NGAIHIFR, is shown in ***Figure 5.9***. The ROC curve and ion intensity boxplot for the peptide at *m/z* 928.48 (AUC=0.271) was found to be lower intensity in the Indigenous cancer tissues as compared to non-Indigenous cancer tissues.

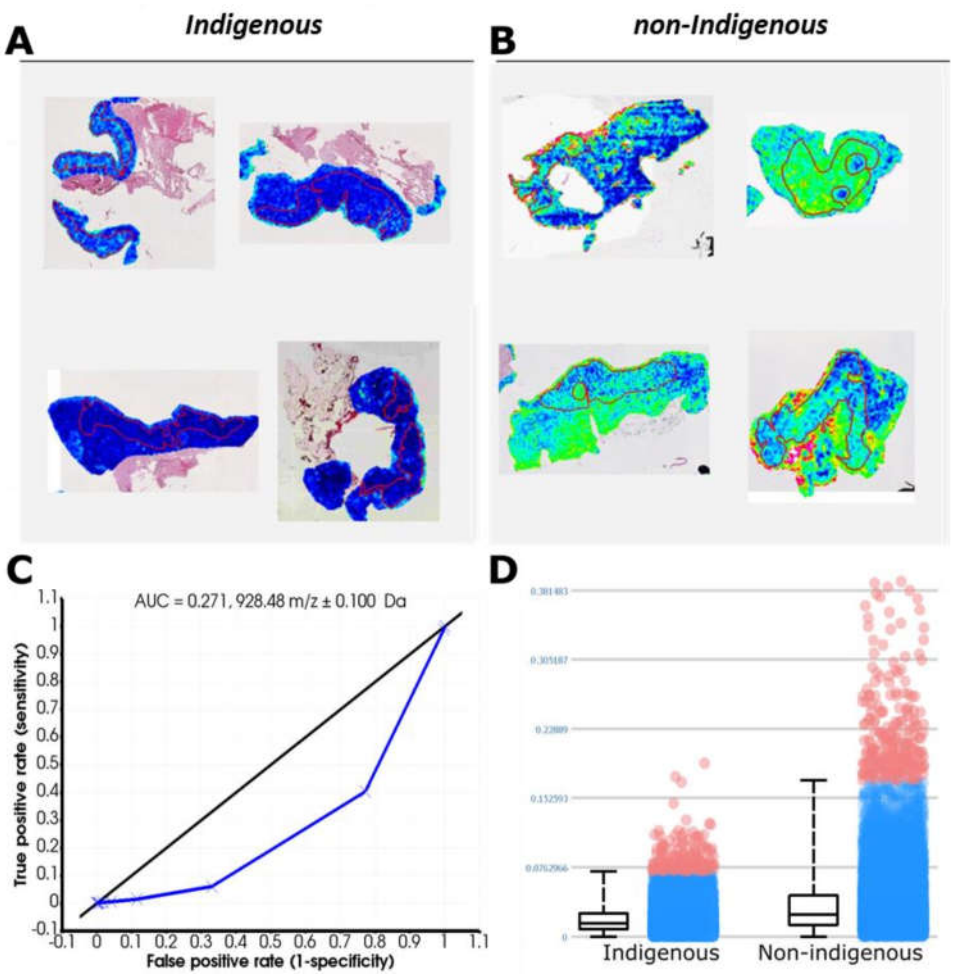


Figure 5.9: MALDI-MSI annotated ion intensity map from a representative Indigenous (n=4) and non-Indigenous (n=4) patient VSCC tissue section for POSTN peptide *m/z* 928.48. (A) Ion intensity map for *m/z* 928.48, which matches to POSTN peptide NGAIHIFR. Regions of VSCC in four Indigenous cancers colored in red; (B) Ion intensity map for *m/z* 928.48 detected in VSCCs (red) from non-Indigenous vulvar cancers (C) ROC curve and (D) boxplot analysis of the MALDI-MSI data for *m/z* 928.48 comparing Indigenous and non-Indigenous VSCC samples, AUC = 0.271; Blue dots represent the *m/z* 928.48 spectra intensity interval between lower and upper quartiles. Red dots represent spectra intensities outside of the intervals.

From the MALDI-MSI peaklist, one of the *m/z* values matched to an intact tryptic peptide from FGA, which showed differential protein abundance levels between the Indigenous and non-Indigenous VSCC tissues. One MALDI-MSI *m/z* value matched the FGA peptide QHLPLIK (848.52) in the peptide library. A MALDI-MSI annotated ion intensity map for *m/z* 848.52, which matches to the FGA peptide, QHLPLIK, is shown in ***Figure 5.10***. The ROC curve and ion intensity boxplot for the peptide at *m/z* 848.52 (AUC=0.307) was found to be lower intensity in the Indigenous cancer tissues as compared to non-Indigenous cancer tissues.

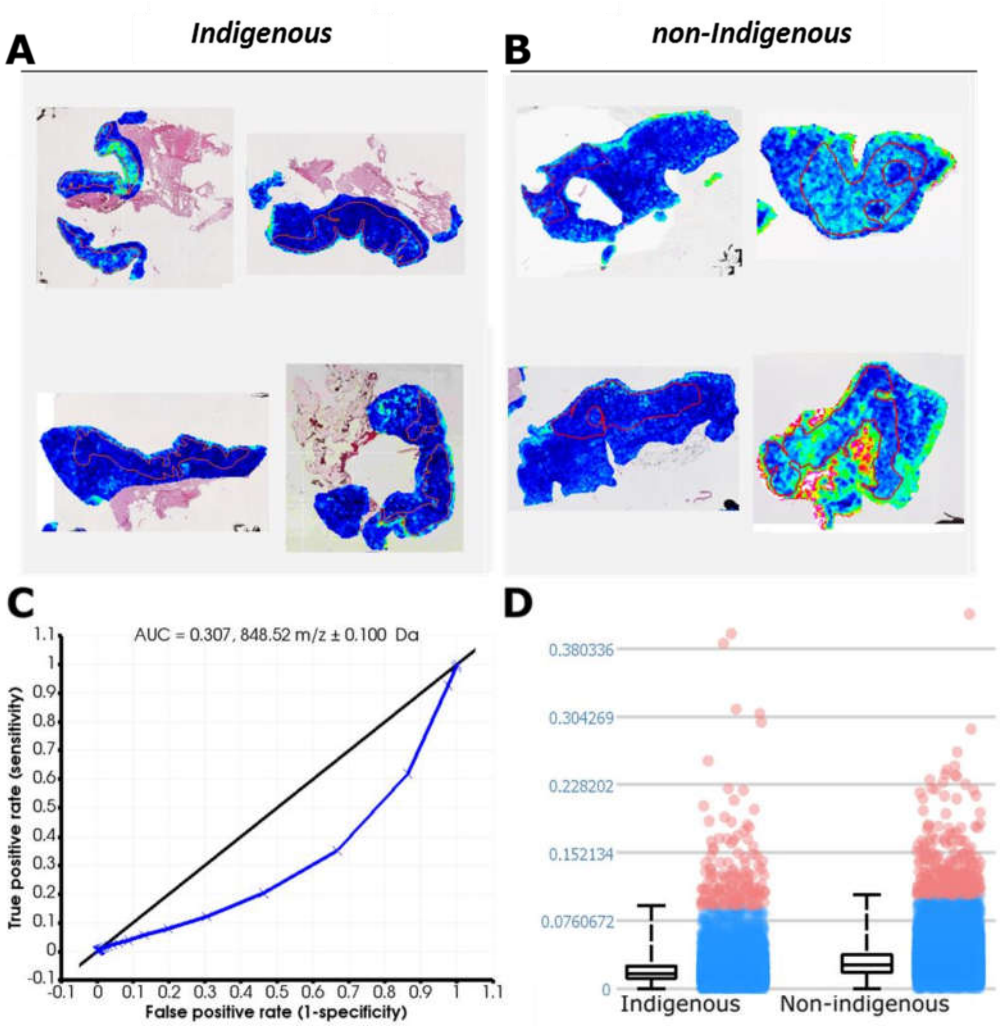


Figure 5.10: MALDI-MSI annotated ion intensity map from a representative Indigenous (n=4) and non-Indigenous (n=4) patient VSCC tissue section for FGA peptide *m/z* 848.52. (A) Ion intensity map for *m/z* 848.52, which matches to FGA peptide QHLPLIK. Regions of VSCC in four Indigenous cancers colored in red; (B) Ion intensity map for *m/z* 848.52 detected in VSCCs (red) from non-Indigenous vulvar cancers (C) ROC curve and (D) boxplot analysis of the MALDI-MSI data for *m/z* 848.52 comparing Indigenous and non-Indigenous VSCC samples, AUC = 0.307; Blue dots represent the *m/z* 848.52 spectra intensity interval between lower and upper quartiles. Red dots represent spectra intensities outside of the intervals.

From the MALDI-MSI peaklist, one of the *m/z* values matched to the intact tryptic peptide sequences from FN1, which showed differential protein abundance levels between the Indigenous and non-Indigenous VSCC patient tissues. One MALDI-MSI *m/z* value matched the FN1 peptide VFAVSHGR (872.44) in the peptide library. A MALDI-MSI annotated ion intensity map for *m/z* 872.44, which matches to the FN1 peptide, VFAVSHGR, is shown in ***Figure 5.11***. The ROC curve and ion intensity boxplot for the peptide at *m/z* 872.44 (AUC=0.267) was found to be lower intensity in the Indigenous cancer tissues as compared to non-Indigenous cancer tissues.

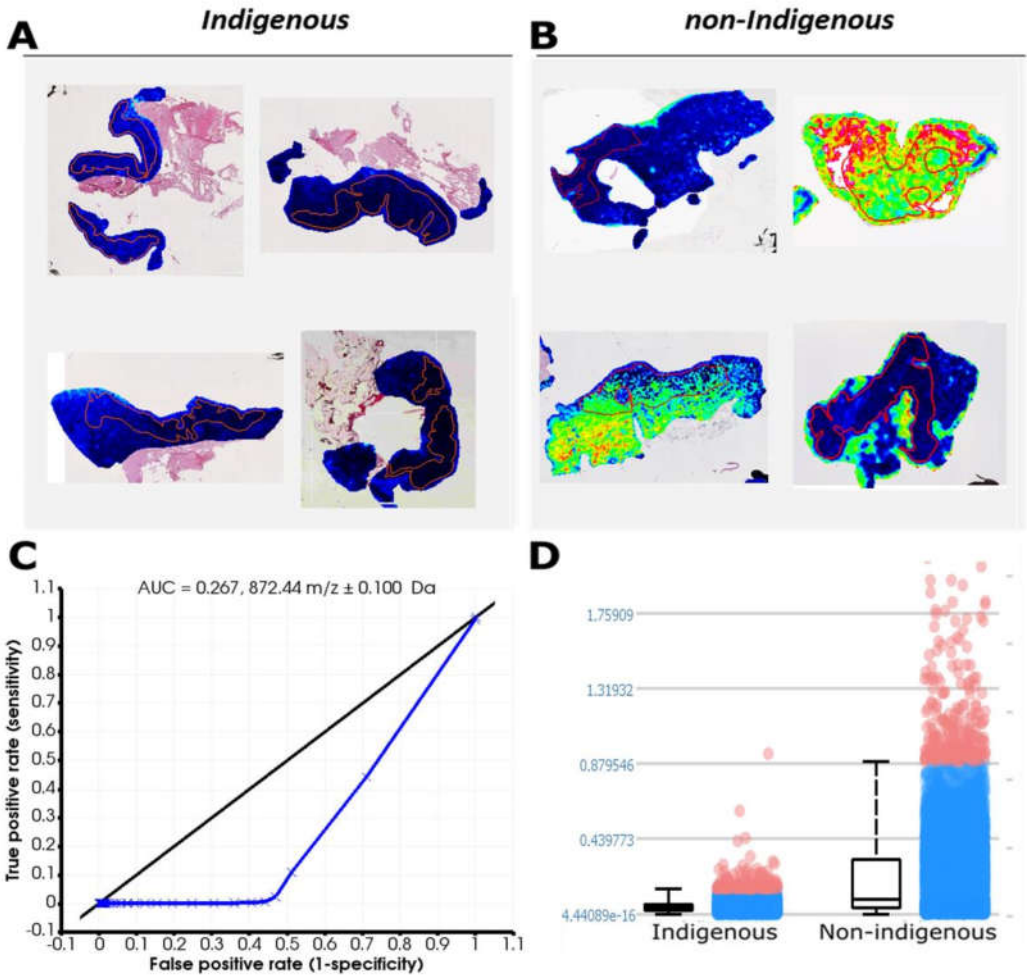


Figure 5.11: MALDI-MSI annotated ion intensity map from a representative Indigenous (n=4) and non-Indigenous (n=4) patient VSCC tissue section for FN1 peptide *m/z* 872.44. (A) Ion intensity map for *m/z* 872.44, which matches to FN1 peptide VFAVSHGR. Regions of VSCC in four Indigenous cancers colored in red; (B) Ion intensity map for *m/z* 872.44 detected in VSCCs (red) from non-Indigenous vulvar cancers (C) ROC curve and (D) boxplot analysis of the MALDI-MSI data for *m/z* 872.44 comparing Indigenous and non-Indigenous VSCC samples, AUC = 0.267; Blue dots represent the *m/z* 872.44 spectra intensity interval between lower and upper quartiles. Red dots represent spectra intensities outside of the intervals.

From the MALDI-MSI peaklist, one of the *m/z* values matched to the intact tryptic peptide sequences from HP, which showed differential protein abundance levels between the Indigenous and non-Indigenous VSCC patient tissues. One MALDI-MSI *m/z* value matched the HP peptide VGYVSGWGR (980.47) in the peptide library. A MALDI-MSI annotated ion intensity map for *m/z* 980.47, which matches to the HP peptide, VFAVSHGR, is shown in ***Figure 5.12***. The ROC curve and ion intensity boxplot for the peptide at *m/z* 980.47 (AUC=0.281) was found to be lower intensity in the Indigenous cancer tissues as compared to non-Indigenous cancer tissues.

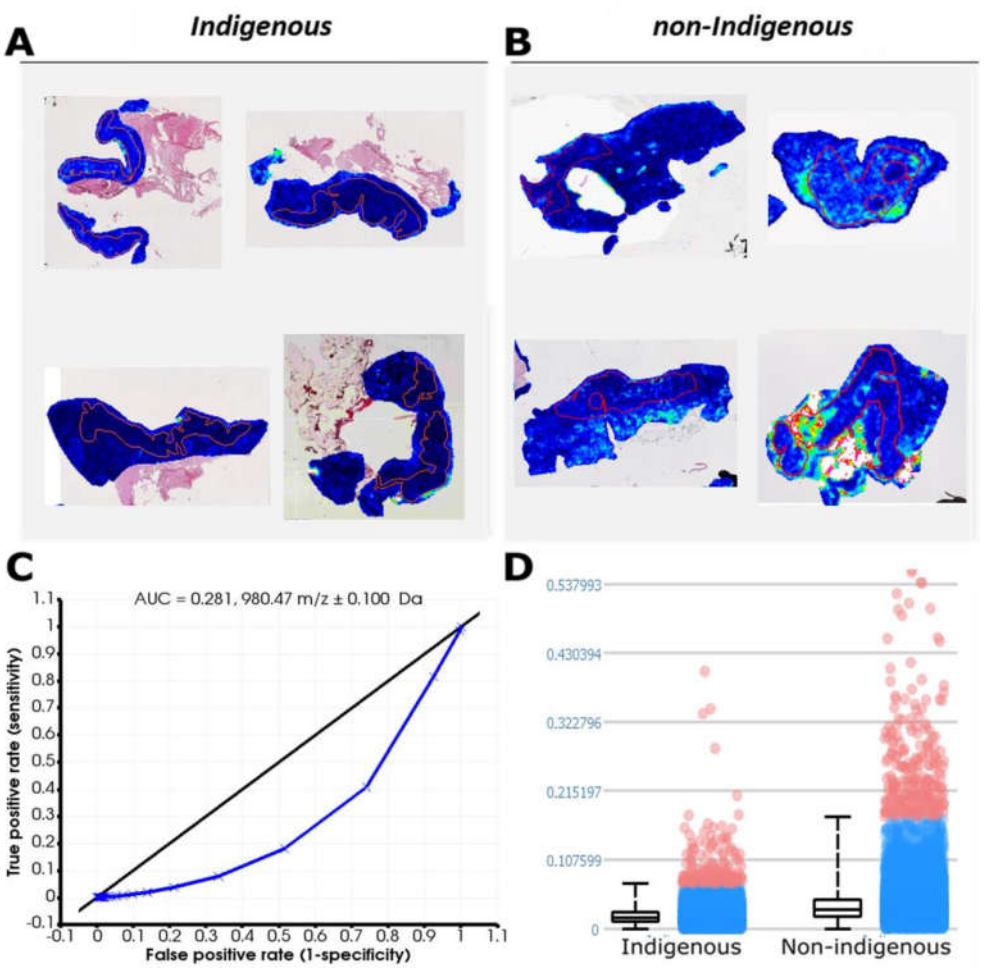


Figure 5.12: MALDI-MSI annotated ion intensity map from a representative Indigenous (n=4) and non-Indigenous (n=4) patient VSCC tissue section for FN1 peptide *m/z* 980.47. (A) Ion intensity map for *m/z* 980.47, which matches to HP peptide VGYVSGWGR. Regions of VSCC in four Indigenous cancers colored in red; (B) Ion intensity map for *m/z* 980.47 detected in VSCCs (red) from non-Indigenous vulvar cancers (C) ROC curve and (D) boxplot analysis of the MALDI-MSI data for *m/z* 980.47 comparing Indigenous and non-Indigenous VSCC samples, AUC = 0.281; Blue dots represent the *m/z* 980.47 spectra intensity interval between lower and upper quartiles. Red dots represent spectra intensities outside of the intervals.

From the MALDI-MSI peaklist, three *m/z* values matched to sequenced peptides from the protein NDRG1, which showed significantly differential abundance between the LNM positive and negative Indigenous VSCC tissues. One MALDI-MSI *m/z* value matched NDRG1 peptide, SHTSEGAHLDITPNSGAAGNSAGPK (2375.11), in the peptide library. A MALDI-MSI annotated ion intensity map for *m/z* 2376.11, which matches the NDRG1 peptide SHTSEGAHLDITPNSGAAGNSAGPK is shown in ***Figure 5.13***. The ROC curve and ion intensity boxplot for the peptide at *m/z* 2376.11 was found to be lower intensity in the LNM positive Indigenous cancer tissues as compared to LNM negative Indigenous cancer tissues.

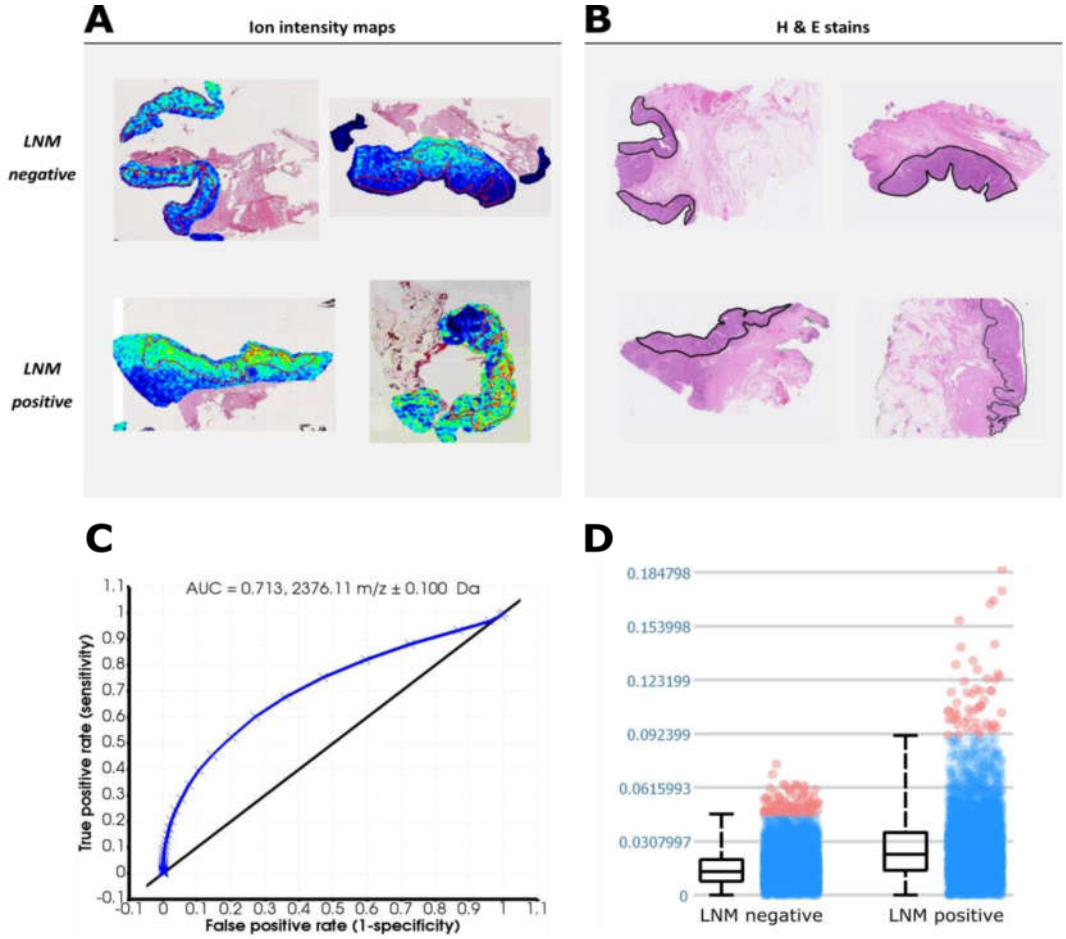


Figure 5.13: MALDI-MSI annotated ion intensity map from a representative Indigenous vulvar tissue section for NDRG1 peptide *m/z* 2376.11. (A) Ion intensity map for *m/z* 2376.11, which matches to NDRG1 peptide SHTSEGAHLDITPNSGAAGNSAGPK. Regions of VSCC of Indigenous samples in red; (B) H & E stain of Indigenous VSCC (black) samples used in MALDI-MSI data acquisition; (C) ROC curve and (D) boxplot analysis of MALDI-MSI data for *m/z* 2376.11 comparing LNM positive and negative Indigenous VSCC samples, AUC = 0.713; Blue dots represent *m/z* 1410.67 spectra intensity interval between lower and upper quartiles. Red dots represent spectra intensities outside of the intervals.

From the MALDI-MSI peaklist, four *m/z* values were matched to sequenced peptides from the protein RAP1A, which showed significantly differential abundance between the LNM positive and negative Indigenous VSCC tissues. One MALDI-MSI *m/z* value matched the RAP1A peptide LVVLGSGGVGK (984.60) in the peptide library. A MALD-MSI annotated ion intensity map for the matched peptide *m/z* 985.60, which matches to the RAP1A peptide LVVLGSGGVGK is shown in ***Figure 5.14***. The ROC curve (AUC=0.318) and ion intensity boxplot for the peptide at *m/z* 985.60 was found to be lower intensity in the LNM positive Indigenous cancer tissues as compared to LNM negative Indigenous cancer tissues.

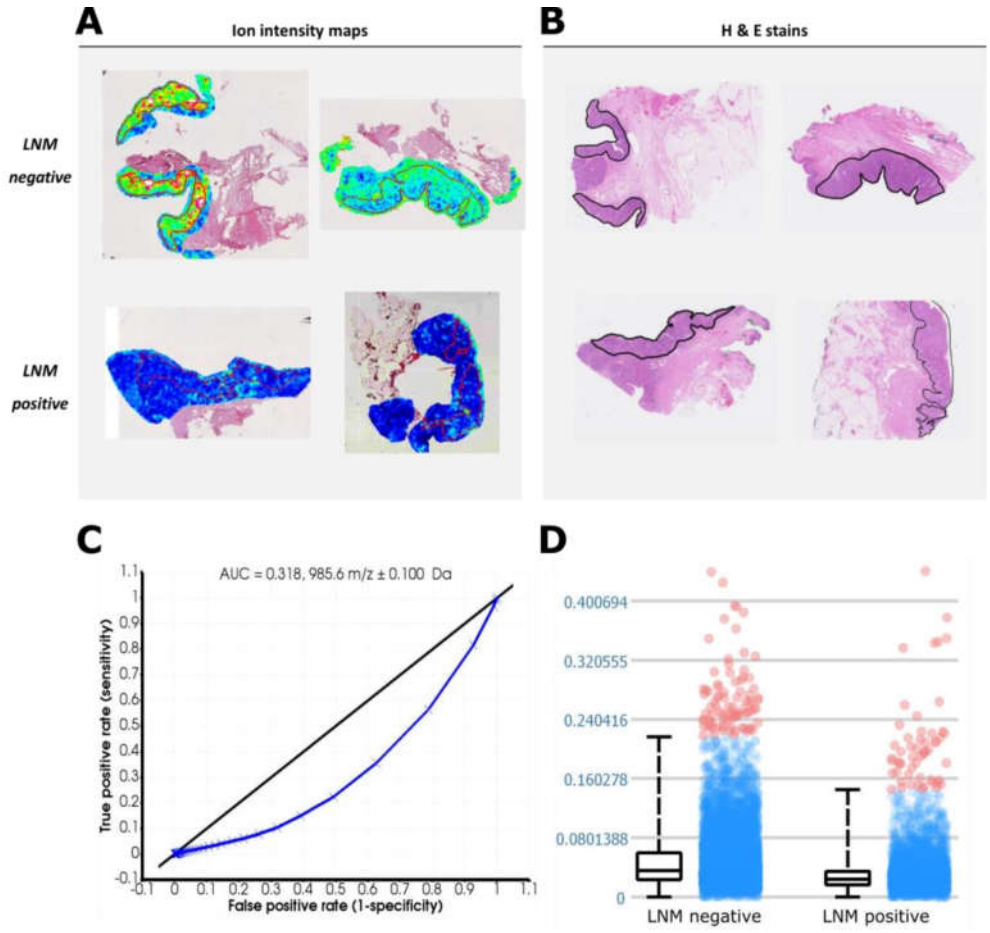


Figure 5.14: MALDI-MSI annotated ion intensity map from a representative Indigenous vulvar tissue section for RAP1A peptide *m/z* 985.60. (A) Ion intensity map for *m/z* 985.60, which matches to RAP1A peptide LVVLGSGGVGK. Regions of VSCC of Indigenous samples in red; (B) H & E stain of Indigenous VSCC (black) samples used in MALDI-MSI data acquisition; (C) ROC curve and (D) boxplot analysis of MALDI-MSI data for *m/z* 985.60 comparing LNM positive and negative Indigenous VSCC samples, AUC = 0.318; Blue dots represent *m/z* 985.60 spectra intensity interval between lower and upper quartiles. Red dots represent spectra intensities outside of the intervals.

The peptide peaks that matched to the proteins FGA, FN1 and HP were not highlighted in the correct tumour regions (ROIs) on the patient tissue samples in MALDI-MSI results. These inconsistent results are more likely caused by other tryptic peptides which cannot be separated and identified in the low mass resolution MALDI instrument. Therefore, these three proteins were removed from the protein candidate list. In summary, the MALDI-MSI validation of proteins POSTN, NDRG1 and RAP1A could be confirmed as having alternative abundance levels in Indigenous vulvar carcinogenesis and metastasis based on their spatial distribution in the peptide ion intensity maps. To further validate the Indigenous associated protein abundance levels and their spatial distribution on tissue IHC was performed on consecutively cut tissue sections.

5.3.6 Immunohistochemistry (IHC) validation of proteins of interest

The protein abundance levels and spatial distribution of POSTN, NDRG1 and RAP1A were further validated by performing immunohistochemistry (IHC) staining on consecutive VSCC tissue sections. The differential abundance of POSTN between Indigenous and non-Indigenous VSCC tissues revealed stronger positive staining of POSTN in the extracellular matrix of the non-Indigenous samples, while the Indigenous samples showed very little POSTN expression in the stroma. IHC analysis of POSTN revealed a significant difference ($p = 0.0016$) in the level of total positive staining between the Indigenous vulvar cancers (6.095 ± 1.439%) and non-Indigenous cancers (17.58 ± 2.850%). Results of the POSTN IHC analysis from four Indigenous and four non-Indigenous representative VSCC samples are provided in ***Figure 5.15.***

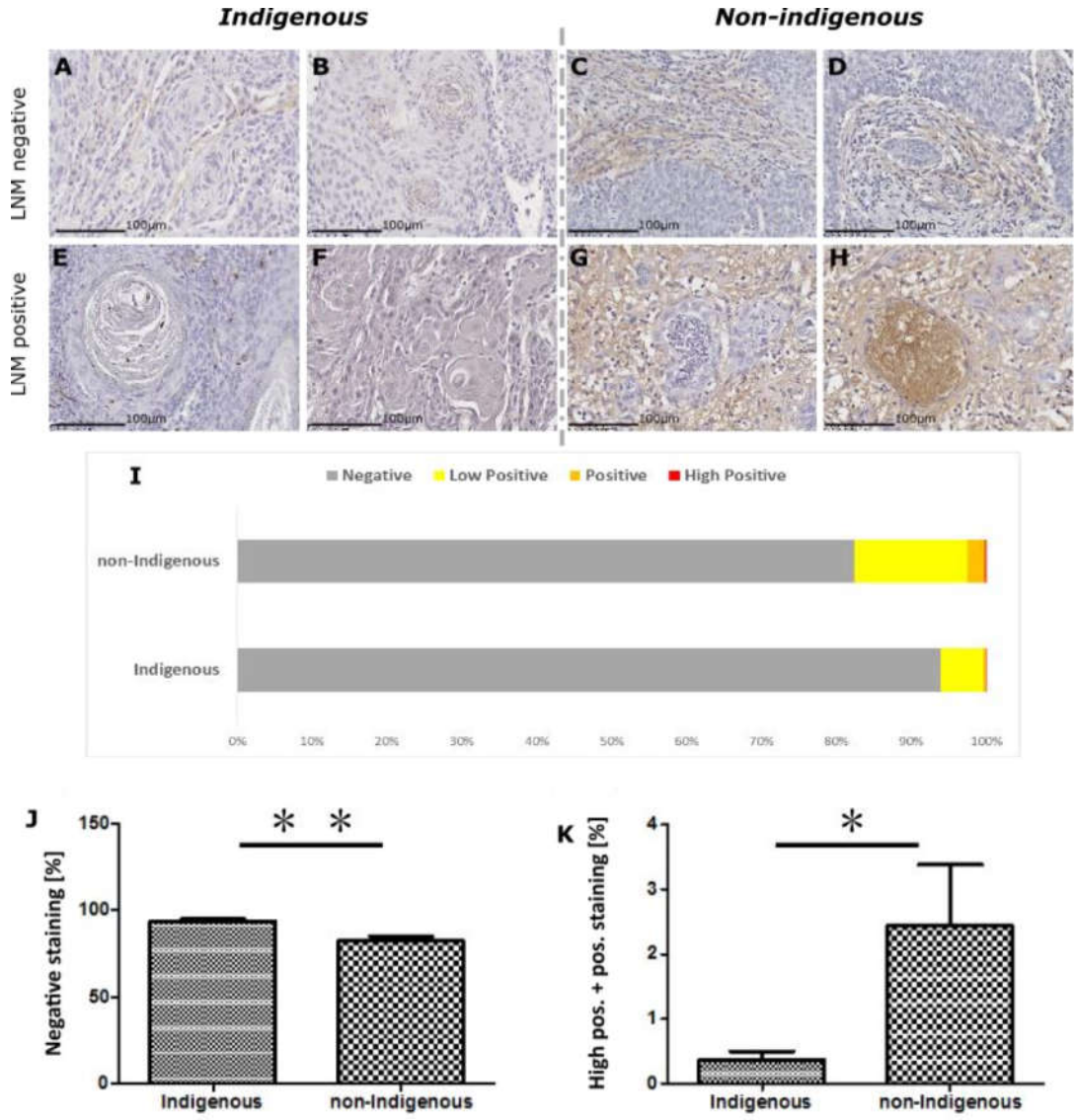


Figure 5.15: IHC analysis of POSTN on VSCC tissue sections containing samples of Indigenous (n=4) and non-Indigenous samples (n=4), at 40× magnification. (A) and (B) POSTN staining of LNM negative Indigenous patient sample as well as (C) and (D) non-Indigenous LNM negative patient sample; (E) and (F) POSTN staining of LNM positive Indigenous patient sample; (G) and (H) POSTN staining of non-Indigenous LNM positive patient sample; (I) Quantitative analysis of POSTN staining across tissue sections was performed using IHC profiler-Image J. For each tissue section, three representative photo-micrographic images at 40× magnification were used and each image was assigned a staining score of negative, low positive, positive or high positive; (J) The per cent of negative staining was significantly different between Indigenous and non-Indigenous VSCC patient samples ($p = 0.0016$); (K) The per cent of positive and high positive staining was significantly different between Indigenous and non-Indigenous VSCC patient samples ($p = 0.0408$).

The differential abundance of NDRG1 between the LNM positive and negative Indigenous VSCC samples, as revealed in LFQ analysis, was further investigated by IHC staining of the Indigenous VSCC tissues analyzed in the LFQ experiments. Stronger positive staining of NDRG1 was mainly observed in the cytoplasm of the VSCC tumours from the LNM positive Indigenous samples as compared to the LNM negative Indigenous samples. IHC analysis of NDRG1 revealed a significant difference ($p = 0.0021$) in the level of total positive staining between LNM positive Indigenous samples (54.62 ± 5.075%) and LNM Indigenous negative samples (14.43 ± 8.313%). Results of IHC analysis and NDRG1 staining from two representative LNM positive and negative Indigenous samples are provided in ***Figure 5.16***.

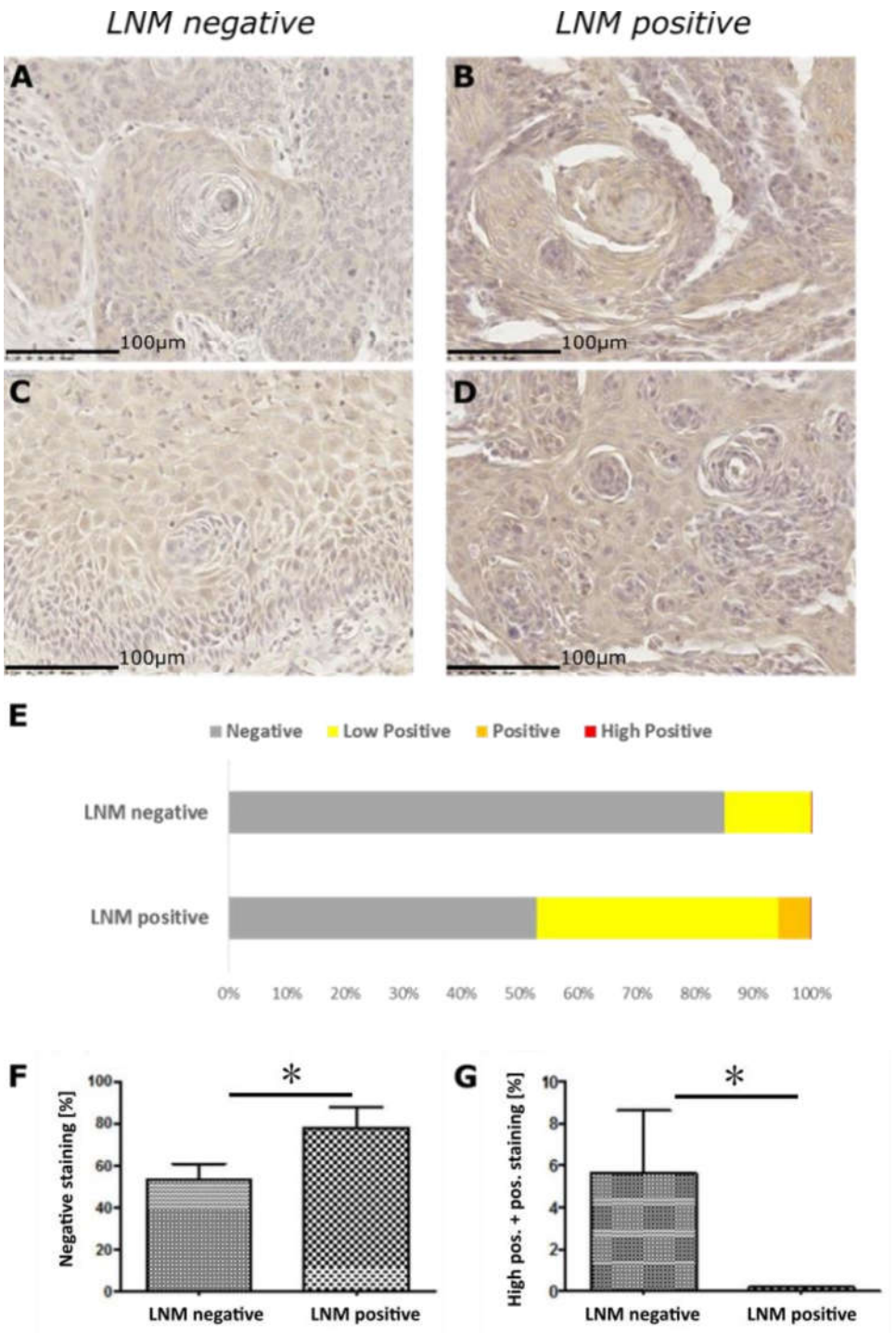


Figure 5.16: IHC analysis of NDRG1 for Indigenous VSCC tissue sections containing samples of LNM positive and negative VSCC samples, at 40× magnification. (A) NDRG1 staining of LNM negative Indigenous sample and (B) LNM positive Indigenous sample; (C) NDRG1 staining of LNM negative Indigenous sample; (D) LNM negative (this should be LNM positive from a different Indigenous patient?) from a different sample; (E) Quantitative analysis of NDRG1 staining across tissue sections was performed using IHC profiler-Image J. For each tissue section, three representative photo-micrographic images at 40× magnification were used and each image was assigned a staining score of negative, low positive, positive and high positive; (F) The per cent of

negative staining was significantly different between LNM positive and negative VSCC samples (p = 0.0361); (G) The per cent of positive and high positive staining was significantly different between LNM positive and negative VSCC samples (p = *0.0478*).

The differential abundance of RAP1A between the LNM positive and negative Indigenous VSCC samples, as revealed in LFQ analysis, was further investigated by IHC staining of the Indigenous VSCC tissues analyzed in the LFQ experiment. Stronger positive staining of RAP1A was observed in the tumour cell cytoplasm of the LNM negative Indigenous samples, while the LNM positive Indigenous samples showed significantly lower protein abundance. IHC analysis of RAP1A revealed a significant difference (p=*0.0046*) in the level of total positive staining between LNM positive Indigenous samples (22.70 ± 3.146%) and LNM Indigenous negative samples (37.24 ± 2.483%). Results for IHC analysis and RAP1A staining from two representative LNM positive and negative Indigenous samples are provided in ***Figure 5.17***.

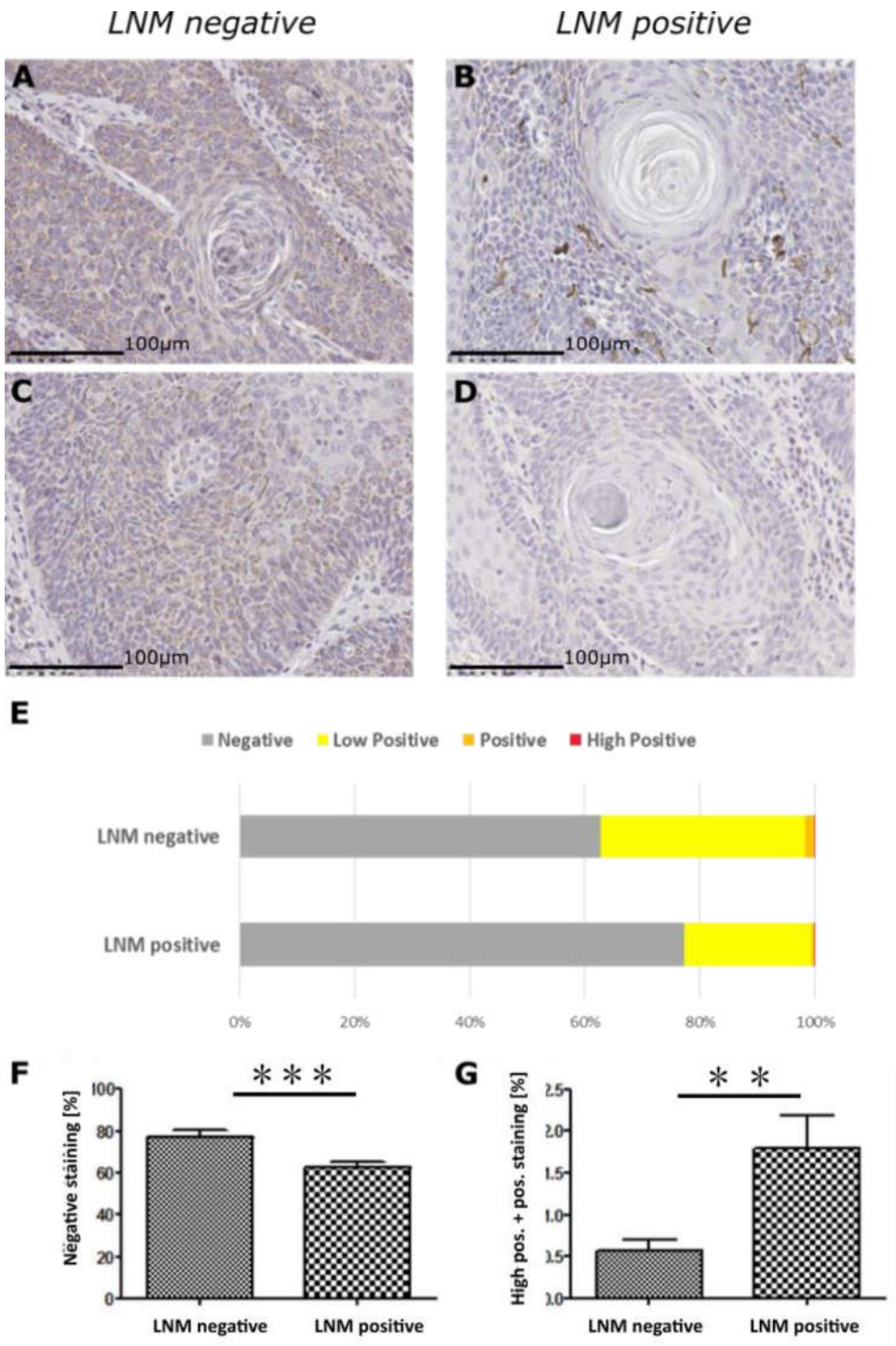


Figure 5.17: IHC analysis of RAP1A on Indigenous VSCC tissue sections containing samples of LNM positive and negative VSCC samples, at 40× magnification. (A) RAP1A staining of LNM negative Indigenous sample and (B) LNM positive Indigenous sample; (C) RAP1A staining of different LNM negative Indigenous sample; (D) LNM negative (that should be positive from a different indigenous patient?) from different sample; (E) Quantitative analysis of RAP1A staining across tissue sections was performed using IHC profiler-Image J. For each tissue section, three representative photo-micrographic images at 40× magnification were used and each image was assigned a staining score of negative, low positive, positive or high positive; (F) The per cent of

negative staining was significantly different between LNM positive and negative VSCC samples (p = 0.0046); (G) The per cent of positive and high positive staining was significantly different between LNM positive and negative VSCC samples (p = *0.0139*).

5.4 Discussion

In this chapter, HPV-independent VSCC tissues collected from Indigenous and non-Indigenous patients were compared by label-free quantification mass spectrometry with the aim of identifying why the two cohorts demonstrate different epidemiological characteristics. Such analyses may highlight molecular pathways that are involved in the earlier diagnostic age and rapid development of vulvar cancer in Indigenous women. Protein variations highlighted in the LFQ analysis were further validated by MALDI-MSI and IHC staining on consecutive tissue sections from the sample cohort. The LFQ comparison of Indigenous and non-Indigenous patients in this study revealed significantly low abundance levels of the extracellular matrix protein Periostin (POSTN) in the Indigenous vulvar cancer patients compared with non-Indigenous patients. In these analyses, the protein N-myc downstream regulated 1 (NDRG1) also displayed higher abundance levels in the LNM positive Indigenous VSCC tissues compared to the negative tissues. Similar results were achieved from non-Indigenous patient TMA analysis described in Chapter 4. The proteome analysis presented in this chapter also revealed proteins that were differentially abundant in the Indigenous patients with and without positive LNM status, that were not reported as differentially abundant in Chapter 4. A lower level of Ras-related protein 1A (RAP1A) was found in the LNM positive Indigenous samples compared to the LNM negative samples, which has not been found in non-Indigenous vulvar cancers yet. Identified proteins of interests that appeared to be specific to the Indigenous VSCC tissues were validated using MALDI-MSI and IHC on consecutive tumour sections.

The differentially abundant proteins of interest between the Indigenous and non-Indigenous VSCC tissues were analyzed using the Reactome Pathway Analysis software. This analysis highlighted signalling pathways involved in neutrophil degranulation and metabolism, specifically involving the proteins POSTN, NDRG1, and RAP1A. Neutrophils are key inflammatory cells and they are the most abundant type of white blood cell, constituting an essential part of a human innate immune system that defends bacterial and viral infections (Murray et al., 2007). Neutrophils are normally recruited during the inflammatory response, producing cytokines and chemokine which may mediate extracellular matrix reconstruction in the local microenvironment by promoting matrix metalloproteinase and keratinocyte growth factor (KGF/FGF-7) (Coussens and Werb, 2002). However, neutrophils and their associated chemokine production have also displayed the potential ability to regulate tumour growth, angiogenesis, and metastasis (Vaguliene et al., 2013). There is an increasing evidence to

suggest neutrophil degranulation triggers the pathogenesis of many severe inflammatory disorders, including rheumatoid arthritis and Behcet's Disease, which causes damage to nearby tissues (Wright et al., 2010). Currently, the neutrophil-lymphocyte ratio has been used in gynecological cancer patients as a prognostic reference for treatment outcomes while recruitment of neutrophils in the extracellular matrix has been frequently reported in cancer-induced inflammation which could increase the possibility of tumour angiogenesis, immunosuppression, and metastasis (Guthrie et al., 2013). Neutrophils have been found to be recruited by chemo-attractants released in the tumour microenvironment and triggered phosphoinositide 3-kinase (PI3K) pathways that boost tumour proliferation and angiogenesis in lung cancer (Houghton et al., 2010). On the other hand, TGF-β and other cytokines that are produced by neutrophils could induce the E-cadherin degradation and further promote epithelial-to-mesenchymal transition (EMT) in pancreatic and ovarian cancer cell lines (Große-Steffen et al., 2012, Mayer et al., 2016, Klebanoff et al., 1986). Previous western blot analysis has also shown the high extracellular concentration of the inflammatory protein, S100A9, is able to induce neutrophil degranulation by activating MAPK-dependent pathways after chronic exposure to high inflammatory levels, which could further increase the risk of carcinogenesis (Simard et al., 2010, Zhu et al., 2013). In the LFQ comparison of LNM negative vulvar cancer patients, 12 proteins of interest were also found to be related to neutrophil degranulation, including S100A9 which showed higher LFQ intensity levels in the two Indigenous LNM negative patients as compared to be the non-Indigenous patients. These results suggest protein alternations related to neutrophil signalling pathways might facilitate the rapid development of vulvar cancer in Indigenous patients.

Obesity and diabetes have been reported as major health problems amongst Australian Indigenous women and recent statistical studies indicated that obesity seems to correlate with the high incidence rates of vulvar cancer (Zhu et al., 2013). Despite the tumorigenesis effect of obesity in Indigenous patients being unclear, the analysis presented in this chapter found the insulin metabolism pathway to be one of the pathways highlighted in the Reactome analysis. The Insulin/Insulin-like Growth Factors pathway has been reported to enhance carcinogenesis of many cancers as it is also able to control apoptosis by regulating Ras/MAPK signalling pathways (Giovannucci, 2003, Perrin et al., 2013, Longo, 2004). To date, the role of the Insulin signalling pathway in vulvar cancer development has not been fully studied, but recent work has shown that levels of insulin receptors are closely related with PTEN tumour suppressor gene expression and PI3K/Akt pathways activation in gynecological cancers (Mehrian-Shai et

al., 2007, Gil et al., 1999). In summary, the highlighted protein abundance variations associated with the insulin signalling pathways indicate a potential relationship between the unique metabolic characteristics of Indigenous VSCC and their rapid development.

The extracellular matrix ligand protein POSTN was found to be significantly lower in abundance in the Indigenous samples as compared to the non-Indigenous samples. Currently, the function of POSTN during vulvar cancer development is still poorly understood. POSTN has been shown to work with missense p53 mutation and mediated STAT1 induced esophageal squamous cell carcinoma invasion (Wong et al., 2013). POSTN is also believed as to be an extracellular matrix component that directly interacts with cytomembrane integrins as chemo-attractors in chronic inflammation (Liu et al., 2014). As the key extracellular matrix protein in the tumour microenvironment, POSTN is able to bind integrins by Akt/PKB- and FAK signalling pathways and mediate EMT changes by altering cell-matrix adhesion in cancer metastasis (Morra and Moch, 2011). Interestingly, another FAK signalling regulator, MBL, also demonstrates unique genetic mutations among Indigenous Australian (Dommett et al., 2006). Inhibiting the POSTN gene expression in a xenograft model results in a dramatic increase in metastatic cancer development, whereas overexpression of POSTN showed increased cell adhesion and inhibition of cell migration (Kyutoku et al., 2011). Although previous research shows POSTN might facilitate tumour metastasis, there is increasing evidence to suggest POSTN could be an important metastatic suppressor in many types of cancer. For example, POSTN expression was reported as having a positive relationship with E-cadherin expression, which enhances cell-cell adhesion and Akt phosphorylation suppression in metastatic cancer (Kim et al., 2011). The down-regulation of POSTN expression via RNA inference showed a remarkable relationship with higher grade bladder cancer and overexpression of POSTN could suppress the metastasis of invasive bladder cell lines in an implanted animal model (Kim et al., 2005). Other research has shown that POSTN interacts with the TAB1/TAK1 signalling pathway which may suppress tumour cell invasiveness and metastasis (Isono et al., 2009). In the IHC analysis of vulvar cancer tissues, POSTN was mainly detected in the stromal cells of the non-Indigenous VSCCs, but at overall lower levels than in the non-Indigenous VSCC tissues.

One of the HPV-independent vulvar cancer proteins, NDRG1, found to be differentially abundant in the analysis presented in Chapter4 (of primary VSCC tissues with and without LNM) was also similarly found to be significantly more abundant in the Indigenous VSCC

tissues with LNM compared to those without LNM. This suggests NDRG1 abundance is directly related to vulvar cancer metastasis, and not the ethnicity of patients. Conversely, the protein RAP1A was around 4-fold lower in abundance in the two LNM positive Indigenous samples as compared to LNM negative samples. The significantly lower amount of RAP1A in LNM positive Indigenous tissues was corroborated by the MALDI-MSI and IHC staining ($p<0.05$) results. Expression of Rap1 may induce production of many important cytokines by activating nuclear factor kappa B (NF-κB) signalling in macrophages cells, whereas a Rap1 knockdown model showed significant suppression of NF-κB activity and decreased the production of pro-inflammatory cytokines in atherosclerosis lesions (Cai et al., 2015). The decrease of Rap1 activity also showed negative effects on RAPL expression which may lead to lymphoid chemotaxis inhibition and further decreases integrin-mediated tumour cell adhesion (Tanaka et al., 2005). Moreover, siRNA experiments show that Rap1 dependent PI3Kγmediated integrin activation plays a key role in regulating tumour-induced inflammation in myeloid cells (Schmid and Varner, 2012). In vulvar cancer samples, a significantly low abundance of RAP1A expression was detected in the LNM positive samples and a decrease of RAP1A expression may facilitate Indigenous vulvar cancer distance metastasis by circumventing the immune system.

Due to limited access to Indigenous vulvar cancer samples, this study was unable to determine whether RAP1A, NDRG1 and POSTN have potential use in vulvar cancer LNM prediction for Indigenous women. However, differently abundant proteins highlighted in the LFQ results revealed important biological processes through Reactome analysis that are likely involved in the development and metastasis of vulvar cancer. Most of the proteins found to be differentially abundant between the Indigenous and non-Indigenous samples were involved in innate immune reactions, especially the process of neutrophil degradation. Identification of protein involved in innate immunity proteins provide evidence for a connection between inflammation and Indigenous vulvar cancer, as carcinogenesis of HPV-independent vulvar cancer is understood to originate from chronic inflammation. Previous research found that the inflammatory condition was widely present in tissues with oncogenic change and this may induce inflammation in a hypoxic microenvironment that further facilitates tumour development and metastasis (Ai et al., 2016, Balkwill, 2012). Further research on innate immunity regulation found many oncogenetic mutations, including the RAS and MYC genes, in downstream inflammatory signalling which may favour tumour development by inducing inflammatory chemokine production and extracellular matrix remodelling (Coussens and Werb, 2002).

Interestingly, RAP1A showed high abundance in LNM negative Indigenous vulvar cancer samples and NDRG1 was related to vulvar cancer LNM. Expression of POSTN is involved in many inflammatory diseases, including those of the airways, skin, atherosclerosis and fibrosis inflammation, and also induces tissue microenvironment changes during chronic inflammation (Liu et al., 2014). Therefore, the rapid development of vulvar cancer in young Indigenous patients may be influenced by inflammatory conditions.

The major problem of current Australian Indigenous cancer studies is that most biomarker candidates simply revealed by a very small group of patients compared with similar studies with hundreds of tissue or serum samples collected from non-Indigenous patients (Diaz et al., 2015). Although the abnormal inflammatory condition has been highlighted as an important risk factor in many other Indigenous cancer studies, there is no direct evidence showed that the inflammatory associated proteins are overexpressed in Indigenous vulvar cancer patients (Cox et al., 2014a). Thus, it is hard to know the abundance differences of identified proteins in other Indigenous vulvar cancer patients without the support of further proteomics or genomics characterization results.

However, considering the numbers of accessible Indigenous patient samples in future studies is very low, more comprehensive proteome studies of identified proteins of interest with their regulators on available Indigenous VSCCs are needed. Currently, a major challenge of applying IHC in Indigenous vulvar cancer protein validations is the lack of high efficiency and Indigenous-specific antibodies in clinical studies. The IHC has been considered as the golden standard in protein biomarker validations, but it mainly reveals the spatial distribution of target biomarker rather than accurate quantitative data without external references, so the diagnostic value of morphology-based IHC assays is hardly achieved (Sullivan and Chung, 2008). Furthermore, the storage of cancer tissues after surgery, samples fixation and preparation, the efficiency of protein epitope retrieval, the choice of monoclonal or polyclonal antibodies, the digital detection system, could also affect the robustness of IHC based protein validation results. In order to validate abundance differences of proteins of interests on limited Indigenous VSCC samples, the MRM (Multiple Reaction Monitoring) MS will be a promising biomarker validation method with its high sensitivity and analytical throughputs on Indigenous tissue samples than IHC validation (Issaq and Veenstra, 2008). Compare with IHC validation methods, like enzyme-linked immunosorbent assays (ELISA) microarray, MRM-MS is less expensive and time-consuming technique for multiple protein validations. Due to the

improvement of analytical speed and mass resolution of MS techniques, nearly a hundred protein candidates could be specifically quantified from limited numbers of samples like Indigenous VSCCs. For example, a recent study shows more than a hundred high or moderate abundance proteins are able to be quantified in a single MS run of the undepleted serum sample (Percy et al., 2014). Besides, the usage of MRM-MS could comprehend the problems of no available antibody or immune assay poorly performed in the Indigenous vulvar cancer research. Further MRM-MS analysis on these valuable Indigenous VSCC tissue samples will not only help to quantify identified proteins of interest in the LFQ experiments but also a complementary or alternative technique to IHC for vulvar cancer biomarker evaluation.

5.5 Conclusion

In summary, this is the first proteomic study of Indigenous vulvar cancer patient samples. Compared with non-Indigenous samples, the four representative Indigenous VSCC samples, i.e. two LNM negative and two LNM positive samples, showed younger diagnostic age and different tissue morphological patterns. To comprehensively study these Indigenous cases, *nano*-LC-MS/MS-based proteomic profiling and label-free quantitation (LFQ) analysis were performed to explore similarities and differences between Indigenous and non-Indigenous patient samples. Results were further validated using MALDI-MSI and immunohistochemistry (IHC). Here, a significantly lower level of the extracellular matrix protein, POSTN, was detected from four Indigenous vulvar cancer patient samples compared with its expression level in LNM positive and negative non-Indigenous patients. Interestingly, NDRG1, previously identified as a potential LNM predictor from Chapter 4, also showed higher abundance levels in two LNM positive Indigenous samples compared to two LNM negative Indigenous samples. Additionally, a significantly lower abundance of RAP1A in the LNM positive Indigenous samples compared to LNM negative samples was observed. Reactome analysis of the proteins found to be differentially abundant revealed a high proportion of them to be part of the inflammatory process of neutrophil degradation, highlighting a potential relationship between inflammation and Indigenous vulvar cancer.

This page intentionally left blank

CHAPTER 6: DISCUSSION AND FUTURE DIRECTIONS

6.1 Main findings

In this thesis, the proteomic fingerprints of HPV-independent vulvar cancer across different stages of progression and metastasis has been characterized, and the differential protein abundances between Indigenous and non-Indigenous VSCCs has also been analyzed. In the published manuscript referred to in Chapter 3, MALDI-MSI revealed the cytoskeleton component, cytokeratin 5, to be a marker of differentiated vulvar intraepithelial neoplasia (dVIN) as the protein demonstrated to be significantly decreased in abundance in adjacent VSCC tissues and benign vulvar skin. Following analyses showed tryptic peptides from β-actin and vimentin to be significantly increased in intensity in the VSCC tissues. The increased level of β-actin in the metastatic LNs and VSCC was further validated by IHC staining and DIA based LFQ analysis on consecutive tissue sections.

In Chapter 4, two TMAs constituting 16 LNM positive and 12 negative HPV-independent primary VSCC tissues were used to establish a MALDI-MSI based classification assay for lymph node metastasis based on primary cancer. Here, a CCA-LDA classification could achieve an LNM classification accuracy of 100% when analyzing the 28 primary vulvar cancer tissues in the TMA. The development of CCA-LDA provides a new solution in vulvar cancer patient classification using MALDI-MSI results. However, this best result only achieved when the log-transformed data and leave-one-out validation are used in this TMA classification, this result cannot be used to represent its real performance on other vulvar cancer patient classification. Top-ranked peptides used for the CCA-LDA classification were identified as from the proteins, NDRG1, Annexin A8, vimentin, collagen type 1 and cytokeratin 6, which have shown fair AUC scores in the ROC analysis of TMA imaging results. The identified protein abundance changes and spatial distribution were validated by IHC staining on consecutive TMA sections. Notably, protein levels of NDRG1 and Annexin A8 were also found to be significantly increased in the six LNM positive primary cancers compared with six LNM negative cancers in LFQ MS analysis with abundance changes of 6.88 and 8.11 fold respectively.

In the last chapter (Chapter 5), although only four representative Indigenous samples were accessible at the time of this study, an increased abundance of NDRG1 was also observed in the LNM positive Indigenous VSCC tissues as compared to the LNM negative Indigenous VSCC tissues. Additionally, the level of RAP1A showed more than a 4 fold change between LNM positive and negative Indigenous VSCC tissues, a finding which has not been previously reported. Importantly, compared with non-Indigenous patients, significantly increased levels of POSTN were observed. Finally, biological processes related to the proteins found to be deferentially abundant between the Non-Indigenous and Indigenous VSCC tissues were highlighted in a Reactome pathway analysis, which may shed light on the high levels of vulvar cancer prevalence and younger diagnostic ages of Indigenous women with VSCC as discussed below.

6.2 Application of MALDI-MSI on vulvar cancer research

To my knowledge, this is the first time MALDI-MSI has been applied to research vulvar cancer. This is also the first time that *nano*-LC-ESI-MS/MS based protein LFQ and MALDI-MSI have been used to analyses VSCC tissues from Indigenous Australians. Currently, the genomic and proteomic markers have been widely considered in patient classifications as the molecular evidence complement to the histopathological analysis. Unfortunately, no DNA or RNA sequencing profile has been established specifically for the HPV-independent vulvar cancer until now, whilst available protein markers are still unable to discriminate primary tumour samples from different pathological stages (Sandberg et al., 2012, Knopp et al., 2004). In this thesis, the proteomic patterns detected in the MALDI-MSI showed a high accuracy in classifying VSCCs with and without LNM using the CCA-LDA method. Although the performance of tumour classification using MALDI-MSI data still challenged by many experimental variations and requires further inspections on a larger patient cohort, it has demonstrated a much better LNM predication outcome comparing with previous results achieved by single protein biomarkers and histological parameters (Knopp et al., 2006, te Grootenhuis et al.). Furthermore, another promising advantage of using MALDI-MSI data in vulvar cancer classification is combining proteomic patterns with histopathological information into the prediction model. A recent study on endometrial cancer demonstrated a significant improvement in LNM prediction outcomes after incorporated ACTN4 peptide intensities with lymphovascular space invasion observations in one model (Mittal et al., 2017). It is believed the prediction ability of those identified proteomics patterns can be further improved after combine some necessary VSCC histopathological information, such as tumour grade and lymphovascular invasion, in the LNM prediction model.

6.3 Cytoskeleton remodeling as an important biological process in vulvar cancer progression

In this thesis, a majority of the proteins highlighted in the MS and IHC analyses were found to be cytoskeletal components which have been shown to facilitate tumour cell proliferation and distant metastasis. Alternation, in the cytoskeletal network has been increasingly accepted as a key hallmarks in the development of ovarian, breast, endometrial and cervical cancer metastasis, and they have also been identified as development hallmarks of virus-induced vulvar carcinoma (Naora and Montell, 2005, Wang et al., 2007, Lee and Shen, 2012, Pappa et al., 2011). Although the genetic mutations that trigger actin-remodelling in vulvar keratinocytes are still unknown to us, a recent study suggested that the tumour actin filament alternation is closely related to the epithelial-mesenchymal transition (EMT) change during vulvar cancer metastasis (Wicki et al., 2006, Zannoni et al., 2011). Here, β-actin was found to have high expression levels in the primary and secondary tumours of the late stage vulvar cancer patients. Importantly, other VSCC specific proteins highlighted in this study are all involved with the construction and regulation of actin-filaments (***Figure 6.1***). The protein myosin 9 has been reported to play as a vital role in actin-filament network remodelling, and as important interactors in estrogen receptors in metastatic bone and breast cancers (Zhou et al., 2016, Stellato et al., 2015). Myosin 9 was also detected in the proteomics analysis of VSCC samples even though its protein abundance difference between LNM positive and negative cancers was not as significant as other candidates. The further study of the myosin-actin interaction in VSCCs could deliver a better understanding of the molecular mechanism of vulvar cancer metastasis.

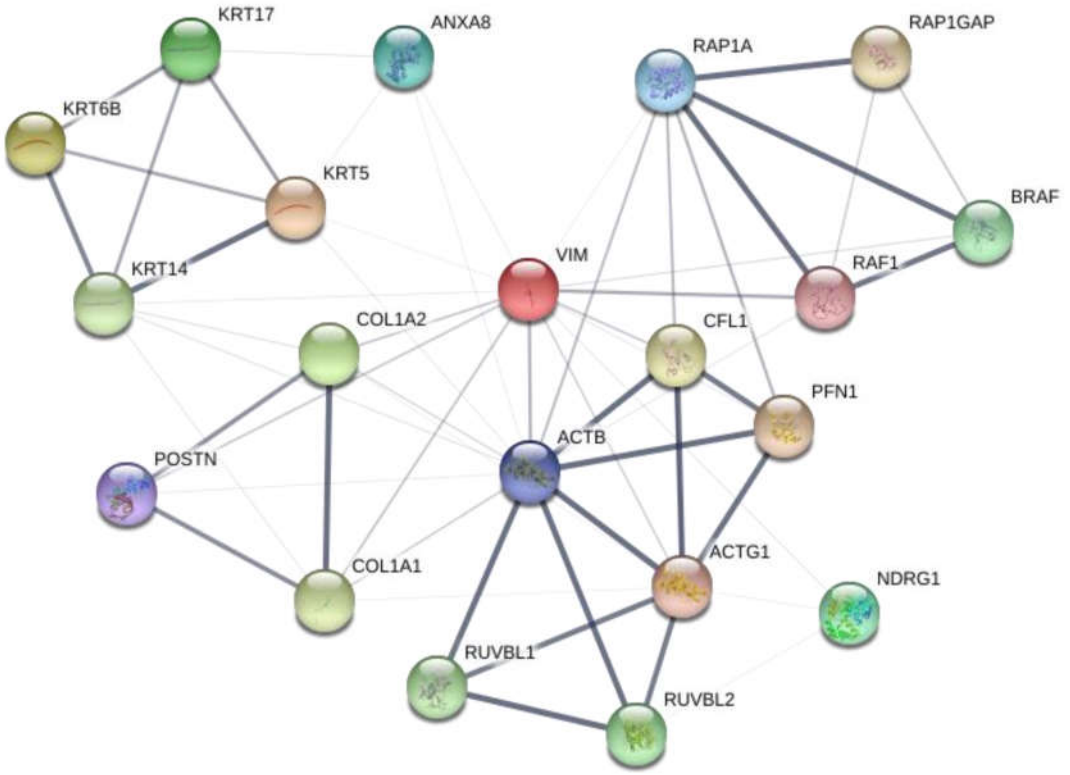


Figure 6.1 Interactions among proteins found to be potentially involved in vulvar cancer progression. Interaction map was created by searching IHC validated protein of interests in this thesis with supporting data in the STRING database (v10.5); line thickness indicates the strength (number of published papers) of data support.

Previous proteomic studies have also shown that loss of E-cadherin and increasing vimentin levels could significantly reduce tumour cell-cell adhesion and facilitate tumour mobility in vulvar cancer (Rodrigues et al., 2013a). As another cytoskeletal protein that is potentially related to actin filament regulation, vimentin has been widely accepted as a biomarker of epithelial-to-mesenchymal transition (EMT) and it frequently detected in the invasive tumour front of vulvar cancer tissues with increased abundance related to poor prognostic outcomes of patients (Rodrigues et al., 2013a). To explore vimentin expression in HPV-independent vulvar cancer patients, the spatial distribution of vimentin was validated using IHC stains, but, the highest vimentin level is found in the stromal area of the LNM negative VSCCs rather than LNM positive ones. Subsequent microscopic inspection indicated vimentin was not only increased in the metastatic VSCCs, but was also overexpressed in leucocytes, such as neutrophils and lymphocytes, which had migrated into the tumour microenvironment.

Tumour-specific proteins, cytokeratin 5 and 6, are highly abundant intermediate components that are vastly expressed in skin epithelial cells. They maintain a series of key functions of keratinocytes, including cell proliferation, epidermis development, keratin repair, and epithelium morphogenesis. Loss of cytokeratin 5 or 6 within epithelial cells results in severe skin disorders (Chu and Weiss, 2002b, Vasca et al., 2014). Previously, cytokeratin 5 and 14 have been detected in LC-MS experiments and were found to be unique to vulvar amyloidosis lesions as keratin-related amyloid materials (Quddus et al., 2014). As a marker of epithelial growth cytokeratin 5 and 6 have displayed overexpression in benign epithelial neoplasms of the breast, cervical, and uterus cancer (Chu and Weiss, 2002a). A previous study found decreasing cytokeratin 5 in VSCC from precursor differentiated VIN to VSCC when the malignant tumour lost normal epithelium morphological characteristics (Zhang et al., 2016a). Further lower abundance levels of cytokeratin 6 were revealed in LNM positive vulvar cancer patients, which displayed about two-fold of differences in IHC analysis between LNM positive and LNM negative samples. In addition, cytokeratin 17 was reported as a basal cell differentiation regulator in epithelium through binding 14-3-3σ and activating Akt/mTOR pathway in HPV-dependent tumorigenesis (Hobbs et al., 2016). As a promoter of cell proliferation, cytokeratin 17 has shown the clinically relevant with cancers invasion, metastasis, and drug resistance (Goyal et al., 2016). Previous IHC experiments showed increasing of cytokeratin 17 expression was associated with lymph node metastasis in thyroid carcinoma (Kim et al., 2015). Abnormal expression of cytokeratin 17 and 5/6 already identified in RNA sequencing results of breast cancer patients as poor clinical outcome markers (Dai et al., 2014). Although there is no direct evidence shows ANXA8 could regulate keratin expressions, a recent study suggested that ANXA8 is up-regulated by p63 and further inhibiting cytokeratin 5 expression in many epithelial cancers (Trink et al., 2007). Besides, the qRT-PCR analysis of oral squamous cell carcinoma samples also found that cytokeratin 6 and ANXA8 might be potential biomarkers of metastatic patients (Oka et al., 2016).

The extracellular matrix protein, collagen type 1 is widely expressed in mammal tissues and is important in many biological processes involved in ordinary tissue development, including embryonic skeletal development and negative regulation of cell-substrate adhesion (Salasznyk et al., 2004). The function of collagen type 1 during vulvar cancer progression has not been fully studied. Previous studies have suggested that collagen type I could accelerate the downregulation of E-cadherin by inducing its negative regulators in a pancreatic cancer cell line (Hosper et al., 2013). Moreover, some research found that type I collagen-rich tissues

showed stabilization of the EMT regulator, the Snail protein, which releases tumour cells from cell cycle capture (Medici and Nawshad, 2010). Periostin has been recognized as a direct upstream regulator of collagen proteins during fibrillogenesis and tumorigenesis (Norris et al., 2007, Ruan et al., 2009). In the LFQ analysis presented in Chapter 5, periostin showed significantly lower expression levels in Indigenous vulvar cancer patients from both early and late stages. As the key extracellular matrix protein in the tumour microenvironment, periostin is able to bind integrins in the Akt/PKB- and FAK signalling pathways and mediate EMT changes by altering cell-matrix adhesion in cancer metastasis (Morra and Moch, 2011). The expression of periostin in stroma is believed to create a tumour-supportive microenvironment and is able to stabilize p53 and E-cadherin proteins (Lv et al., 2014). Lower periostin abundance levels could be one of the key causes of the rapid development of vulvar cancer in young Indigenous women.

6.4 Tumor hypoxia proteins in vulvar cancer metastasis

Apart from proteins revealed in the MALDI-MSI, there were many proteins highlighted in LFQ analysis of VSCC patient samples which could be of interest in future studies. For example, the 40S ribosome protein S5 is an important functional protein which maintains regular cell growth and proliferation, whilst the impairment of 40S ribosome expression could active the tumorigenesis via p53-dependent nucleolar stress (de las Heras-Rubio et al., 2014, Nicolas et al., 2016). In Chapter 4, the 40S ribosome protein S5 displayed a more than 5 fold decrease in the six LNM positive vulvar cancers compared with the rest LNM negative cancers. Phosphoglycerate kinase 1 (PGK1) is a critical enzyme involved in glycolysis that synthesizes pyruvate from D-glyceraldehyde 3-phosphate. It was recently detected in many malignant tumours and was strongly associated with tumour hypoxia that results in cancer metastasis (Semenza, 2003, Winter et al., 2007, Chi et al., 2006). Real-time qPCR and western blotting of PGK1 expression in 46 breast cancer reveal that PGK1 is significantly increased in malignant tumours, especially in LNM positive patients (Sun et al., 2015). Higher PGK1 expression may result from aerobic microenvironment in VSCC and overexpression of PGK1 may induce chemo-resistance and anti-apoptotic effects of malignant tumour cells (Weber, 2016). In the vulvar cancer cohort, a higher PGK1 abundance was found in the tumour regions, especially in the LNM positive tissues whose PGK1 abundance were 2-fold higher than LNM negative tissues. Laminins are also known as EMT involved proteins and increasing laminins in malignant VSCCs could facilitate tumour invasion by reducing cell-cell attachment, gaining migration abilities in the extracellular matrix during cancer progression (Oliveras-Ferraros et al., 2014). Detection of LAMB3 expression in the LNM positive VSCCs may provide evidence of EMT changes in vulvar cancer metastasis. Genetic and IHC studies have indicated that LAMB3 with ITGB1 can be used as pro-metastatic markers for lung cancer (Wang et al., 2013). Another metabolism protein revealed to have higher abundance in the LNM positive VSCCs is the Galectin-3-binding protein (LG3BP). In distant metastasis, LG3BP may mediate galectin dependent tumour cell aggregation through mediating cell-extracellular glycoconjagulated and actin skeleton reorganizing with focal adhesion kinase (Lin et al., 2015, Funasaka et al., 2014). The expression of LG3BP has been used in a diagnostic panel with other proteins in lung cancer (Kearney et al., 2015). In Chapter 4, the LG3DP protein showed 5 times higher LFQ intensity in the LNM positive samples compared with the other samples. On the contrary, lower abundance of the metabolism protein sodium/potassium-transporting ATPase subunit alpha

(AT1A) in malignant tumours compared to benign lesions has been found in many different cancer types, other than vulvar cancer. For example, the amount of AT1A mRNA is 1.5 to 2 times greater in benign lesions in breast cancer (Bogdanov et al., 2017). Lower levels of the protein AT1A in tumour regions has been found to be associated with tumour hypoxia and nutrient deprivation (Izuishi et al., 2000). Interestingly, the vulvar cancer LNM hallmark protein, NDRG1, leads to remarkably decreased levels of phosphorylated Akt1in hypoxic tumours, which could further trigger actin-remodelling in pancreatic cancer cell lines (Liu et al., 2017a). These proteins could be important in vulvar cancer progression as they have not only shown significant abundance differences in the LFQ analyses but are involved in tumour hypoxia, which may trigger the malignant growth of cervical cancer (Höckel et al., 1996).

6.5 Neutrophil degranulation promotes Indigenous vulvar cancer progression

Apart from NDRG1, Ras-related protein rap1α (RAP1A) is another protein able to modulate AKT1 protein levels promoting tumour formations (Lou et al., 2002). Significantly lower RAP1A levels were observed in the LNM positive Indigenous vulvar cancer tissues compared with the LNM negative Indigenous cancer tissues here, whilst further Reactome pathway analysis suggested the RAP1A expression changes are closely related to neutrophil degranulation. The neutrophil has been recognized as a key inflammatory cell type and is also the most abundant type of white blood cell, which constitutes an essential part of the innate immune system that defends bacterial infections (Murray et al., 2007). Neutrophils are normally recruited in the inflammatory response, producing cytokines and chemokines which may mediate extracellular matrix reconstruction in the local microenvironment by promoting matrix metalloproteinases and keratinocyte growth factor (KGF/FGF-7) (Coussens and Werb, 2002). However, neutrophils and the associated chemokine production have shown the potential ability to regulate tumour growth, angiogenesis, and metastasis (Vaguliene et al., 2013). There is an increasing evidence to suggest neutrophil degranulation results in the pathogenesis of many severe inflammatory disorders, including rheumatoid arthritis and Behcet's Disease, and cause damage to nearby tissues (Wright et al., 2010). Currently, the neutrophil-lymphocyte ratio has been used in gynecological cancer patients as a prognostic reference for treatment outcomes while recruitment of neutrophils in the extracellular matrix has been frequently reported in cancer-induced inflammation which could increase the possibility of tumour angiogenesis, immunosuppression, and metastasis (Guthrie et al., 2013). Neutrophils have been recruited by chemo-attractants released in the tumour microenvironment and triggered phosphoinositide 3-kinase (PI3K) pathways that boost tumour proliferation and angiogenesis in lung cancer (Houghton et al., 2010). On the other hand, TGF-β and other cytokines that are produced by neutrophils could induce the E-cadherin degradation and further promote EMT in pancreatic and ovarian cancer cell lines (Große-Steffen et al., 2012, Mayer et al., 2016, Klebanoff et al., 1986). Previous western blot analysis has shown the high extracellular concentration of the inflammatory protein, S100A9, is able to induce neutrophil degranulation by activating MAPK-dependent pathways after chronic exposure to inflammation, which could further increase the risk of carcinogenesis (Simard et al., 2010, Zhu et al., 2013). In the LFQ comparison of LNM negative vulvar cancer patients, 14 protein variations were also found to be related to neutrophil degranulation, including S100A9.

6.6 Clinical considerations and future directions

Vulvar cancer is a heavy burden for Australian women, especially Indigenous Australian women from remote communities. There will be approximately 6020 Australian women diagnosed with vulvar cancer in 2017 and 1150 patients will die from the disease (Board, 2016). Although the new FIGO staging system has been modified with tissue morphological patterns of HPV-independent vulvar cancer and its precursor lesions, many vulvar cancer patients are still suffering misdiagnosis and overtreatments (Ordi et al., 2009, Van der Steen et al., 2010). For example, precancerous LS and dVIN lesions shared many tissue morphological patterns similar to the benign lesions which make it hard to be identified from biopsy samples (Ordi et al., 2009). Unlike the well-characterized stages of HPV induced vulvar cancer development, proteomic features of HPV-independent VSCC and dVIN, are still unclear and largely unstudied. Previously, pathologists have been trying to build a quantitative relationship between histopathological observations, such as tumour grades or sizes, and VSCC development stages (Tantipalakorn et al., 2009, Benedet et al., 2000, Courtney-Brooks et al., 2010). Unfortunately, prediction models using histological parameters have shown very low accuracy in stage II vulvar cancer patients as histological parameters used in FIGO system are highly descriptive and irrelevant to molecular pathology (Baiocchi et al., 2013). Evidence from clinical imaging, such as MRI or CT scan, have been considered as complementary parameters to tissue biopsy results, but they are not sensitive enough, challenged by higher false positive rates, and not able to determine distant metastases (Andersen et al., 2015, Thaker et al., 2015). It has been suggested that molecular evidence improve the classification vulvar cancer patients. To achieve the aim to deliver a more accurate patient classification system and a better understanding of HPV-independent vulvar cancer pathology, molecular biomarkers with high discriminative abilities and clinical relevance to cancer development stages need to be introduced into vulvar cancer management. For example, overexpression of p16 has been identified as a key marker for identifying HPV-dependent VSCCs, while Ki-67 is up-regulated in vulvar cancers with distance metastasis (Rufforny et al., 2005, Modesitt et al., 2000a). The clinical relevance of other tumour proliferation genes, such as p53 and MIB1, have also been used in vulvar cancer analysis (Hoevenaars et al., 2008). Previous proteomics based research on vulvar cancer have had several limitations: firstly, current vulvar cancer studies are based on HPV-dependent samples or combined both cancer etiological types together rather than specifically targeted high metastatic risk HPV-independent type, secondly, as no genomic

database of vulvar cancer has been established, most biomarkers tested on vulvar cancer patients are identified from other gynecological cancer research, lastly, available genetic biomarkers of vulvar cancer progression, p14 as an example, are more closely related to the cell cycle regulation rather than the unique tumour morphology and metabolism changes presenting in the cancer metastasis (Lavorato-Rocha et al., 2013). Therefore, the identification of proteins associated with primary vulvar cancer tissues from different developmental stages will not only complement current knowledge of HPV-independent vulvar cancer, but also provide knowledge that may aid in vulvar cancer diagnosis and classification.

In this thesis, proteomics signatures of HPV-independent VSCCs from both Indigenous and non-Indigenous patients were highlighted in this MS based proteomics study. However, a lot of works is still required before the findings presented here can be applied to vulvar cancer diagnosis and classification. Future directions include exploring the diagnostic potential and clinical relevance of the highlighted proteins. Fundamentally and importantly, a much larger sample cohort containing VSCC patients with different genomic and histopathological backgrounds is necessary to reliably verify the results presented here. Although protein predictors revealed in this MS based HPV-independent vulvar cancer proteomics characterization could be potential biomarkers in future vulvar cancer early diagnosis and prognostic predictions, their abundance differences were only confirmed by a very small percentage of Australian patients, which leads to a very high specificity for patients selected in this cohort (Board, 2016). Besides, protein candidates described in this study have neither been validated in genetic-based multivariate analyses nor validated prediction outcomes in large clinical cohorts of Indigenous or non-Indigenous patients, which might reject the correlation between protein predictors and vulvar cancer developmental stages (Trietsch, 2017). For example, in the global validation of ovarian cancer biomarker validations on 1287 patients, only few biomarker genes published in the previous analysis are actually capable of predicting outcome in the large sample set (Győrffy et al., 2012). Therefore, aiming to eliminate the biased result caused by the inherent variability across Australia women and the molecular heterogeneity between individual tumours, ideally more than a thousand vulvar cancer patients, including all four stage patients, should be recruited for next step experiments for a more robust biomarker validation (Ludwig and Weinstein, 2005). However, because of the rareness of vulvar cancer tissue samples, the initial biomarker validation steps will be performed on a few hundred patients which will be grouped according to the following guidelines (***Figure 6.2***).

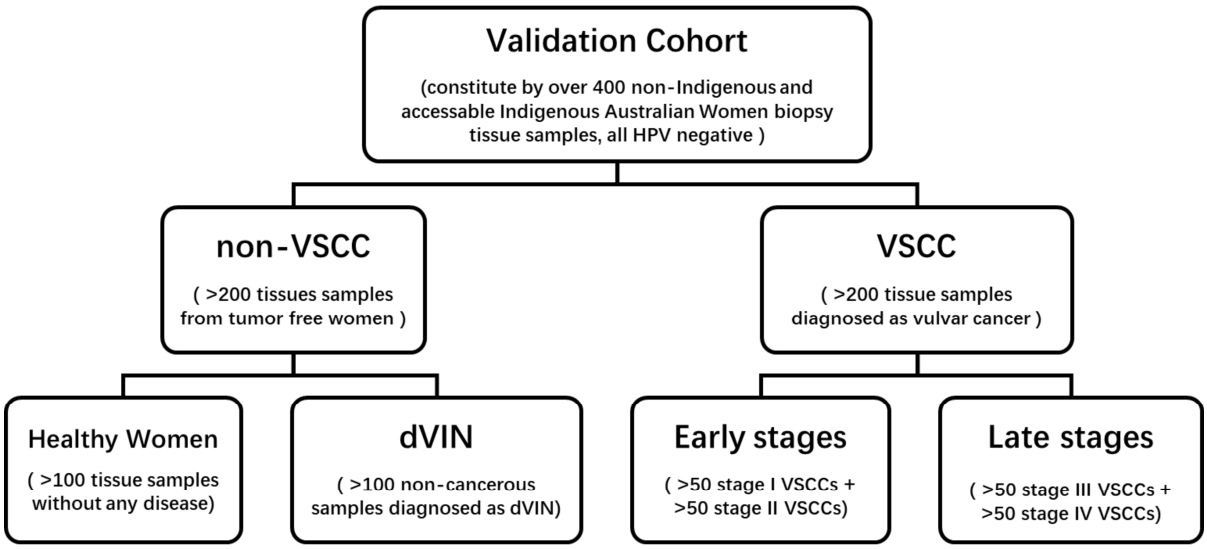


Figure 6.2: The study design of future vulvar cancer biomarker validation and expected patient numbers in each group of the validation cohort.

Currently, the results that show the differential abundance of NDRG1, RAP1A and POSTN in Indigenous VSCC patients have been concluded from only four representative samples. A significantly larger cohort of Indigenous VSCC tissues, especially those collected from patients living in remote communities with a high incidence rate, is required to provide more reliable results. In order to validate the protein expression changes described in this thesis, genomics and transcriptomics data is necessary as complementary evidence (Binder et al., 2015b). Unfortunately, no transcriptome database of HPV-independent vulvar cancer has been established, so it is difficult to explore the molecular pathological causes behind the changes observed at the protein level. Future RNA-sequencing of VSCC samples would likely highlight genetic mutations associated with observed protein signatures. Additionally, knockdown of candidate proteins in vulvar cancer cell lines would help elucidate the molecular functions and potential downstream targets. All of this would provide a better understanding of vulvar cancer progression overall.

Lastly, the ultimate goal of this thesis was to identify proteomics patterns from the primary tumour tissues which could indicate different progression stages of HPV-independent vulvar cancer. In the best case scenario, the CCA-LDA method applied in Chapter 4 could achieve

100% LNM classification rate from primary tumour tissues, the diagnostic sensitivity and specificity of potential biomarkers have not been applied to a broader cohort of VSCC tissues yet. This would require other high-throughput proteomics analysis methods, such as ELISA or protein microarray, to evaluate a significantly larger number of samples. Furthermore, how to improve the performance of protein markers in patient classification will be the factor that determines their potential as vulvar cancer diagnostic markers. Combining other pathological information, such as tumour sizes and grades, in the patient classification might be a promising solution as it has been successfully used in the endometrial cancer metastasis prediction (Mittal et al., 2017). To date, no molecular markers have been performed on the vulvar cancer tissue samples to help pathologists determining the progression stage of cancers other than the biopsy, while misclassification of vulvar cancer LNM status is still a challenge that affects the treatment and prognostic outcomes of vulvar cancer patients. Until now, the molecular markers are only used in classifying different etiological types of vulvar cancer in the clinical practice (Sandberg et al., 2012, Skapa et al., 2007). Furthermore, limited numbers of molecular studies and the high incidence rate of vulvar cancer in the Indigenous Australian community means a better understanding of the disease etiology is urgently required. Ultimately, the research findings presented here will incrementally aid the development of a more accurate diagnostic and classification system for HPV-independent vulvar cancer.

This page intentionally left blank

REFERENCE

REFERENCES:

AEBERSOLD, R. & MANN, M. 2003. Mass spectrometry-based proteomics. *Nature,* 422**,** 198-207.

AI, R., SUN, Y., GUO, Z., WEI, W., ZHOU, L., LIU, F., HENDRICKS, D. T., XU, Y. & ZHAO, X. 2016. NDRG1 overexpression promotes the progression of esophageal squamous cell carcinoma through modulating Wnt signaling pathway. *Cancer biology & therapy,* 17**,** 943-954.

AICHLER, M. & WALCH, A. 2015. MALDI Imaging mass spectrometry: current frontiers and perspectives in pathology research and practice. *Laboratory investigation,* 95**,** 422.

AIHW 2016. Gynaecological cancers in Australia An overview. *Cancer Serious.* Canberra: Australian Institute of Health and Welfare.

AKAGI, E. M., LAVORATO-ROCHA, A. M., DE MELO MAIA, B., RODRIGUES, I. S., CARVALHO, K. C., STIEPCICH, M. M., BAIOCCHI, G., SATO-KUWABARA, Y., ROGATTO, S. R. & SOARES, F. A. 2014. ROCK1 as a novel prognostic marker in vulvar cancer. *BMC cancer,* 14**,** 1.

ALLURI, P., LARIOS, J., MALIK, R., RAE, J. & CHINNAIYAN, A. 2016. Targeting Estrogen Receptor (ER) Mutations for Treatment of Endocrine Therapy Resistance in Breast Cancer. *International Journal of Radiation Oncology• Biology• Physics,* 96**,** S53.

ALMEIDA, R. A., GORDO, M., SILVA, F., ARAÚJO, R. C. D., RAMADA, M. H., ABRÃO, F. Y., COSTA, T. O., KOOLEN, H. H., DE SOUZA, A. D. & BLOCH JR, C. 2015. Cinerascetins, New Peptides from Hypsiboas cinerascens: MALDI LIFT-TOF-MS/MS de novo Sequence and Imaging Analysis. *Journal of the Brazilian Chemical Society,* 26**,** 2290-2297.

ANDERSEN, K., ZOBBE, V., THRANOV, I. R. & PEDERSEN, K. D. 2015. Relevance of computerized tomography in the preoperative evaluation of patients with vulvar cancer: a prospective study. *Cancer Imaging,* 15**,** 8.

ANDERSON, I., CRENGLE, S., LEIALOHA KAMAKA, M., CHEN, T.-H., PALAFOX, N. & JACKSON-PULVER, L. Indigenous health in Australia, New Zealand, and the Pacific. *The Lancet,* 367**,** 1775-1785.

ANDL, C. D., MIZUSHIMA, T., NAKAGAWA, H., OYAMA, K., HARADA, H., CHRUMA, K., HERLYN, M. & RUSTGI, A. K. 2003. Epidermal growth factor receptor mediates increased cell proliferation, migration, and aggregation in esophageal keratinocytes in vitro and in vivo. *Journal of Biological Chemistry,* 278**,** 1824-1830.

ANSINK, A. Vulvar squamous cell carcinoma. Seminars in dermatology, 1996. 51-59.

ARMITAGE, E. G. & CIBOROWSKI, M. 2017. Applications of metabolomics in cancer studies. *Metabolomics: From Fundamentals to Clinical Applications.* Springer.

ASARA, J. M., CHRISTOFK, H. R., FREIMARK, L. M. & CANTLEY, L. C. 2008. A label‐free quantification method by MS/MS TIC compared to SILAC and spectral counting in a proteomics screen. *Proteomics,* 8**,** 994-999.

AVIKI, E., ESSELEN, K., BARCIA, S., HOROWITZ, N., FELTMATE, C., BERKOWITZ, R., NUCCI, M. & MUTO, M. 2015. Trends and outcomes with utilization of plastic surgery for vulvar closure following vulvectomy for invasive squamous cell cancer. *Gynecologic oncology,* 137**,** 163.

BAIOCCHI, G., CESTARI, F. S., ROCHA, R., LAVORATO-ROCHA, A., MAIA, B., CESTARI, L., KUMAGAI, L., FALOPPA, C., FUKAZAWA, E. & BADIGLIAN-FILHO, L. 2013. Prognostic value of the number and laterality of metastatic inguinal lymph nodes in vulvar cancer: revisiting the FIGO staging system. *European Journal of Surgical Oncology,* 39**,** 780-785.

BALKWILL, F. R. 2012. The chemokine system and cancer. *The Journal of pathology,* 226**,** 148-157.

BALLUFF, B., RAUSER, S., MEDING, S., ELSNER, M., SCHÖNE, C., FEUCHTINGER, A., SCHUHMACHER, C., NOVOTNY, A., JÜTTING, U. & MACCARRONE, G. 2011a. MALDI imaging identifies prognostic seven-protein signature of novel tissue markers in intestinal-type gastric cancer. *The American journal of pathology,* 179**,** 2720-2729.

BALLUFF, B., SCHÖNE, C., HÖFLER, H. & WALCH, A. 2011b. MALDI imaging mass spectrometry for direct tissue analysis: technological advancements and recent applications. *Histochemistry & Cell Biology,* 136**,** 227-244.

BANTSCHEFF, M., SCHIRLE, M., SWEETMAN, G., RICK, J. & KUSTER, B. 2007. Quantitative mass spectrometry in proteomics: a critical review. *Anal Bioanal Chem,* 389**,** 1017-1031.
BARKATI, M., FORTIN, I., MILESHKIN, L., BERNSHAW, D., CARRIER, J.-F. & NARAYAN, K. 2013. Hemoglobin level in cervical cancer: a surrogate for an infiltrative phenotype. *International Journal of Gynecological Cancer,* 23**,** 724-729.
BENABDELLAH, F., SEYER, A., QUINTON, L., TOUBOUL, D., BRUNELLE, A. & LAPRÉVOTE, O. 2010. Mass spectrometry imaging of rat brain sections: nanomolar sensitivity with MALDI versus nanometer resolution by TOF–SIMS. *Analytical and bioanalytical chemistry,* 396**,** 151-162.
BENEDET, J., BENDER, H., JONES III, H., NGAN, H. & PECORELLI, S. 2000. FIGO staging classifications and clinical practice guidelines in the management of gynecologic cancers. *Int J Gynecol Obstet,* 70**,** 209-262.
BENJAMINI, Y. 2010. Discovering the false discovery rate. *Journal of the Royal Statistical Society: series B (statistical methodology),* 72**,** 405-416.
BERG, D., HIPP, S., MALINOWSKY, K., BÖLLNER, C. & BECKER, K.-F. 2010. Molecular profiling of signalling pathways in formalin-fixed and paraffin-embedded cancer tissues. *European Journal of Cancer,* 46**,** 47-55.
BERNHARDT, O. M., SELEVSEK, N., GILLET, L. C., RINNER, O., PICOTTI, P., AEBERSOLD, R. & REITER, L. 2012. Spectronaut: A fast and efficient algorithm for MRM-like processing of data independent acquisition (SWATH-MS) data. *Biognosys. ch*.
BHARADWAJ, S. & PRASAD, G. 2002. Tropomyosin-1, a novel suppressor of cellular transformation is downregulated by promoter methylation in cancer cells. *Cancer letters,* 183**,** 205-213.
BINDER, P. S., PRAT, J. & MUTCH, D. G. 2015a. The future role of molecular staging in gynecologic cancer. *International Journal of Gynecology & Obstetrics,* 131**,** S127-S131.
BINDER, P. S., PRAT, J. & MUTCH, D. G. 2015b. Molecular staging of gynaecological cancer – what is the future? *Best practice & research. Clinical obstetrics & gynaecology,* 29**,** 776-789.
BLEEKER, M. C., VISSER, P. J., OVERBEEK, L. I., VAN BEURDEN, M. & BERKHOF, J. 2016. Lichen sclerosus: incidence and risk of vulvar squamous cell carcinoma. *Cancer Epidemiology and Prevention Biomarkers,* 25**,** 1224-1230.
BOARD, P. A. T. E. 2016. Vulvar Cancer Treatment (PDQ®).
BODELON, C., MADELEINE, M. M., JOHNSON, L. G., DU, Q., GALLOWAY, D. A., MALKKI, M., PETERSDORF, E. W. & SCHWARTZ, S. M. 2014. Genetic variation in the TLR and NF‐κB pathways and cervical and vulvar cancer risk: A population‐based case–control study. *International journal of cancer,* 134**,** 437-444.
BODELON, C., MADELEINE, M. M., JOHNSON, L. G., DU, Q., MALKKI, M., PETERSDORF, E. W. & SCHWARTZ, S. M. 2012. Genetic variation in CD83 and risks of cervical and vulvar cancers: a population-based case–control study. *Gynecologic oncology,* 124**,** 525-528.
BOGDANOV, A., MOISEENKO, F. & DUBINA, M. 2017. Abnormal expression of ATP1A1 and ATP1A2 in breast cancer. *F1000Research,* 6.
BORRAS, G., MOLINA, R., XERCAVINS, J., BALLESTA, A. & IGLESIAS, J. 1995. Tumor antigens CA 19.9, CA 125, and CEA in carcinoma of the uterine cervix. *Gynecologic oncology,* 57**,** 205-211.
BOTELHO, M. C., LAVORATO-ROCHA, A. M., RODRIGUES, I. S., MAIA, B. M., CARVALHO, K. C., PUGA, R., SOARES, F. A. & ROCHA, R. M. 2015. S100P: Cause or effect of vulvar carcinoma prognostic status. *Cancer Research,* 75**,** 516-516.
BRAGA-NETO, U. M. & DOUGHERTY, E. R. 2004. Is cross-validation valid for small-sample microarray classification? *Bioinformatics,* 20**,** 374-380.
BRODT, P., FALLAVOLLITA, L., BRESALIER, R. S., METERISSIAN, S., NORTON, C. R. & WOLITZKY, B. A. 1997. Liver endothelial E‐selectin mediates carcinoma cell adhesion and promotes liver metastasis. *International Journal of Cancer,* 71**,** 612-619.
BRYAN, R. T., WEI, W., SHIMWELL, N. J., COLLINS, S. I., HUSSAIN, S. A., BILLINGHAM, L. J., MURRAY, P. G., DESHMUKH, N., JAMES, N. D. & WALLACE, D. M. A. 2011. Assessment of high‐

throughput high‐resolution MALDI‐TOF‐MS of urinary peptides for the detection of muscle‐invasive bladder cancer. *PROTEOMICS-Clinical Applications,* 5**,** 493-503.
BUSFIELD, F., DUFFY, D. L., KESTING, J. B., WALKER, S. M., LOVELOCK, P. K., GOOD, D., TATE, H., WATEGO, D., MARCZAK, M. & HAYMAN, N. 2002. A Genomewide search for type 2 diabetes–susceptibility genes in Indigenous Australians. *The American Journal of Human Genetics,* 70**,** 349-357.
CACCIATORE, S., ZADRA, G., BANGO, C., PENNEY, K. L., TYEKUCHEVA, S., YANES, O. & LODA, M. 2017. Metabolic profiling in formalin-fixed and paraffin-embedded prostate cancer tissues. *Molecular Cancer Research*.
CAI, Y., SUKHOVA, G. K., WONG, H. K., XU, A., TERGAONKAR, V., VANHOUTTE, P. M. & TANG, E. H. C. 2015. Rap1 induces cytokine production in pro-inflammatory macrophages through NFκB signaling and is highly expressed in human atherosclerotic lesions. *Cell Cycle,* 14**,** 3580-3592.
CAMPELLO, R. J., MOULAVI, D. & SANDER, J. Density-based clustering based on hierarchical density estimates. Pacific-Asia conference on knowledge discovery and data mining, 2013. Springer, 160-172.
CANAVAN, T. P. & COHEN, D. 2002. Vulvar cancer. *American family physician,* 66**,** 1269-1274.
CAPRIOLI, R. M., FARMER, T. B. & GILE, J. 1997. Molecular imaging of biological samples: localization of peptides and proteins using MALDI-TOF MS. *Analytical chemistry,* 69**,** 4751-4760.
CARLSON, B. C., HOFER, M. D., BALLEK, N., YANG, X. J., MEEKS, J. J. & GONZALEZ, C. M. 2013. Protein Markers of Malignant Potential in Penile and Vulvar Lichen Sclerosus. *The Journal of Urology,* 190**,** 399-406.
CHAFE, W., RICHARDS, A., MORGAN, L. & WILKINSON, E. 1988. Unrecognized invasive carcinoma in vulvar intraepithelial neoplasia (VIN). *Gynecologic oncology,* 31**,** 154-162.
CHAO, A., WANG, T. H., LEE, Y. S., HSUEH, S., CHAO, A. S., CHANG, T. C., KUNG, W. H., HUANG, S. L., CHAO, F. Y. & WEI, M. L. 2006. Molecular characterization of adenocarcinoma and squamous carcinoma of the uterine cervix using microarray analysis of gene expression. *International journal of cancer,* 119**,** 91-98.
CHEN, J., LI, S., YANG, Z., LU, G. & HU, H. 2008. Correlation between NDRG1 and PTEN expression in endometrial carcinoma. *Cancer science,* 99**,** 706-710.
CHEN, S., O'SULLIVAN, H., TABRIZI, S. N., FAIRLEY, C. K., QUINN, M. A. & GARLAND, S. M. 1999. Prevalence and genotyping of HPV in cervical cancer among Australian women. *International Journal of Gynecology & Obstetrics,* 67**,** 163-168.
CHI, J.-T., WANG, Z., NUYTEN, D. S., RODRIGUEZ, E. H., SCHANER, M. E., SALIM, A., WANG, Y., KRISTENSEN, G. B., HELLAND, Å. & BØRRESEN-DALE, A.-L. 2006. Gene expression programs in response to hypoxia: cell type specificity and prognostic significance in human cancers. *PLoS medicine,* 3**,** e47.
CHU, P. G. & WEISS, L. M. 2002a. Expression of cytokeratin 5/6 in epithelial neoplasms: an immunohistochemical study of 509 cases. *Modern pathology,* 15**,** 6.
CHU, P. G. & WEISS, L. M. 2002b. Expression of cytokeratin 5/6 in epithelial neoplasms: an immunohistochemical study of 509 cases. *Mod Pathol,* 15**,** 6-10.
COLLARINO, A., DONSWIJK, M. L., VAN DRIEL, W. J., STOKKEL, M. P. & OLMOS, R. A. V. 2015. The use of SPECT/CT for anatomical mapping of lymphatic drainage in vulvar cancer: possible implications for the extent of inguinal lymph node dissection. *European journal of nuclear medicine and molecular imaging,* 42**,** 2064-2071.
CONDON, J., BARNES, A. & CUNNINGHAM, J. 2003. Cancer in Indigenous Australians: a review. *Cancer Causes & Control,* 14**,** 109-121.
CONDON, J., RUMBOLD, A., THORN, J., O'BRIEN, M., DAVY, M. & ZARDAWI, I. 2009a. A cluster of vulvar cancer and vulvar intraepithelial neoplasia in young Australian Indigenous women. *Cancer Causes & Control,* 20**,** 67-74.

CONDON, J. R., ARMSTRONG, B. K., BARNES, T. & ZHAO, Y. 2005. Cancer incidence and survival for indigenous Australians in the Northern Territory. *Australian and New Zealand journal of public health,* 29**,** 123-128.

CONDON, J. R., RUMBOLD, A. R., THORN, J. C., O'BRIEN, M. M., DAVY, M. J. & ZARDAWI, I. 2009b. A cluster of vulvar cancer and vulvar intraepithelial neoplasia in young Australian Indigenous women. *Cancer Causes & Control,* 20**,** 67-74.

COORY, M., THOMPSON, A. & GANGULY, I. 2000. Cancer among people living in rural and remote Indigenous communities in Queensland. *The Medical Journal of Australia,* 173**,** 301-304.

COSME, J., TCHERNYSHYOV, I. & GRAMOLINI, A. 2016. Basic Concepts in Mass Spectrometry and Protein Quantitation. *Manual of Cardiovascular Proteomics.* Springer.

COURTNEY-BROOKS, M., SUKUMVANICH, P., BERIWAL, S., ZORN, K. K., RICHARD, S. D. & KRIVAK, T. C. 2010. Does the number of nodes removed impact survival in vulvar cancer patients with node-negative disease? *Gynecologic oncology,* 117**,** 308-311.

COUSSENS, L. M. & WERB, Z. 2002. Inflammation and cancer. *Nature,* 420**,** 860.

COX, A. J., MOSCOVIS, S. M., BLACKWELL, C. C. & SCOTT, R. J. 2014a. Cytokine gene polymorphism among Indigenous Australians. *Innate immunity,* 20**,** 431-439.

COX, J., HEIN, M. Y., LUBER, C. A., PARON, I., NAGARAJ, N. & MANN, M. 2014b. Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Molecular & Cellular Proteomics,* 13**,** 2513-2526.

CUNNINGHAM, J., RUMBOLD, A. R., ZHANG, X. & CONDON, J. R. 2008. Incidence, aetiology, and outcomes of cancer in Indigenous peoples in Australia. *The Lancet Oncology,* 9**,** 585-595.

DA SILVA, P. L., DO AMARAL, V. C., GABRIELLI, V., GUEVARA, M. M. M., MANNELLA, P., BARACAT, E. C., SOARES-JR, J. M. & SIMONCINI, T. 2015. Prolactin promotes breast cancer cell migration through actin cytoskeleton remodeling. *Frontiers in endocrinology,* 6.

DAI, X., CHEN, A. & BAI, Z. 2014. Integrative investigation on breast cancer in ER, PR and HER2-defined subgroups using mRNA and miRNA expression profiling. *Scientific reports,* 4.

DALING, J. R., SHERMAN, K. J., HISLOP, T. G., MADEN, C., MANDELSON, M. T., BECKMANN, A. M. & WEISS, N. S. 1992. Cigarette smoking and the risk of anogenital cancer. *American journal of epidemiology,* 135**,** 180-189.

DAVIS, B. W., GELBER, R. D., GOLDHIRSCH, A., HARTMANN, W. H., LOCHER, G. W., REED, R., GOLOUH, R., SÜE‐SÖDERBERGH, J., HOLLOWAY, L. & RUSSELL, I. 1986. Prognostic significance of tumor grade in clinical trials of adjuvant therapy for breast cancer with axillary lymph node metastasis. *Cancer,* 58**,** 2662-2670.

DE CICCO, C., SIDERI, M., BARTOLOMEI, M., GRANA, C., CREMONESI, M., FIORENZA, M., MAGGIONI, A., BOCCIOLONE, L., MANGIONI, C. & COLOMBO, N. 2000. Sentinel node biopsy in early vulvar cancer. *British journal of cancer,* 82**,** 295.

DE LAS HERAS-RUBIO, A., PERUCHO, L., PACIUCCI, R., VILARDELL, J. & LLEONART, M. E. 2014. Ribosomal proteins as novel players in tumorigenesis. *Cancer and Metastasis Reviews,* 33**,** 115-141.

DE NOO, M. E., MERTENS, B. J., ÖZALP, A., BLADERGROEN, M. R., VAN DER WERFF, M. P., VAN DE VELDE, C. J., DEELDER, A. M. & TOLLENAAR, R. A. 2006. Detection of colorectal cancer using MALDI-TOF serum protein profiling. *European Journal of Cancer,* 42**,** 1068-1076.

DIAZ, A., MOORE, S. P., MARTIN, J. H., GREEN, A. C., GARVEY, G. & VALERY, P. C. 2015. Factors associated with cancer-specific and overall survival among Indigenous and non-Indigenous gynecologic cancer patients in Queensland, Australia: a matched cohort study. *International Journal of Gynecological Cancer,* 25**,** 542-547.

DOMMETT, R., KLEIN, N. & TURNER, M. 2006. Mannose‐binding lectin in innate immunity: past, present and future. *HLA,* 68**,** 193-209.

DORING, G. 1994. The role of neutrophil elastase in chronic inflammation. *American journal of respiratory and critical care medicine,* 150**,** S114.

EBERT, M. P., NIEMEYER, D., DEININGER, S. O., WEX, T., KNIPPIG, C., HOFFMANN, J., SAUER, J., ALBRECHT, W., MALFERTHEINER, P. & RÖCKEN, C. 2006. Identification and confirmation of increased fibrinopeptide a serum protein levels in gastric cancer sera by magnet bead assisted MALDI-TOF mass spectrometry. *Journal of proteome research,* 5**,** 2152-2158.

ECKEY, C., HÖHN, A., HÖCKEL, M., SCHERLING, R., HENTSCHEL, B. & HORN, L. 2014. Epithelial-mesenchymal transition (EMT) in vulvar cancer with and without inguinal lymph node metastases. *Geburtshilfe und Frauenheilkunde,* 74**,** A34.

EDWARDS, C. L., TORTOLERO-LUNA, G., LINARES, A. C., MALPICA, A., BAKER, V. V., COOK, E., JOHNSON, E. & FOLLEN, M. M. 1996. Vulvar intraepithelial neoplasia and vulvar cancer. *Obstetrics and gynecology clinics of North America,* 23**,** 295-324.

ELLEN, T. P., KE, Q., ZHANG, P. & COSTA, M. 2007. NDRG1, a growth and cancer related gene: regulation of gene expression and function in normal and disease states. *Carcinogenesis,* 29**,** 2-8.

ENDO, D., TODO, Y., OKAMOTO, K., MINOBE, S., KATO, H. & NISHIYAMA, N. 2015. Prognostic factors for patients with cervical cancer treated with concurrent chemoradiotherapy: a retrospective analysis in a Japanese cohort. *Journal of gynecologic oncology,* 26**,** 12-18.

FANG, B. A., KOVAČEVIĆ, Ž., PARK, K. C., KALINOWSKI, D. S., JANSSON, P. J., LANE, D. J., SAHNI, S. & RICHARDSON, D. R. 2014. Molecular functions of the iron-regulated metastasis suppressor, NDRG1, and its potential as a molecular target for cancer therapy. *Biochimica et Biophysica Acta (BBA)-Reviews on Cancer,* 1845**,** 1-19.

FOWLER, C. B., O'LEARY, T. J. & MASON, J. T. 2011. Protein mass spectrometry applications on FFPE tissue sections. *Methods in molecular biology (Clifton, N.J.) U6 - ctx_ver=Z39.88-2004&ctx_enc=info%3Aofi%2Fenc%3AUTF-8&rfr_id=info:sid/summon.serialssolutions.com&rft_val_fmt=info:ofi/fmt:kev:mtx:journal&rft.genre=article&rft.atitle=Protein+mass+spectrometry+applications+on+FFPE+tissue+sections&rft.jtitle=Methods+in+molecular+biology+%28Clifton%2C+N.J.%29&rft.au=Fowler%2C+Carol+B&rft.au=O%27Leary%2C+Timothy+J&rft.au=Mason%2C+Jeffrey+T&rft.date=2011-01-01&rft.volume=724&rft.spage=281&rft_id=info:pmid/21370020&rft.externalDocID=21370020 U7 - Journal Article U8 - FETCH-pubmed_primary_213700201,* 724**,** 281.

FRANCK, J., ARAFAH, K., ELAYED, M., BONNEL, D., VERGARA, D., JACQUET, A., VINATIER, D., WISZTORSKI, M., DAY, R. & FOURNIER, I. 2009. MALDI imaging mass spectrometry state of the art technology in clinical proteomics. *Molecular & Cellular Proteomics,* 8**,** 2023-2033.

FREDERICK, B. A., HELFRICH, B. A., COLDREN, C. D., ZHENG, D., CHAN, D., BUNN, P. A. & RABEN, D. 2007. Epithelial to mesenchymal transition predicts gefitinib resistance in cell lines of head and neck squamous cell carcinoma and non–small cell lung carcinoma. *Molecular cancer therapeutics,* 6**,** 1683-1691.

FROST, J., LUDEMAN, L., HILLABY, K., GORNALL, R., LLOYD, G., KENDALL, C., SHORE, A. C. & STONE, N. 2016. Identification of cancer associated molecular changes in histologically benign vulval disease found in association with vulval squamous cell carcinoma using Fourier transform infrared spectroscopy. *Analytical Methods,* 8**,** 8452-8460.

FUCHS, B., SÜß, R. & SCHILLER, J. 2010. An update of MALDI-TOF mass spectrometry in lipid research. *Progress in lipid research,* 49**,** 450-475.

FUNASAKA, T., RAZ, A. & NANGIA-MAKKER, P. 2014. Galectin-3 in angiogenesis and metastasis. *Glycobiology,* 24**,** 886-891.

GAO, L., MA, W., CHEN, J., WANG, K., LI, J., WANG, S., BEKES, F., APPELS, R. & YAN, Y. 2010. Characterization and comparative analysis of wheat high molecular weight glutenin subunits by SDS-PAGE, RP-HPLC, HPCE, and MALDI-TOF-MS. *J Agric Food Chem,* 58**,** 2777-2786.

GAO, Y., WU, X., YANG, Y., RUAN, Y., YANG, X., ZHU, Y., YANG, H., ZHANG, Z., YANG, C. & LI, Z. 2016. High expression of flotillin-1 is associated with lymph node metastasis and poor prognosis in vulvar squamous cell carcinoma. *Int J Clin Exp Pathol,* 9**,** 9538-9546.

GARLAND, S. M., BROTHERTON, J. M., CONDON, J. R., MCINTYRE, P. B., STEVENS, M. P., SMITH, D. W. & TABRIZI, S. N. 2011. Human papillomavirus prevalence among indigenous and non-indigenous Australian women prior to a national HPV vaccination program. *BMC medicine,* 9**,** 104.

GESSEL, M. M., NORRIS, J. L. & CAPRIOLI, R. M. 2014. MALDI imaging mass spectrometry: spatial molecular analysis to enable a new age of discovery. *Journal of proteomics,* 107**,** 71-82.

GIKA, H. G., THEODORIDIS, G. A., PLUMB, R. S. & WILSON, I. D. 2014. Current practice of liquid chromatography–mass spectrometry in metabolomics and metabonomics. *Journal of pharmaceutical and biomedical analysis,* 87**,** 12-25.

GIL, E. B., MALONE LINK, E., LIU, L. X., JOHNSON, C. D. & LEES, J. A. 1999. Regulation of the insulin-like developmental pathway of Caenorhabditis elegans by a homolog of the PTEN tumor suppressor gene. *Proceedings of the National Academy of Sciences,* 96**,** 2925-2930.

GILLET, L. C., NAVARRO, P., TATE, S., RÖST, H., SELEVSEK, N., REITER, L., BONNER, R. & AEBERSOLD, R. 2012. Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. *Molecular & Cellular Proteomics,* 11**,** O111. 016717.

GIOVANNUCCI, E. 2003. Nutrition, insulin, insulin-like growth factors and cancer. *Hormone and Metabolic Research,* 35**,** 694-704.

GOEMINNE, L. J., GEVAERT, K. & CLEMENT, L. 2016. Peptide-level robust ridge regression improves estimation, sensitivity, and specificity in data-dependent quantitative label-free shotgun proteomics. *Molecular & Cellular Proteomics,* 15**,** 657-668.

GONZALEZ, M. E., MARTIN, E. E. & KLEER, C. G. 2015. Human breast carcinoma-associated mesenchymal stem cells promote breast cancer cell proliferation, irreversible EMT and invasion through Collagen I. AACR.

GOODWIN, R. J. 2012. Sample preparation for mass spectrometry imaging: small mistakes can lead to big consequences. *JOURNAL OF PROTEOMICS,* 75**,** 4893-4911.

GOODWIN, R. J., DUNGWORTH, J. C., COBB, S. R. & PITT, A. R. 2008a. Time-dependent evolution of tissue markers by MALDI-MS imaging. *PROTEOMICS,* 8**,** 3801-8.

GOODWIN, R. J. A., PENNINGTON, S. R. & PITT, A. R. 2008b. Protein and peptides in pictures: Imaging with MALDI mass spectrometry. *PROTEOMICS,* 8**,** 3785-3800.

GORDINIER, M. E., STEINHOFF, M. M., HOGAN, J. W., PEIPERT, J. F., GAJEWSKI, W. H., FALKENBERRY, S. S. & GRANAI, C. 1997. S-Phase fraction, p53, and HER-2/neu status as predictors of nodal metastasis in early vulvar cancer. *Gynecologic oncology,* 67**,** 200-202.

GOYAL, A., SOLUS, J. F., CHAN, M. P., DOYLE, L. A., SCHAFFER, A., THAKURIA, M., HORN, T. D., DUNCAN, L. M. & NAZARIAN, R. M. 2016. Cytokeratin 17 is highly sensitive in discriminating cutaneous lymphadenoma (a distinct trichoblastoma variant) from basal cell carcinoma. *Journal of cutaneous pathology,* 43**,** 422-429.

GROSECLOSE, M. R., MASSION, P. P., CHAURAND, P. & CAPRIOLI, R. M. 2008. High‐throughput proteomic analysis of formalin‐fixed paraffin‐embedded tissue microarrays using MALDI imaging mass spectrometry. *PROTEOMICS,* 8**,** 3715-3724.

GROßE-STEFFEN, T., GIESE, T., GIESE, N., LONGERICH, T., SCHIRMACHER, P., HÄNSCH, G. M. & GAIDA, M. M. 2012. Epithelial-to-mesenchymal transition in pancreatic ductal adenocarcinoma and pancreatic tumor cell lines: the role of neutrophils and neutrophil-derived elastase. *Clinical and Developmental Immunology,* 2012.

GRULICH, A. E., JIN, F., CONWAY, E. L., STEIN, A. N. & HOCKING, J. 2010. Cancers attributable to human papillomavirus infection. *Sexual Health,* 7**,** 244-252.

GUSTAFSSON, J. O., EDDES, J. S., MEDING, S., KOUDELKA, T., OEHLER, M. K., MCCOLL, S. R. & HOFFMANN, P. 2012. Internal calibrants allow high accuracy peptide matching between MALDI imaging MS and LC-MS/MS. *Journal of proteomics,* 75**,** 5093-5105.

GUSTAFSSON, J. O., OEHLER, M. K., RUSZKIEWICZ, A., MCCOLL, S. R. & HOFFMANN, P. 2011a. MALDI Imaging Mass Spectrometry (MALDI-IMS)-Application of Spatial Proteomics for Ovarian Cancer Classification and Diagnosis. *Int J Mol Sci,* 12**,** 773-94.

GUSTAFSSON, J. O., OEHLER, M. K., RUSZKIEWICZ, A., MCCOLL, S. R. & HOFFMANN, P. 2011b. MALDI Imaging Mass Spectrometry (MALDI-IMS)—Application of Spatial Proteomics for Ovarian Cancer Classification and Diagnosis. *International journal of molecular sciences,* 12**,** 773-794.

GUSTAFSSON, J. O. R., OEHLER, M. K., MCCOLL, S. R. & HOFFMANN, P. 2010. Citric acid antigen retrieval (CAAR) for tryptic peptide imaging directly on archived formalin-fixed paraffin-embedded tissue. *Journal of proteome research,* 9**,** 4315.

GUTHRIE, G. J., CHARLES, K. A., ROXBURGH, C. S., HORGAN, P. G., MCMILLAN, D. C. & CLARKE, S. J. 2013. The systemic inflammation-based neutrophil–lymphocyte ratio: experience in patients with cancer. *Critical reviews in oncology/hematology,* 88**,** 218-230.

GYŐRFFY, B., LÁNCZKY, A. & SZÁLLÁSI, Z. 2012. Implementing an online tool for genome-wide validation of survival-associated biomarkers in ovarian-cancer using microarray data from 1287 patients. *Endocrine-related cancer,* 19**,** 197-208.

HACKER, N. F., BEREK, J. S., LAGASSE, L. D., LEUCHTER, R. S. & MOORE, J. G. 1986. Management of regional lymph nodes and their prognostic influence on vulvar cancer. *Gynaecological Oncology.* Springer.

HAMPL, M., DECKERS-FIGIEL, S., HAMPL, J. A., REIN, D. & BENDER, H. G. 2008. New aspects of vulvar cancer: changes in localization and age of onset. *Gynecologic oncology,* 109**,** 340-345.

HAN, G., LIM, D., LEITAO, M. M., ABU‐RUSTUM, N. R. & SOSLOW, R. A. 2014. Histological features associated with occult lymph node metastasis in FIGO clinical stage I, grade I endometrioid carcinoma. *Histopathology,* 64**,** 389-398.

HAUSPY, J., BEINER, M., HARLEY, I., EHRLICH, L., RASTY, G. & COVENS, A. 2007. Sentinel lymph node in vulvar cancer. *Cancer,* 110**,** 1015-1023.

HAY, C. M., LACHANCE, J. A., LUCAS, F., SMITH, K. A. & JONES, M. A. 2016. Biomarkers p16, human papillomavirus and p53 predict recurrence and survival in early stage squamous cell carcinoma of the vulva. *Journal of lower genital tract disease,* 20**,** 252-256.

HEFLER-FRISCHMUTH, K., SEEBACHER, V., POLTERAUER, S., TEMPFER, C., REINTHALLER, A. & HEFLER, L. 2010. The inflammation-based modified Glasgow Prognostic Score in patients with vulvar cancer. *European Journal of Obstetrics and Gynecology and Reproductive Biology,* 149**,** 102-105.

HEFLER, L., TEMPFER, C., OBERMAIR, A., FRISCHMUTH, K., SLIUTZ, G., REINTHALLER, A., LEODOLTER, S. & KAINZ, C. 1999. Serum concentrations of vascular endothelial growth factor in vulvar cancer. *Clinical cancer research,* 5**,** 2806-2809.

HEFLER, L. A., SLIUTZ, G., LEODOLTER, S., SPEISER, P., JOURA, E., REINTHALLER, A. & KOHLBERGER, P. 2005. Squamous cell carcinoma antigen serum levels as prognostic parameter in patients with early stage vulvar cancer. *Gynecologic oncology,* 97**,** 904-907.

HEROUX, M. S., MARLA, A. C., HALLIGAN, B. D., AL-GIZAWIY, M., JENNIFER, M., MALKIN, M. G., KATHLEEN, M. S. & MIRZA, S. P. 2014. Comprehensive Characterization of Glioblastoma Tumor Tissues for Biomarker Identification using Mass Spectrometry-based Label-free Quantitative Proteomics 2. *Cancer,* 39**,** 40.

HILDESHEIM, A., HAN, C.-L., BRINTON, L. A., KURMAN, R. J. & SCHILLER, J. T. 1997. Human papillomavirus type 16 and risk of preinvasive and invasive vulvar cancer: results from a seroepidemiological case-control study. *Obstetrics & Gynecology,* 90**,** 748-754.

HILL, J. J., TREMBLAY, T.-L., PEN, A., LI, J., ROBOTHAM, A. C., LENFERINK, A. E., WANG, E., O'CONNOR-MCCOURT, M. & KELLY, J. F. 2011. Identification of Vascular Breast Tumor Markers by Laser Capture Microdissection and Label-Free LC− MS. *Journal of proteome research,* 10**,** 2479-2493.

HOBBS, R., BATAZZI, A., HAN, M. & COULOMBE, P. 2016. Loss of keratin 17 induces tissue-specific cytokine polarization and cellular differentiation in HPV16-driven cervical tumorigenesis in vivo. *Oncogene,* 35**,** 5653.

HÖCKEL, M., SCHLENGER, K., ARAL, B., MITZE, M., SCHÄFFER, U. & VAUPEL, P. 1996. Association between tumor hypoxia and malignant progression in advanced cancer of the uterine cervix. *Cancer research,* 56**,** 4509-4515.

HOEVENAARS, B. M., VAN DER AVOORT, I. A., DE WILDE, P., MASSUGER, L. F., MELCHERS, W. J., DE HULLU, J. A. & BULTEN, J. 2008. A panel of p16INK4A, MIB1 and p53 proteins can distinguish between the 2 pathways leading to vulvar squamous cell carcinoma. *International journal of cancer,* 123**,** 2767-2773.

HOLST, S., HEIJS, B., DE HAAN, N., VAN ZEIJL, R. J., BRIAIRE-DE BRUIJN, I. H., VAN PELT, G. W., MEHTA, A. S., ANGEL, P. M., MESKER, W. E. & TOLLENAAR, R. A. 2016. Linkage-specific in situ sialic acid derivatization for N-glycan mass spectrometry imaging of formalin-fixed paraffin-embedded tissues. *Analytical chemistry,* 88**,** 5904-5913.

HOLTHOFF, E. R., BYRUM, S., KELLY, T., TACKETT, A., QUICK, C. M. & POST, S. R. 2016. Proteomic Profiles of Vulvar Squamous Cell Carcinoma Variants. *The FASEB Journal,* 30**,** 1081.1-1081.1.

HØRDING, U., BOCK, J. E., KRINGSHOLM, B., ANDREASSON, B., VISFELDT, J. & DAUGAARD, S. 1993. Human papillomavirus in vulvar squamous‐cell carcinoma and in normal vulvar tissues: A search for a possible impact of HPV on vulvar cancer prognosis. *International journal of cancer,* 55**,** 394-396.

HØRDING, U., JUNGE, J., DAUGAARD, S., LUNDVALL, F., POULSEN, H. & BOCK, J. E. 1994. Vulvar squamous cell carcinoma and papillomaviruses: indications for two different etiologies. *Gynecologic oncology,* 52**,** 241-246.

HOSPER, N. A., VAN DEN BERG, P. P., DE ROND, S., POPA, E. R., WILMER, M. J., MASEREEUW, R. & BANK, R. A. 2013. Epithelial-to-mesenchymal transition in fibrosis: collagen type I expression is highly upregulated after EMT, but does not contribute to collagen deposition. *Experimental cell research,* 319**,** 3000-3009.

HOUGHTON, A. M., RZYMKIEWICZ, D. M., JI, H., GREGORY, A. D., EGEA, E. E., METZ, H. E., STOLZ, D. B., LAND, S. R., MARCONCINI, L. A. & KLIMENT, C. R. 2010. Neutrophil elastase–mediated degradation of IRS-1 accelerates lung tumor growth. *Nature medicine,* 16**,** 219.

HOWLADER N, N. A., KRAPCHO M, MILLER D, BISHOP K, KOSARY CL, YU M, RUHL J, TATALOVICH Z, MARIOTTO A, LEWIS DR, CHEN HS, FEUER EJ, CRONIN KA 2017. SEER Cancer Statistics Review (CSR) 1975-2014. Bethesda, MD: National Cancer Institute.

HUNG, C. F., WU, T., MONIE, A. & RODEN, R. 2008. Antigen‐specific immunotherapy of cervical and ovarian cancer. *Immunological reviews,* 222**,** 43-69.

IKEDA, K., MONDEN, T., KANOH, T., TSUJIE, M., IZAWA, H., HABA, A., OHNISHI, T., SEKIMOTO, M., TOMITA, N., SHIOZAKI, H. & MONDEN, M. 1998. Extraction and analysis of diagnostically useful proteins from formalin-fixed, paraffin-embedded tissue sections. *J Histochem Cytochem,* 46**,** 397-403.

ISHII, N., MAIER, D., MERLO, A., TADA, M., SAWAMURA, Y., DISERENS, A. C. & MEIR, E. G. 1999. Frequent Co‐Alterations of TP53, p16/CDKN2A, p14ARF, PTEN Tumor Suppressor Genes in Human Glioma Cell Lines. *Brain pathology,* 9**,** 469-479.

ISONO, T., KIM, C. J., ANDO, Y., SAKURAI, H., OKADA, Y. & INOUE, H. 2009. Suppression of cell invasiveness by periostin via TAB1/TAK1. *International journal of oncology,* 35**,** 425-432.

ISSAQ, H. J. & VEENSTRA, T. D. 2008. Would you prefer multiple reaction monitoring or antibodies with your biomarker validation? *Expert review of proteomics,* 5**,** 761-763.

IZUISHI, K., KATO, K., OGURA, T., KINOSHITA, T. & ESUMI, H. 2000. Remarkable tolerance of tumor cells to nutrient deprivation: possible new biochemical target for cancer therapy. *Cancer research,* 60**,** 6201-6207.

JACOB, J. T., COULOMBE, P. A., KWAN, R. & OMARY, M. B. 2018. Types I and II keratin intermediate filaments. *Cold Spring Harbor Perspectives in Biology,* 10**,** a018275.

JIA, J., ZHAO, G., YUE, M., YANG, H. & WANG, J. 2015. NDRG1 promotes the multidrug resistance of neuroblastoma cells with upregulated expression of drug resistant proteins. *Biomedicine & Pharmacotherapy,* 76**,** 46-51.

JONES, R. W. & ROWAN, D. M. 1994. Vulvar intraepithelial neoplasia III: a clinical study of the outcome in 113 cases with relation to the later development of invasive vulvar carcinoma. *Obstetrics & Gynecology,* 84**,** 741-745.

JONES, R. W., SADLER, L., GRANT, S., WHINERAY, J., EXETER, M. & ROWAN, D. 2005. Clinically identifying women with vulvar lichen sclerosus at increased risk of squamous cell carcinoma. *Obstetrical & gynecological survey,* 60**,** 98-99.

JOURA, E. A. 2002. Epidemiology, diagnosis and treatment of vulvar intraepithelial neoplasia. *Current Opinion in Obstetrics and Gynecology,* 14**,** 39-43.

KAP, M., SMEDTS, F., OOSTERHUIS, W., WINTHER, R., CHRISTENSEN, N., REISCHAUER, B., VIERTLER, C., GROELZ, D., BECKER, K.-F., ZATLOUKAL, K., LANGER, R., SLOTTA-HUSPENINA, J., BODO, K., DE JONG, B., OELMULLER, U. & RIEGMAN, P. 2011. Histological Assessment of PAXgene Tissue Fixation and Stabilization Reagents. *PLoS One,* 6**,** e27704.

KEARNEY, P. E., FANG, K. C., LI, X.-J. & HAYWARD, C. 2015. Compositions, methods and kits for diagnosis of lung cancer. Google Patents.

KEDRIN, D., VAN RHEENEN, J., HERNANDEZ, L., CONDEELIS, J. & SEGALL, J. E. 2007. Cell motility and cytoskeletal regulation in invasion and metastasis. *Journal of mammary gland biology and neoplasia,* 12**,** 143-152.

KELLIE, J. F., TRAN, J. C., LEE, J. E., AHLF, D. R., THOMAS, H. M., NTAI, I., CATHERMAN, A. D., DURBIN, K. R., ZAMDBORG, L. & VELLAICHAMY, A. 2010. The emerging process of Top Down mass spectrometry for protein analysis: biomarkers, protein-therapeutics, and achieving high throughput. *Molecular BioSystems,* 6**,** 1532-1539.

KENTER, G. G., WELTERS, M. J., VALENTIJN, A. R. P., LOWIK, M. J., BERENDS-VAN DER MEER, D. M., VLOON, A. P., ESSAHSAH, F., FATHERS, L. M., OFFRINGA, R. & DRIJFHOUT, J. W. 2009. Vaccination against HPV-16 oncoproteins for vulvar intraepithelial neoplasia. *New England Journal of Medicine,* 361**,** 1838-1847.

KHANNA, N., RAUH, L. A., LACHIEWICZ, M. P. & HOROWITZ, I. R. 2016a. Margins for cervical and vulvar cancer. *J Surg Oncol*.

KHANNA, N., RAUH, L. A., LACHIEWICZ, M. P. & HOROWITZ, I. R. 2016b. Margins for cervical and vulvar cancer. *Journal of surgical oncology,* 113**,** 304-309.

KHEIRELSEID, E. A., CHANG, K. H., NEWELL, J., KERIN, M. J. & MILLER, N. 2010. Identification of endogenous control genes for normalisation of real-time quantitative PCR data in colorectal cancer. *BMC molecular biology,* 11**,** 12.

KIM, C. J., SAKAMOTO, K., TAMBE, Y. & INOUE, H. 2011. Opposite regulation of epithelial-to-mesenchymal transition and cell invasiveness by periostin between prostate and bladder cancer cells. *International journal of oncology,* 38**,** 1759-1766.

KIM, C. J., YOSHIOKA, N., TAMBE, Y., KUSHIMA, R., OKADA, Y. & INOUE, H. 2005. Periostin is down‐regulated in high grade human bladder cancers and suppresses in vitro cell invasiveness and in vivo metastasis of cancer cells. *International journal of cancer,* 117**,** 51-58.

KIM, H.-S., LEE, J.-J., DO, S.-I., KIM, K., DO, I.-G., KIM, D.-H., CHAE, S. W. & SOHN, J. H. 2015. Overexpression of cytokeratin 17 is associated with the development of papillary thyroid carcinoma and the presence of lymph node metastasis. *International Journal of Clinical and Experimental Pathology,* 8**,** 5695-5701.

KLEBANOFF, S. J., VADAS, M. A., HARLAN, J. M., SPARKS, L. H., GAMBLE, J. R., AGOSTI, J. M. & WALTERSDORPH, A. M. 1986. Stimulation of neutrophils by tumor necrosis factor. *The Journal of Immunology,* 136**,** 4220-4225.

KNOPP, S., BJØRGE, T., NESLAND, J. M., TROPÉ, C., SCHEISTRØEN, M. & HOLM, R. 2004. p16INK4a and p21Waf1/Cip1 expression correlates with clinical outcome in vulvar carcinomas. *Gynecologic oncology,* 95**,** 37-45.

KNOPP, S., NESLAND, J. M., TROPÉ, C. & HOLM, R. 2006. p14ARF, a prognostic predictor in HPV-negative vulvar carcinoma. *American journal of clinical pathology,* 126**,** 266-276.

KO, H.-W., LEE, H.-H., HUO, L., XIA, W., YANG, C.-C., HSU, J. L., LI, L.-Y., LAI, C.-C., CHAN, L.-C. & CHENG, C.-C. 2016. GSK3β inactivation promotes the oncogenic functions of EZH2 and enhances methylation of H3K27 in human breast cancers. *Oncotarget,* 7**,** 57131.

KOHRENHAGEN, N., VOELKER, H.-U., KAPP, M., DIETL, J. & KAEMMERER, U. 2010. The expression of galectin-1 in vulvar neoplasia. *Anticancer research,* 30**,** 1547-1552.

KOKKA, F., SINGH, N., FARUQI, A., GIBBON, K. & ROSENTHAL, A. N. 2011a. Is differentiated vulval intraepithelial neoplasia the precursor lesion of human papillomavirus-negative vulval squamous cell carcinoma? *Int J Gynecol Cancer,* 21**,** 1297-305.

KOKKA, F., SINGH, N., FARUQI, A., GIBBON, K. & ROSENTHAL, A. N. 2011b. Is differentiated vulval intraepithelial neoplasia the precursor lesion of human papillomavirus-negative vulval squamous cell carcinoma? *International journal of gynecological cancer,* 21**,** 1297-1305.

KOKKAT, T. J., PATEL, M. S., MCGARVEY, D., LIVOLSI, V. A. & BALOCH, Z. W. 2013. Archived formalin-fixed paraffin-embedded (FFPE) blocks: A valuable underexploited resource for extraction of DNA, RNA, and protein. *Biopreservation and biobanking,* 11**,** 101-106.

KOVACEVIC, Z. & RICHARDSON, D. R. 2006. The metastasis suppressor, Ndrg-1: a new ally in the fight against cancer. *Carcinogenesis,* 27**,** 2355-2366.

KOVACEVIC, Z., SIVAGURUNATHAN, S., MANGS, H., CHIKHANI, S., ZHANG, D. & RICHARDSON, D. R. 2011. The metastasis suppressor, N-myc downstream regulated gene 1 (NDRG1), upregulates p21 via p53-independent mechanisms. *Carcinogenesis,* 32**,** 732-740.

KRAUSS, M., SINGER, H. & HOLLENDER, J. 2010. LC–high resolution MS in environmental analysis: from target screening to the identification of unknowns. *Analytical and bioanalytical chemistry,* 397**,** 943-951.

KRIEGSMANN, M., CASADONTE, R., KRIEGSMANN, J., DIENEMANN, H., SCHIRMACHER, P., KOBARG, J. H., SCHWAMBORN, K., STENZINGER, A., WARTH, A. & WEICHERT, W. 2016. Reliable entity subtyping in non-small cell lung cancer by MALDI imaging mass spectrometry on formalin-fixed paraffin-embedded tissue specimens. *Molecular & Cellular Proteomics***,** mcp. M115. 057513.

KRIEGSMANN, M., WANDERNOTH, P., LISENKO, K., CASADONTE, R., LONGUESPÉE, R., ARENS, N. & KRIEGSMANN, J. 2017. Detection of HPV subtypes by mass spectrometry in FFPE tissue specimens: a reliable tool for routine diagnostics. *Journal of clinical pathology,* 70**,** 417-423.

KUDO, I., ESUMI, M., KUSUMI, Y., FURUSAKA, T. & OSHIMA, T. 2017. Particular gene upregulation and p53 heterogeneous expression in TP53-mutated maxillary carcinoma. *Oncology Letters,* 14**,** 4633-4640.

KYUTOKU, M., TANIYAMA, Y., KATSURAGI, N., SHIMIZU, H., KUNUGIZA, Y., IEKUSHI, K., KOIBUCHI, N., SANADA, F., OSHITA, Y. & MORISHITA, R. 2011. Role of periostin in cancer progression and metastasis: inhibition of breast cancer progression and metastasis by anti-periostin antibody in a murine model. *International journal of molecular medicine,* 28**,** 181-186.

LANCASTER, L. 2010. Gynaecological cancer in Australia. *Australian Journal of Cancer Nursing, The,* 11**,** 5.

LAVORATO-ROCHA, A. M., DE MELO MAIA, B., RODRIGUES, I. S. A., STIEPCICH, M. M. Á., BAIOCCHI, G., DA SILVA CESTARI, F. M., CARVALHO, K. C., SOARES, F. A. & ROCHA, R. M. 2013. Prognostication of vulvar cancer based on p14ARF status: molecular assessment of transcript and protein. *Annals of surgical oncology,* 20**,** 31-39.

LEE, J.-C., CHUNG, L.-C., CHEN, Y.-J., FENG, T.-H. & JUANG, H.-H. 2014. N-myc downstream-regulated gene 1 downregulates cell proliferation, invasiveness, and tumorigenesis in human oral squamous cell carcinoma. *Cancer letters,* 355**,** 242-252.

LEE, M.-Y. & SHEN, M.-R. 2012. Epithelial-mesenchymal transition in cervical carcinoma. *American journal of translational research,* 4**,** 1.

LEE, S. S., LIM, J., TAN, S., CHA, J., YEO, S. Y., AGNEW, H. D. & HEATH, J. R. 2009. Accurate MALDI-TOF/TOF Sequencing of One-Bead− One-Compound Peptide Libraries with Application to the Identification of Multiligand Protein Affinity Agents Using in Situ Click Chemistry Screening. *Analytical chemistry,* 82**,** 672-679.

LEE, Y. Y., WILCZYNSKI, S. P., CHUMAKOV, A., CHIH, D. & KOEFFLER, H. P. 1994. Carcinoma of the vulva: HPV and p53 mutations. *Oncogene,* 9**,** 1655-1659.

LEEMANS, C. R., BRAAKHUIS, B. J. & BRAKENHOFF, R. H. 2011. The molecular biology of head and neck cancer. *Nature Reviews Cancer,* 11**,** 9-22.

LEIBOWTTCH, M., NEILL, S., PELISSE, M. & MOYAL‐BARACCO, M. 1990. The epithelial changes associated with squamous cell carcinoma of the vulva: a review of the clinical, histological and viral findings in 78 women. *BJOG: An International Journal of Obstetrics & Gynaecology,* 97**,** 1135-1139.

LEMAIRE, R., DESMONS, A., TABET, J. C., DAY, R., SALZET, M. & FOURNIER, I. 2007. Direct analysis and MALDI imaging of formalin-fixed, paraffin-embedded tissue sections. *Journal of proteome research,* 6**,** 1295-1305.

LEONARD, S., PEREIRA, M., FOX, R., GORDON, N., YAP, J., KEHOE, S., LUESLEY, D., WOODMAN, C. & GANESAN, R. 2016. Over-expression of DNMT3A predicts the risk of recurrent vulvar squamous cell carcinomas. *Gynecologic oncology,* 143**,** 414-420.

LERMA, E., ESTELLER, M., HERMAN, J. G. & PRAT, J. 2002. Alterations of the p16 INK4a/Rb/cyclin-D1 pathway in vulvar carcinoma, vulvar intraepithelial neoplasia, and lichen sclerosus. *Human pathology,* 33**,** 1120-1125.

LERMA, E., MATIAS-GUIU, X., LEE, S. J. & PRAT, J. 1999. Squamous cell carcinoma of the vulva: study of ploidy, HPV, p53, and pRb. *International journal of gynecological pathology: official journal of the International Society of Gynecological Pathologists,* 18**,** 191-197.

LEUCHTER, R. S., HACKER, N. F., VOET, R. L., BEREK, J. S., TOWNSEND, D. E. & LAGASSE, L. D. 1982. Primary carcinoma of the Bartholin gland: a report of 14 cases and review of the literature. *Obstetrics & Gynecology,* 60**,** 361-368.

LI, W. & TSE, F. L. 2010. Dried blood spot sampling in combination with LC‐MS/MS for quantitative analysis of small molecules. *Biomedical Chromatography,* 24**,** 49-65.

LIEGL, B. & REGAUER, S. 2006. p53 immunostaining in lichen sclerosus is related to ischaemic stress and is not a marker of differentiated vulvar intraepithelial neoplasia (d-VIN). *Histopathology,* 48**,** 268-74.

LIM, H., ENG, J., YATES, J. R., TOLLAKSEN, S. L., GIOMETTI, C. S., HOLDEN, J. F., ADAMS, M. W., REICH, C. I., OLSEN, G. J. & HAYS, L. G. 2003. Identification of 2D-gel proteins: a comparison of MALDI/TOF peptide mass mapping to μ LC-ESI tandem mass spectrometry. *Journal of the American Society for Mass Spectrometry,* 14**,** 957-970.

LIN, T.-W., CHANG, H.-T., CHEN, C.-H., CHEN, C.-H., LIN, S.-W., HSU, T.-L. & WONG, C.-H. 2015. Galectin-3 binding protein and galectin-1 interaction in breast cancer cell aggregation and metastasis. *Journal of the American Chemical Society,* 137**,** 9685-9693.

LIPECKA, J., CHHUON, C., BOURDERIOUX, M., BESSARD, M. A., ENDERT, P., EDELMAN, A. & GUERRERA, I. C. 2016. Sensitivity of mass spectrometry analysis depends on the shape of the filtration unit used for filter aided sample preparation (FASP). *Proteomics,* 16**,** 1852-1857.

LIU, A. Y., ZHENG, H. & OUYANG, G. 2014. Periostin, a multifunctional matricellular protein in inflammatory and tumor microenvironments. *Matrix biology,* 37**,** 150-156.

LIU, N. Q., DEKKER, L. J. M., STINGL, C., GÜZEL, C., DE MARCHI, T., MARTENS, J. W. M., FOEKENS, J. A., LUIDER, T. M. & UMAR, A. 2013. Quantitative Proteomic Analysis of Microdissected Breast Cancer Tissues: Comparison of Label-Free and SILAC-based Quantification with Shotgun, Directed, and Targeted MS Approaches. *Journal of Proteome Research,* 12**,** 4627-4641.

LIU, W., ZHANG, B., HU, Q., QIN, Y., XU, W., SHI, S., LIANG, C., MENG, Q., XIANG, J. & LIANG, D. 2017a. A new facet of NDRG1 in pancreatic ductal adenocarcinoma: suppression of glycolytic metabolism. *International journal of oncology,* 50**,** 1792-1800.

LIU, Y., XUN, X.-H., YI, J.-M., XIANG, Y. & HUA, J. 2017b. Discovery of lung squamous carcinoma biomarkers by profiling the plasma peptide with LC/MS/MS. *Chinese Chemical Letters,* 28**,** 1093-1098.

LOMBARD-BANEK, C., REDDY, S., MOODY, S. A. & NEMES, P. 2016. Label-free quantification of proteins in single embryonic cells with neural fate in the cleavage-stage frog (Xenopus laevis) embryo using capillary electrophoresis electrospray ionization high-resolution mass spectrometry (CE-ESI-HRMS). *Molecular & Cellular Proteomics,* 15**,** 2756-2768.

LONGO, V. D. 2004. Ras: the other pro-aging pathway. *Science's SAGE KE,* 2004**,** pe36.

LONGUESPÉE, R., BOYON, C., DESMONS, A., KERDRAON, O., LEBLANC, E., FARRÉ, I., VINATIER, D., DAY, R., FOURNIER, I. & SALZET, M. 2014. Spectroimmunohistochemistry: A novel form of MALDI mass spectrometry imaging coupled to immunohistochemistry for tracking antibodies. *Omics: a journal of integrative biology,* 18**,** 132-141.

LOU, L., URBANI, J., RIBEIRO-NETO, F. & ALTSCHULER, D. L. 2002. cAMP inhibition of Akt is mediated by activated and phosphorylated Rap1b. *Journal of Biological Chemistry,* 277**,** 32799-32806.

LUDWIG, J. A. & WEINSTEIN, J. N. 2005. Biomarkers in cancer staging, prognosis and treatment selection. *Nature Reviews Cancer,* 5**,** 845.

LV, H., LIU, R., FU, J., YANG, Q., SHI, J., CHEN, P., JI, M., SHI, B. & HOU, P. 2014. Epithelial cell-derived periostin functions as a tumor suppressor in gastric cancer through stabilizing p53 and E-cadherin proteins via the Rb/E2F1/p14ARF/Mdm2 signaling pathway. *Cell Cycle,* 13**,** 2962-2974.

MAGDELDIN, S. & YAMAMOTO, T. 2012. Toward deciphering proteomes of formalin‐fixed paraffin‐embedded (FFPE) tissues. *PROTEOMICS,* 12**,** 1045-1058.

MAKAR, A., SCHEISTROEN, M., VAN DEN WEYNGAERT, D. & TROPE, C. 2001. Surgical management of stage I and II vulvar cancer: the role of the sentinel node biopsy. Review of literature. *International Journal of Gynecological Cancer,* 11**,** 255-262.

MANN, M. 2006. Functional and quantitative proteomics using SILAC. *Nature reviews Molecular cell biology,* 7**,** 952-958.

MARKO-VARGA, G., FEHNIGER, T. E., REZELI, M., DÖME, B., LAURELL, T. & VÉGVÁRI, Á. 2011. Drug localization in different lung cancer phenotypes by MALDI mass spectrometry imaging. *Journal of proteomics,* 74**,** 982-992.

MASIR, N., GHODDOOSI, M., MANSOR, S., ABDUL-RAHMAN, F., FLORENCE, C. S., MOHAMED-ISMAIL, N. A., TAMBY, M. R. & MD-LATAR, N. H. 2012. RCL2, a potential formalin substitute for tissue fixation in routine pathological specimens. *Histopathology,* 60**,** 804-15.

MATUSZEWSKI, B., CONSTANZER, M. & CHAVEZ-ENG, C. 1998. Matrix effect in quantitative LC/MS/MS analyses of biological fluids: a method for determination of finasteride in human plasma at picogram per milliliter concentrations. *Analytical Chemistry,* 70**,** 882-889.

MAYER, C., DARB-ESFAHANI, S., MEYER, A.-S., HÜBNER, K., ROM, J., SOHN, C., BRAICU, I., SEHOULI, J., HÄNSCH, G. M. & GAIDA, M. M. 2016. Neutrophil granulocytes in ovarian cancer-induction of epithelial-to-mesenchymal-transition and tumor cell migration. *Journal of Cancer,* 7**,** 546.

MBEUNKUI, F., METGE, B. J., SHEVDE, L. A. & PANNELL, L. K. 2007. Identification of differentially secreted biomarkers using LC-MS/MS in isogenic cell lines representing a progression of breast cancer. *J Proteome Res,* 6**,** 2993.

MCALPINE, J. N., KIM, S. Y., AKBARI, A., ESHRAGH, S., REUSCHENBACH, M., VON KNEBEL DOEBERITZ, M., PRIGGE, E. S., JORDAN, S., SINGH, N. & MILLER, D. M. 2017. HPV-independent Differentiated Vulvar Intraepithelial Neoplasia (dVIN) is Associated With an Aggressive Clinical Course. *International Journal of Gynecological Pathology,* 36**,** 507-516.

MCDONALD, S., MAGUIRE, G., DUARTE, N., WANG, X. L. & WENDY, H. 2004. C-reactive protein, cardiovascular risk, and renal disease in a remote Australian Aboriginal community. *Clinical Science,* 106**,** 121-128.

MCGRATH, P. & RAWSON, N. 2013. The experience of relocation for specialist treatment for indigenous women diagnosed with vulvar cancer in East Arnhem Land. *Journal of psychosocial oncology,* 31**,** 540-555.

MCGRATH, P., RAWSON, N. & ADIDI, L. 2015. Diagnosis and treatment for vulvar cancer for indigenous women from East Arnhem Land, Northern Territory: Bioethical reflections. *Journal of bioethical inquiry,* 12**,** 343-352.

MCWHIRTER, R. E., THOMSON, R. J., MARTHICK, J. R., RUMBOLD, A. R., BROWN, M. A., TAYLOR-THOMSON, D., MAYPILAMA, E. L., CONDON, J. R. & DICKINSON, J. L. 2014. Runs of homozygosity and a cluster of vulvar cancer in young Australian Aboriginal women. *Gynecologic Oncology,* 133**,** 421-426.

MEDICI, D. & NAWSHAD, A. 2010. Type I collagen promotes epithelial–mesenchymal transition through ILK-dependent activation of NF-κB and LEF-1. *Matrix Biology,* 29**,** 161-165.

MEHRIAN-SHAI, R., CHEN, C. D., SHI, T., HORVATH, S., NELSON, S. F., REICHARDT, J. K. V. & SAWYERS, C. L. 2007. Insulin growth factor-binding protein 2 is a candidate biomarker for PTEN status and PI3K/Akt pathway activation in glioblastoma and prostate cancer. *Proceedings of the National Academy of Sciences,* 104**,** 5563-5568.

MERRILL JR, A. H., SULLARDS, M. C., ALLEGOOD, J. C., KELLY, S. & WANG, E. 2005. Sphingolipidomics: High-throughput, structure-specific, and quantitative analysis of sphingolipids by liquid chromatography tandem mass spectrometry. *Methods,* 36**,** 207-224.

MIRNEZAMI, R., SPAGOU, K., VORKAS, P., LEWIS, M., KINROSS, J., WANT, E., SHION, H., GOLDIN, R., DARZI, A. & TAKATS, Z. 2014. Chemical mapping of the colorectal cancer microenvironment via MALDI imaging mass spectrometry (MALDI‐MSI) reveals novel cancer‐associated field effects. *Molecular oncology,* 8**,** 39-49.

MITTAL, P., KLINGLER-HOFFMANN, M., ARENTZ, G., WINDERBAUM, L., KAUR, G., ANDERSON, L., SCURRY, J., LEUNG, Y., STEWART, C. J. & CARTER, J. 2017. Annexin A2 and alpha actinin 4 expression correlates with metastatic potential of primary endometrial cancer. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics,* 1865**,** 846-857.

MITTAL, P., KLINGLER-HOFFMANN, M., ARENTZ, G., WINDERBAUM, L., KAUR, G., ANDERSON, L., SCURRY, J., LEUNG, Y., STEWART, C. J. R., CARTER, J., HOFFMANN, P. & OEHLER, M. K. Annexin A2 and alpha actinin 4 expression correlates with metastatic potential of primary endometrial cancer. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*.

MITTAL, P., KLINGLER‐HOFFMANN, M., ARENTZ, G., ZHANG, C., KAUR, G., OEHLER, M. K. & HOFFMANN, P. 2016. Proteomics of endometrial cancer diagnosis, treatment, and prognosis. *PROTEOMICS-Clinical Applications,* 10**,** 217-229.

MODESITT, S. C., GROBEN, P. A., WALTON, L. A., FOWLER JR, W. C. & VAN LE, L. 2000a. Expression of Ki-67 in vulvar carcinoma and vulvar intraepithelial neoplasia III: correlation with clinical prognostic factors. *Gynecologic oncology,* 76**,** 51-55.

MODESITT, S. C., GROBEN, P. A., WALTON, L. A., FOWLER, W. C. & VAN LE, L. 2000b. Expression of Ki-67 in vulvar carcinoma and vulvar intraepithelial neoplasia III: correlation with clinical prognostic factors. *Gynecologic oncology,* 76**,** 51-55.

MODESITT, S. C., WATERS, A. B., WALTON, L., FOWLER JR, W. C. & VAN LE, L. 1998. Vulvar intraepithelial neoplasia III: occult cancer and the impact of margin status on recurrence. *Obstetrics & Gynecology,* 92**,** 962-966.

MONK, B. J., BURGER, R. A., LIN, F., PARHAM, G., VASILEV, S. A. & WILCZYNSKI, S. P. 1995. Prognostic significance of human papillomavirus DNA in vulvar carcinoma. *Obstetrics & Gynecology,* 85**,** 709-715.

MOREE, W. J., PHELAN, V. V., WU, C.-H., BANDEIRA, N., CORNETT, D. S., DUGGAN, B. M. & DORRESTEIN, P. C. 2012. Interkingdom metabolic transformations captured by microbial imaging mass spectrometry. *Proceedings of the National Academy of Sciences,* 109**,** 13811-13816.

MORITA, Y., IKEGAMI, K., GOTO-INOUE, N., HAYASAKA, T., ZAIMA, N., TANAKA, H., UEHARA, T., SETOGUCHI, T., SAKAGUCHI, T., IGARASHI, H., SUGIMURA, H., SETOU, M. & KONNO, H. 2010. Imaging mass spectrometry of gastric carcinoma in formalin-fixed paraffin-embedded tissue microarray. *Cancer Sci,* 101**,** 267-73.

MORRA, L. & MOCH, H. 2011. Periostin expression and epithelial-mesenchymal transition in cancer: a review and an update. *Virchows Archiv,* 459**,** 465.

MOUNEIMNE, G., HANSEN, SCOTT D., SELFORS, LAURA M., PETRAK, L., HICKEY, MICHELE M., GALLEGOS, LISA L., SIMPSON, KAYLENE J., LIM, J., GERTLER, FRANK B., HARTWIG, JOHN H., MULLINS, R. D. & BRUGGE, JOAN S. 2012. Differential Remodeling of Actin Cytoskeleton Architecture by Profilin Isoforms Leads to Distinct Effects on Cell Migration and Invasion. *Cancer Cell,* 22**,** 615-630.

MULLER, P. A. & VOUSDEN, K. H. 2013. p53 mutations in cancer. *Nature cell biology,* 15**,** 2-8.

MUÑOZ, N., BOSCH, F. X., DE SANJOSÉ, S., HERRERO, R., CASTELLSAGUÉ, X., SHAH, K. V., SNIJDERS, P. J. & MEIJER, C. J. 2003. Epidemiologic classification of human papillomavirus types associated with cervical cancer. *New England Journal of Medicine,* 348**,** 518-527.

MUNOZ, N., CASTELLSAGUÉ, X., DE GONZÁLEZ, A. B. & GISSMANN, L. 2006. HPV in the etiology of human cancer. *Vaccine,* 24**,** S1-S10.

MURRAY, C. K., HOFFMASTER, R. M., SCHMIT, D. R., HOSPENTHAL, D. R., WARD, J. A., CANCIO, L. C. & WOLF, S. E. 2007. Evaluation of white blood cell count, neutrophil percentage, and elevated temperature as predictors of bloodstream infection in burn patients. *Archives of Surgery,* 142**,** 639-642.

NAGAI, M. A., GERHARD, R., FREGNANI, J. H. T., NONOGAKI, S., RIERGER, R. B., NETTO, M. M. & SOARES, F. A. 2011. Prognostic value of NDRG1 and SPARC protein expression in breast cancer patients. *Breast cancer research and treatment,* 126**,** 1-14.

NAKABAYASHI, K. & HIYAMA, K. 2011. Method of determining lymph node metastasis in cervical cancer, device for determining the same, and computer program. Google Patents.

NAORA, H. & MONTELL, D. J. 2005. Ovarian cancer metastasis: integrating insights from disparate model organisms. *Nature Reviews Cancer,* 5**,** 355.

NEUBERT, H., HALKET, J. M., OCANA, M. F. & PATEL, R. K. 2004. MALDI post-source decay and LIFT-TOF/TOF investigation of α-cyano-4-hydroxycinnamic acid cluster interferences. *Journal of the American Society for Mass Spectrometry,* 15**,** 336-343.

NG, D. H., HUMPHRIES, J. D., SELLEY, J. N., WARWOOD, S., KNIGHT, D., BYRON, A. & HUMPHRIES, M. J. 2017. Improved LC-MS chromatographic alignment increases the accuracy of label-free quantitative proteomics: Comparison of spectral counting versus ion intensity-based proteomic quantification strategies. *bioRxiv***,** 111476.

NICOLAS, E., PARISOT, P., PINTO-MONTEIRO, C., DE WALQUE, R., DE VLEESCHOUWER, C. & LAFONTAINE, D. L. 2016. Involvement of human ribosomal proteins in nucleolar structure and p53-dependent nucleolar stress. *Nature communications,* 7**,** 11390.

NOFECH-MOZES, S., KHALIFA, M. A., ISMIIL, N., SAAD, R. S., HANNA, W. M., COVENS, A. & GHORAB, Z. 2008. Immunophenotyping of serous carcinoma of the female genital tract. *Modern Pathology,* 21**,** 1147.

NORRIS, R. A., DAMON, B., MIRONOV, V., KASYANOV, V., RAMAMURTHI, A., MORENO‐RODRIGUEZ, R., TRUSK, T., POTTS, J. D., GOODWIN, R. L. & DAVIS, J. 2007. Periostin regulates collagen fibrillogenesis and the biomechanical properties of connective tissues. *Journal of cellular biochemistry,* 101**,** 695-711.

O'CONNELL, T. X., NATHAN, L. S., SATMARY, W. A. & GOLDSTEIN, A. T. 2008. Non-neoplastic epithelial disorders of the vulva. *Am Fam Physician,* 77**,** 321-6.

OKA, R., NAKASHIRO, K.-I., GODA, H., IWAMOTO, K., TOKUZEN, N. & HAMAKAWA, H. 2016. Annexin A8 is a novel molecular marker for detecting lymph node metastasis in oral squamous cell carcinoma. *Oncotarget,* 7**,** 4882.

OLAWAIYE, A., LEE, L. M., KRASNER, C. & HOROWITZ, N. 2007. Treatment of squamous cell vulvar cancer with the anti-EGFR tyrosine kinase inhibitor Tarceva. *Gynecologic oncology,* 106**,** 628-630.

OLIVERAS-FERRAROS, C., VAZQUEZ-MARTIN, A., CUYÀS, E., COROMINAS-FAJA, B., RODRÍGUEZ-GALLEGO, E., FERNÁNDEZ-ARROYO, S., MARTIN-CASTILLO, B., JOVEN, J. & MENENDEZ MENENDEZ, J. 2014. Acquired resistance to metformin in breast cancer cells triggers transcriptome reprogramming toward a degradome-related metastatic stem-like profile. *Cell Cycle,* 13**,** 1132-1144.

ONG, S.-E., FOSTER, L. J. & MANN, M. 2003. Mass spectrometric-based approaches in quantitative proteomics. *Methods,* 29**,** 124-130.

OONK, M. H., HOLLEMA, H., DE HULLU, J. A. & VAN DER ZEE, A. G. 2006. Prediction of lymph node metastases in vulvar cancer: a review. *International Journal of Gynecological Cancer,* 16**,** 963-971.

OONK, M. H., VAN HEMEL, B. M., HOLLEMA, H., DE HULLU, J. A., ANSINK, A. C., VERGOTE, I., VERHEIJEN, R. H., MAGGIONI, A., GAARENSTROOM, K. N. & BALDWIN, P. J. 2010a. Size of sentinel-node metastasis and chances of non-sentinel-node involvement and survival in early stage vulvar cancer: results from GROINSS-V, a multicentre observational study. *The lancet oncology,* 11**,** 646-652.

OONK, M. H., VAN HEMEL, B. M., HOLLEMA, H., DE HULLU, J. A., ANSINK, A. C., VERGOTE, I., VERHEIJEN, R. H., MAGGIONI, A., GAARENSTROOM, K. N., BALDWIN, P. J., VAN DORST, E. B., VAN DER VELDEN, J., HERMANS, R. H., VAN DER PUTTEN, H. W., DROUIN, P., RUNNEBAUM, I. B., SLUITER, W. J. & VAN DER ZEE, A. G. 2010b. Size of sentinel-node metastasis and chances of non-sentinel-node involvement and survival in early stage vulvar cancer: results from GROINSS-V, a multicentre observational study. *The Lancet Oncology,* 11**,** 646-652.

ORDI, J., ALEJO, M., FUSTÉ, V., LLOVERAS, B., DEL PINO, M., ALONSO, I. & TORNÉ, A. 2009. HPV-negative vulvar intraepithelial neoplasia (VIN) with basaloid histologic pattern: an unrecognized variant of simplex (differentiated) VIN. *The American journal of surgical pathology,* 33**,** 1659-1665.

ORIGONI, M., SIDERI, M., GARSIA, S., CARINELLI, S. G. & FERRARI, A. G. 1992. Prognostic value of pathological patterns of lymph node positivity in squamous cell carcinoma of the vulva stage III and IVA FIGO. *Gynecologic oncology,* 45**,** 313-316.

OZEKI, C., SAWAI, Y., SHIBATA, T., KOHNO, T., OKAMOTO, K., YOKOTA, J., TASHIRO, F., TANUMA, S., SAKAI, R., KAWASE, T., KITABAYASHI, I., TAYA, Y. & OHKI, R. 2011. Cancer susceptibility polymorphism of p53 at codon 72 affects phosphorylation and degradation of p53 protein. *J Biol Chem,* 286**,** 18251-60.

PAPPA, K. I., JACOB-HIRSCH, J., VLACHOS, G. D., CHRISTODOULOU, I., PARTSINEVELOS, G., AMARIGLIO, N., MARKAKI, S., ANTSAKLIS, A. & ANAGNOU, N. P. 2011. Expression profiling of vulvar carcinoma: clues for deranged extracellular matrix remodeling and effects on multiple signaling pathways combined with discrete patient subsets. *Translational oncology,* 4**,** IN1-IN6.

PAPPA, K. I., RODOLAKIS, A., CHRISTODOULOU, I., GAZOULI, M., MARKAKI, S., ANTSAKLIS, A. & ANAGNOU, N. P. 2014. Comparative assessment of lymph node micrometastasis in cervical, endometrial and vulvar cancer: insights on the real time qRT-PCR approach versus immunohistochemistry, employing dual molecular markers. *BioMed research international,* 2014.

PARK, J.-S., KIM, E.-J., KWON, H.-J., HWANG, E.-S., NAMKOONG, S.-E. & UM, S.-J. 2000. Inactivation of interferon regulatory factor-1 tumor suppressor protein by HPV E7 oncoprotein Implication for the E7-mediated immune evasion mechanism in cervical carcinogenesis. *Journal of Biological Chemistry,* 275**,** 6764-6769.

PARKER, B. L., YANG, G., HUMPHREY, S. J., CHAUDHURI, R., MA, X., PETERMAN, S. & JAMES, D. E. 2015. Targeted phosphoproteomics of insulin signaling using data-independent acquisition mass spectrometry. *Sci. Signal.,* 8**,** rs6-rs6.

PARKIN, D. M. & BRAY, F. 2006. The burden of HPV-related cancers. *Vaccine,* 24**,** S11-S25.

PATEL, R. 2015. MALDI-TOF MS for the diagnosis of infectious diseases. *Clinical chemistry,* 61**,** 100-111.

PECORELLI, S. 2009. Revised FIGO staging for carcinoma of the vulva, cervix, and endometrium. *International Journal of Gynecology & Obstetrics,* 105**,** 103-104.

PERCY, A. J., CHAMBERS, A. G., YANG, J., HARDIE, D. B. & BORCHERS, C. H. 2014. Advances in multiplexed MRM-based protein biomarker quantitation toward clinical utility. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics,* 1844**,** 917-926.

PERRIN, A., GUNDA, M., YU, B., YEN, K., ITO, S., FORSTER, S., TISSENBAUM, H. & DERRY, W. 2013. Noncanonical control of C. elegans germline apoptosis by the insulin/IGF-1 and Ras/MAPK signaling pathways. *Cell death and differentiation,* 20**,** 97.

PETERSLUND, N. A., KOCH, C., JENSENIUS, J. C. & THIEL, S. 2001. Association between deficiency of mannose-binding lectin and severe infections after chemotherapy. *The Lancet,* 358**,** 637-638.

PINTO, A. P., MIRON, A., YASSIN, Y., MONTE, N., WOO, T. Y., MEHRA, K. K., MEDEIROS, F. & CRUM, C. P. 2010. Differentiated vulvar intraepithelial neoplasia contains Tp53 mutations and is genetically linked to vulvar squamous cell carcinoma. *Mod Pathol,* 23**,** 404-12.

PINTO, P., FRÜHAUF, F., ZIKAN, M., CIBULA, D. & FISCHEROVÁ, D. 2016. OC10. 04: The role of ultrasound in lymph node staging in vulvar cancer. *Ultrasound in Obstetrics & Gynecology,* 48**,** 18-18.

POWELL, J. & WOJNAROWSKA, F. 1999. Lichen sclerosus. *The Lancet,* 353**,** 1777-1783.

PRIESKE, K., HAERINGER, N., GRIMM, D., TRILLSCH, F., EULENBURG, C., BURANDT, E., SCHMALFELDT, B., MAHNER, S., MUELLER, V. & WOELBER, L. 2016. Patterns of distant metastases in vulvar cancer. *Gynecologic oncology,* 142**,** 427-434.

PRINCE, M., SIVANANDAN, R., KACZOROWSKI, A., WOLF, G., KAPLAN, M., DALERBA, P., WEISSMAN, I., CLARKE, M. & AILLES, L. 2007. Identification of a subpopulation of cells with cancer stem cell properties in head and neck squamous cell carcinoma. *Proceedings of the National Academy of Sciences,* 104**,** 973-978.

QUDDUS, M. R., SUNG, C. J., SIMON, R. A. & LAWRENCE, W. D. 2014. Localized amyloidosis of the vulva with and without vulvar intraepithelial neoplasia: report of a series. *Human Pathology,* 45**,** 2037-2042.

RAMBERG, H., BRAADLAND, P. R., STENSLAND, M. E., RAKAEE, M., NYGÅRD, S., ØGREN, O., BERGE, V., SVINDLAND, A., RICHARDSSON, E. & DE SOUZA, G. A. 2017. Scavenging for lethal prostate cancer biomarkers in FFPE tissue. AACR.

RAO, J. Y. & LI, N. 2004. Microfilament actin remodeling as a potential target for cancer drug development. *Current cancer drug targets,* 4**,** 345-354.

RASMUSSEN, C. L., SAND, F. L., HOFFMANN FREDERIKSEN, M., KAAE ANDERSEN, K. & KJÆR, S. K. 2018. Does HPV status influence survival after vulvar cancer? *International journal of cancer,* 142**,** 1158-1165.

RASPAGLIESI, F., DITTO, A., FONTANELLI, R., MACCAURO, M., CARCANGIU, M., PARAZZINI, F. & BOMBARDIERI, E. 2003. False‐negative sentinel node in patients with vulvar cancer: A case study. *International Journal of Gynecological Cancer,* 13**,** 361-363.

RATUSHNY, V., GOBER, M. D., HICK, R., RIDKY, T. W. & SEYKORA, J. T. 2012. From keratinocyte to cancer: the pathogenesis and modeling of cutaneous squamous cell carcinoma. *The Journal of clinical investigation,* 122**,** 464-472.

READ, S. A. & DOUGLAS, M. W. 2014. Virus induced inflammation and cancer development. *Cancer letters,* 345**,** 174-181.

REGAUER, S. & REICH, O. 2013. Etiology of vulvar cancer will impact on treatment options and therapy outcome: two major pathways of vulvar cancer. *Gynecologic oncology,* 131**,** 246-247.

ROBISON, K., ROQUE, D., MCCOURT, C., STUCKEY, A., DISILVESTRO, P. A., SUNG, C. J., STEINHOFF, M., GRANAI, C. & MOORE, R. G. 2014. Long-term follow-up of vulvar cancer patients evaluated with sentinel lymph node biopsy alone. *Gynecologic oncology,* 133**,** 416-420.

RODRIGUES, I., LAVORATO-ROCHA, A., DE M MAIA, B., STIEPCICH, M., DE CARVALHO, F., BAIOCCHI, G., SOARES, F. & ROCHA, R. 2013a. Epithelial-mesenchymal transition-like events in vulvar cancer and its relation with HPV. *British journal of cancer,* 109**,** 184-194.

RODRIGUES, L. S., LAVORATO, A., DE MELO MAIA, B., BAIOCCHI, G., STIEPCICH, M., BERTONHA, F. & MALAGOLI, R. 2013b. Expression profiling of β-catenin, E-cadherin and Slug in the development and progression of vulvar carcinoma. *Cancer Research,* 73**,** 1481-1481.

RODRIGUEZ, A. & LAIO, A. 2014. Clustering by fast search and find of density peaks. *Science,* 344**,** 1492-1496.

ROHNER, T. C., STAAB, D. & STOECKLI, M. 2005. MALDI mass spectrometric imaging of biological tissue sections. *Mechanisms of ageing and development,* 126**,** 177-185.

ROSE, R. J., DAMOC, E., DENISOV, E., MAKAROV, A. & HECK, A. J. R. 2012. High-sensitivity Orbitrap mass analysis of intact macromolecular assemblies. *Nat Meth,* 9**,** 1084-1086.

ROSENTHAL, D., DYSON, S., PITTS, M. & GARLAND, S. 2007. Challenges to Accepting a Human Papilloma Virus (HPV) Vaccine: A Qualitative Study of Australian Women. *Women & Health,* 45**,** 59-73.

RÖST, H. L., ROSENBERGER, G., NAVARRO, P., GILLET, L., MILADINOVIĆ, S. M., SCHUBERT, O. T., WOLSKI, W., COLLINS, B. C., MALMSTRÖM, J. & MALMSTRÖM, L. 2014. OpenSWATH enables automated, targeted analysis of data-independent acquisition MS data. *Nature biotechnology,* 32**,** 219-223.

ROTH, B., JAYARATNA, I., SUNDI, D., CHENG, T., MELQUIST, J., CHOI, W., PORTEN, S., NITTI, G., NAVAI, N. & WSZOLEK, M. 2016. Employing an orthotopic model to study the role of epithelial-mesenchymal transition in bladder cancer metastasis. *Oncotarget*.

ROTHSCHILD, B. L., SHIM, A. H., AMMER, A. G., KELLEY, L. C., IRBY, K. B., HEAD, J. A., CHEN, L., VARELLA-GARCIA, M., SACKS, P. G. & FREDERICK, B. 2006. Cortactin overexpression regulates actin-related protein 2/3 complex activity, motility, and invasion in carcinomas with chromosome 11q13 amplification. *Cancer Research,* 66**,** 8017-8025.

ROTTMANN, M., BECK, T., BURGES, A., DANNECKER, C., KIECHLE, M., MAYR, D., SCHLESINGER-RAAB, A., SCHUBERT-FRITSCHLE, G. & ENGEL, J. 2016. Trends in surgery and outcomes of squamous cell vulvar cancer patients over a 16-year period (1998–2013): a population-based analysis. *Journal of cancer research and clinical oncology,* 142**,** 1331-1341.

ROWLEY, K., WALKER, K. Z., COHEN, J., JENKINS, A. J., O NEAL, D., SU, Q., BEST, J. D. & O DEA, K. 2003. Inflammation and vascular endothelial activation in an Aboriginal population: relationships to coronary disease risk factors and nutritional markers. *Medical Journal of Australia,* 178**,** 495-500.

RUAN, K., BAO, S. & OUYANG, G. 2009. The multifaceted role of periostin in tumorigenesis. *Cellular and molecular life sciences,* 66**,** 2219.

RUBIE, C., KEMPF, K., HANS, J., SU, T., TILTON, B., GEORG, T., BRITTNER, B., LUDWIG, B. & SCHILLING, M. 2005. Housekeeping gene variability in normal and cancerous colorectal, pancreatic, esophageal, gastric and hepatic tissues. *Molecular and cellular probes,* 19**,** 101-109.

RUBIN, M. A., DUNN, R., STRAWDERMAN, M. & PIENTA, K. J. 2002. Tissue microarray sampling strategy for prostate cancer biomarker analysis. *The American journal of surgical pathology,* 26**,** 312-319.

RUFFORNY, I., WILKINSON, E. J., LIU, C., ZHU, H., BUTERAL, M. & MASSOLL, N. A. 2005. Human Papillomavirus Infection and p16INK4a Protein Expression in Vulvar Intraepithelial Neoplasia and Invasive Squamous Cell Carcinoma. *Journal of Lower Genital Tract Disease,* 9**,** 108-113.

RUMBOLD, A., TAN, S., CONDON, J., TAYLOR-THOMSON, D., NICKELS, M., TABRIZI, S., DAVY, M., O'BRIEN, M., CONNORS, C., ZARDAWI, I., STANKOVICH, J. & GARLAND, S. 2012a. Investigating a cluster of vulvar cancer in young women: a cross-sectional study of genital human papillomavirus prevalence. *BMC Infectious Diseases,* 12**,** 243.

RUMBOLD, A. R., TAN, S. E., CONDON, J. R., TAYLOR-THOMSON, D., NICKELS, M., TABRIZI, S. N., DAVY, M. L., O'BRIEN, M. M., CONNORS, C. M. & ZARDAWI, I. 2012b. Investigating a cluster of vulvar cancer in young women: a cross-sectional study of genital human papillomavirus prevalence. *BMC infectious diseases,* 12**,** 243.

SALASZNYK, R. M., WILLIAMS, W. A., BOSKEY, A., BATORSKY, A. & PLOPPER, G. E. 2004. Adhesion to vitronectin and collagen I promotes osteogenic differentiation of human mesenchymal stem cells. *BioMed Research International,* 2004**,** 24-34.

SALNIKOW, K., SU, W., BLAGOSKLONNY, M. V. & COSTA, M. 2000. Carcinogenic metals induce hypoxia-inducible factor-stimulated transcription by reactive oxygen species-independent mechanism. *Cancer research,* 60**,** 3375-3378.

SAND, F. L. & THOMSEN, S. F. 2018. Skin diseases of the vulva: inflammatory, erosive-ulcerating and apocrine gland diseases, zinc and vitamin deficiency, vulvodynia and vestibulodynia. *Journal of Obstetrics and Gynaecology,* 38**,** 149-160.

SANDBERG, A., LINDELL, G., KÄLLSTRÖM, B. N., BRANCA, R. M., DANIELSSON, K. G., DAHLBERG, M., LARSON, B., FORSHED, J. & LEHTIÖ, J. 2012. Tumor proteomics by multivariate analysis on individual pathway data for characterization of vulvar cancer phenotypes. *Molecular & Cellular Proteomics,* 11**,** M112. 016998.

SANDBERG, A., LINDELL, G., KÄLLSTRÖM, B. N., BRANCA, R. M., DANIELSSON, K. G., LARSON, B., FORSHED, J. & LEHTIÖ, J. QUANTITATIVE PROTEOME PROFILING OF HUMAN PAPILLOMA VIRUS INFLUENCE ON VULVAR CARCINOMA ON A GLOBAL AND INDIVIDUAL TUMOUR LEVEL.

SANDRIN, T. R., GOLDSTEIN, J. E. & SCHUMAKER, S. 2013. MALDI TOF MS profiling of bacteria at the strain level: a review. *Mass spectrometry reviews,* 32**,** 188-217.

SANO, T., OYAMA, T., KASHIWABARA, K., FUKUDA, T. & NAKAJIMA, T. 1998. Expression status of p16 protein is associated with human papillomavirus oncogenic potential in cervical and genital lesions. *The American journal of pathology,* 153**,** 1741-1748.

SANTOS, M., MONTAGUT, C., MELLADO, B., GARCÍA, Á., Y CAJAL, S. R., CARDESA, A., PUIG-TINTORÉ, L. M. & ORDI, J. 2004. Immunohistochemical staining for p16 and p53 in premalignant and malignant epithelial lesions of the vulva. *International journal of gynecological pathology,* 23**,** 206-214.

SCHMID, M. C. & VARNER, J. A. 2012. Myeloid cells in tumor inflammation. *Vascular cell,* 4**,** 14.

SCHÖNE, C., HÖFLER, H. & WALCH, A. 2013. MALDI imaging mass spectrometry in cancer research: combining proteomic profiling and histological evaluation. *Clinical biochemistry,* 46**,** 539-545.

SCHWARTZ, R. A. 1995. Verrucous carcinoma of the skin and mucosa. *Journal of the American Academy of Dermatology,* 32**,** 1-21.

SCLEROSUS, L. 2008. Non-neoplastic epithelial disorders of the vulva.

SCURRY, J. P. & VANIN, K. 1997. Vulvar squamous cell carcinoma and lichen sclerosus. *Australasian Journal of Dermatology,* 38**,** S20-S25.

SEMENZA, G. L. 2003. Targeting HIF-1 for cancer therapy. *Nature reviews. Cancer,* 3**,** 721.

SETHI, M. K., KIM, H., PARK, C. K., BAKER, M. S., PAIK, Y.-K., PACKER, N. H., HANCOCK, W. S., FANAYAN, S. & THAYSEN-ANDERSEN, M. 2015. In-depth N-glycome profiling of paired colorectal cancer and non-tumorigenic tissues reveals cancer-, stage- and EGFR-specific protein N-glycosylation. *Glycobiology,* 25**,** 1064-1078.

SHACO-LEVY, R., BEAN, S. M., VOLLMER, R. T., PAPALAS, J. A., BENTLEY, R. C., SELIM, M. A. & ROBBOY, S. J. 2010. Paget disease of the vulva: a histologic study of 56 cases correlating pathologic features and disease course. *International Journal of Gynecological Pathology,* 29**,** 69-78.

SHANKAR, J. & NABI, I. R. 2015. Actin cytoskeleton regulation of epithelial mesenchymal transition in metastatic cancer cells. *PloS one,* 10**,** e0119954.
SHEMESH, T., ROWLEY, K., JENKINS, A., BRIMBLECOMBE, J., BEST, J. & O'DEA, K. 2007. Differential association of C-reactive protein with adiposity in men and women in an Aboriginal community in northeast Arnhem Land of Australia. *International journal of obesity,* 31**,** 103.
SHERMAN, K., DALING, J., MCKNIGHT, B. & CHU, J. 1994. Hormonal factors in vulvar cancer. A case-control study. *The Journal of reproductive medicine,* 39**,** 857-861.
SHETA, R., ROUX-DALVAI, F., WOO, C. M., FOURNIER, F., BOURASSA, S., BERTOZZI, C. R., DROIT, A. & BACHVAROV, D. 2016. Proteomic dataset for altered glycoprotein expression upon GALNT3 knockdown in ovarian cancer cells. *Data in Brief,* 8**,** 342-349.
SHIN, N.-R., LEE, Y.-Y., KIM, S.-H., CHOI, C. H., KIM, T.-J., LEE, J.-W., BAE, D.-S. & KIM, B.-G. 2014. Prognostic value of pretreatment hemoglobin level in patients with early cervical cancer. *Obstetrics & gynecology science,* 57**,** 28-36.
SIDERI, M., JONES, R., WILKINSON, E., PRETI, M., HELLER, D., SCURRY, J., HAEFNER, H. & NEILL, S. 2005. Squamous vulvar intraepithelial neoplasia: 2004 modified terminology, ISSVD Vulvar Oncology Subcommittee. *The Journal of reproductive medicine,* 50**,** 807-810.
SIMARD, J. C., GIRARD, D. & TESSIER, P. A. 2010. Induction of neutrophil degranulation by S100A9 via a MAPK‐dependent mechanism. *Journal of leukocyte biology,* 87**,** 905-914.
SKAPA, P., ZAMECNIK, J., HAMSIKOVA, E., SALAKOVA, M., SMAHELOVA, J., JANDOVA, K., ROBOVA, H., ROB, L. & TACHEZY, R. 2007. Human papillomavirus (HPV) profiles of vulvar lesions: possible implications for the classification of vulvar squamous cell carcinoma precursors and for the efficacy of prophylactic HPV vaccination. *The American journal of surgical pathology,* 31**,** 1834-1843.
SMALL, J. V., ROTTNER, K., KAVERINA, I. & ANDERSON, K. 1998. Assembling an actin cytoskeleton for cell attachment and movement. *Biochimica Et Biophysica Acta (BBA)-Molecular Cell Research,* 1404**,** 271-281.
SMITH, R. D., ANDERSON, G. A., LIPTON, M. S., PASA‐TOLIC, L., SHEN, Y., CONRADS, T. P., VEENSTRA, T. D. & UDSETH, H. R. 2002. An accurate mass tag strategy for quantitative and high‐throughput proteome measurements. *PROTEOMICS,* 2**,** 513-523.
SMOLDERS, J., SCHUURMAN, K. G., VAN STRIEN, M. E., MELIEF, J., HENDRICKX, D., HOL, E. M., VAN EDEN, C., LUCHETTI, S. & HUITINGA, I. 2013. Expression of Vitamin D Receptor and Metabolizing Enzymes in Multiple Sclerosis—Affected Brain Tissue. *Journal of Neuropathology & Experimental Neurology,* 72**,** 91-105.
SOGAWA, K., WATANABE, M., SATO, K., SEGAWA, S., ISHII, C., MIYABE, A., MURATA, S., SAITO, T. & NOMURA, F. 2011. Use of the MALDI BioTyper system with MALDI–TOF mass spectrometry for rapid identification of microorganisms. *Analytical and bioanalytical chemistry,* 400**,** 1905-1911.
SONG, Y. S., LEE, H. W., PARK, Y. R., KIM, D. K., SIM, J., KANG, H. P. & KIM, J. H. 2010. TMA-TAB: A spreadsheet-based document for exchange of tissue microarray data based on the tissue microarray-object model. *Journal of Biomedical Informatics,* 43**,** 435-441.
STELLATO, C., NASSA, G., TARALLO, R., GIURATO, G., RAVO, M., RIZZO, F., MARCHESE, G., ALEXANDROVA, E., CORDELLA, A. & BAUMANN, M. 2015. Identification of cytoplasmic proteins interacting with unliganded estrogen receptor α and β in human breast cancer cells. *Proteomics,* 15**,** 1801-1807.
STOECKLI, M., STAAB, D. & SCHWEITZER, A. 2007. Compound and metabolite distribution measured by MALDI mass spectrometric imaging in whole-body tissue sections. *International Journal of Mass Spectrometry,* 260**,** 195-202.
STROUP, A., HARLAN, L. & TRIMBLE, E. 2008. Demographic, clinical, and treatment trends among women diagnosed with vulvar cancer in the United States. *Gynecologic oncology,* 108**,** 577-583.

SUCKAU, D., RESEMANN, A., SCHUERENBERG, M., HUFNAGEL, P., FRANZEN, J. & HOLLE, A. 2003. A novel MALDI LIFT-TOF/TOF mass spectrometer for proteomics. *Analytical and bioanalytical chemistry,* 376**,** 952-965.

SULLIVAN, C. A. & CHUNG, G. G. 2008. Biomarker validation: in situ analysis of protein expression using semiquantitative immunohistochemistry-based techniques. *Clinical colorectal cancer,* 7**,** 172-177.

SUN, S., LIANG, X., ZHANG, X., LIU, T., SHI, Q., SONG, Y., JIANG, Y., WU, H., LU, X. & PANG, D. 2015. Phosphoglycerate kinase-1 is a predictor of poor survival and a novel prognostic biomarker of chemoresistance to paclitaxel treatment in breast cancer. *British journal of cancer,* 112**,** 1332-1339.

TAMARA, S. O., YANSHOLE, L. V., YANSHOLE, V. V., FURSOVA, A. Z., STEPAKOV, D. A., NOVOSELOV, V. P. & TSENTALOVICH, Y. P. 2016. Spatial distribution of metabolites in the human lens. *Experimental eye research,* 143**,** 68-74.

TANAKA, T., BAI, Z., SRINOULPRASERT, Y., YANG, B., HAYASAKA, H. & MIYASAKA, M. 2005. Chemokines in tumor progression and metastasis. *Cancer science,* 96**,** 317-322.

TANG, Q., LIU, Y., LI, T., YANG, X., ZHENG, G., CHEN, H., JIA, L. & SHAO, J. 2016. A novel co-drug of aspirin and ursolic acid interrupts adhesion, invasion and migration of cancer cells to vascular endothelium via regulating EMT and EGFR-mediated signaling pathways: multiple targets for cancer metastasis prevention and treatment. *Oncotarget,* 7**,** 73114.

TANIGUCHI, K. & KARIN, M. IL-6 and related cytokines as the critical lynchpins between inflammation and cancer. Seminars in immunology, 2014. Elsevier, 54-74.

TANTIPALAKORN, C., ROBERTSON, G., MARSDEN, D. E., GEBSKI, V. & HACKER, N. F. 2009. Outcome and patterns of recurrence for International Federation of Gynecology and Obstetrics (FIGO) stages I and II squamous cell vulvar cancer. *Obstetrics & Gynecology,* 113**,** 895-901.

TE GROOTENHUIS, N. C., VAN DER ZEE, A. G. J., VAN DOORN, H. C., VAN DER VELDEN, J., VERGOTE, I., ZANAGNOLO, V., BALDWIN, P. J., GAARENSTROOM, K. N., VAN DORST, E. B., TRUM, J. W., SLANGEN, B. F. M., RUNNEBAUM, I. B., TAMUSSINO, K., HERMANS, R. H., PROVENCHER, D. M., DE BOCK, G. H., DE HULLU, J. A. & OONK, M. H. M. Sentinel nodes in vulvar cancer: Long-term follow-up of the GROningen INternational Study on Sentinel nodes in Vulvar cancer (GROINSS-V) I. *Gynecologic Oncology,* 140**,** 8-14.

THAKER, N. G., KLOPP, A. H., JHINGRAN, A., FRUMOVITZ, M., IYER, R. B. & EIFEL, P. J. 2015. Survival outcomes for patients with stage IVB vulvar cancer with grossly positive pelvic lymph nodes: Time to reconsider the FIGO staging system? *Gynecologic oncology,* 136**,** 269-273.

TRIETSCH, M. 2017. *Vulvar squamous cell carcinoma: genetics, morphology and clinical behaviour.*

TRIETSCH, M. D., NOOIJ, L. S., GAARENSTROOM, K. N. & VAN POELGEEST, M. I. 2015. Genetic and epigenetic changes in vulvar squamous cell carcinoma and its precursor lesions: A review of the current literature. *Gynecologic oncology,* 136**,** 143-157.

TRIETSCH, M. D., OONK, M., HAWINKELS, L., BOR, R., VAN EENDENBURG, J., IVANOVA, Z., PETERS, A., NIJMAN, H. W., GAARENSTROOM, K. N. & BOSSE, T. 2016. Prognostic value and clinicopathologic characteristics of L1 cell adhesion molecule (L1CAM) in a large series of vulvar squamous cell carcinomas. *Oncotarget*.

TRIETSCH, M. D., SPAANS, V. M., TER HAAR, N. T., OSSE, E. M., PETERS, A. A. W., GAARENSTROOM, K. N. & FLEUREN, G. J. 2014. CDKN2A(p16) and HRAS are frequently mutated in vulvar squamous cell carcinoma. *Gynecologic Oncology,* 135**,** 149-155.

TRINK, B., OSADA, M., RATOVITSKI, E. A. & SIDRANSKY, D. 2007. p63 transcriptional regulation of epithelial integrity and cancer. *Cell Cycle,* 6**,** 240-245.

TSAI, J. H., DONAHER, J. L., MURPHY, D. A., CHAU, S. & YANG, J. 2012. Spatiotemporal regulation of epithelial-mesenchymal transition is essential for squamous cell carcinoma metastasis. *Cancer cell,* 22**,** 725-736.

TURNER, M. W., DINAN, L., HEATLEY, S., JACK, D. L., BOETTCHER, B., LESTER, S., MCCLUSKEY, J. & ROBERTON, D. 2000. Restricted polymorphism of the mannose-binding lectin gene of indigenous Australians. *Human molecular genetics,* 9**,** 1481-1486.

URESHINO, H., MURAKAMI, Y., WATARI, K., IZUMI, H., KAWAHARA, A., KAGE, M., ARAO, T., NISHIO, K., YANAGIHARA, K. & KINOSHITA, H. 2012. N-myc downstream regulated gene 1 (NDRG1) promotes metastasis of human scirrhous gastric cancer cells through epithelial mesenchymal transition. *PloS one,* 7**,** e41312.

VAGULIENE, N., ZEMAITIS, M., LAVINSKIENE, S., MILIAUSKAS, S. & SAKALAUSKAS, R. 2013. Local and systemic neutrophilic inflammation in patients with lung cancer and chronic obstructive pulmonary disease. *BMC immunology,* 14**,** 36.

VALERY, P. C., COORY, M., STIRLING, J. & GREEN, A. C. 2006. Cancer diagnosis, treatment, and survival in Indigenous and non-Indigenous Australians: a matched cohort study. *The Lancet,* 367**,** 1842-1848.

VÄLIKANGAS, T., SUOMI, T. & ELO, L. L. 2017. A comprehensive evaluation of popular proteomics software workflows for label-free proteome quantification and imputation. *Briefings in Bioinformatics*.

VAN DE NIEUWENHOF, H. P., BULTEN, J., HOLLEMA, H., DOMMERHOLT, R. G., MASSUGER, L. F., VAN DER ZEE, A. G., DE HULLU, J. A. & VAN KEMPEN, L. C. 2011. Differentiated vulvar intraepithelial neoplasia is often found in lesions, previously diagnosed as lichen sclerosus, which have progressed to vulvar squamous cell carcinoma. *Mod Pathol,* 24**,** 297-305.

VAN DE NIEUWENHOF, H. P., MASSUGER, L. F., VAN DER AVOORT, I. A., BEKKERS, R. L., CASPARIE, M., ABMA, W., VAN KEMPEN, L. C. & DE HULLU, J. A. 2009a. Vulvar squamous cell carcinoma development after diagnosis of VIN increases with age. *Eur J Cancer,* 45**,** 851-6.

VAN DE NIEUWENHOF, H. P., MASSUGER, L. F., VAN DER AVOORT, I. A., BEKKERS, R. L., CASPARIE, M., ABMA, W., VAN KEMPEN, L. C. & DE HULLU, J. A. 2009b. Vulvar squamous cell carcinoma development after diagnosis of VIN increases with age. *European Journal of Cancer,* 45**,** 851-856.

VAN DE NIEUWENHOF, H. P., VAN KEMPEN, L. C., DE HULLU, J. A., BEKKERS, R. L., BULTEN, J., MELCHERS, W. J. & MASSUGER, L. F. 2009c. The etiologic role of HPV in vulvar squamous cell carcinoma fine tuned. *Cancer Epidemiology and Prevention Biomarkers,* 18**,** 2061-2067.

VAN DEN EINDEN, L. C., VAN DER AVOORT, I. A. & DE HULLU, J. A. 2013. Prevention, identification and treatment of vulvar squamous (pre) malignancies: a review focusing on quality of care. *Expert review of anticancer therapy,* 13**,** 845-859.

VAN DER STEEN, S., VAN DE NIEUWENHOF, H., MASSUGER, L., BULTEN, J. & DE HULLU, J. 2010. New FIGO staging system of vulvar cancer indeed provides a better reflection of prognosis. *Gynecologic oncology,* 119**,** 520-525.

VAN DER ZEE, A. G., OONK, M. H., DE HULLU, J. A., ANSINK, A. C., VERGOTE, I., VERHEIJEN, R. H., MAGGIONI, A., GAARENSTROOM, K. N., BALDWIN, P. J. & VAN DORST, E. B. 2008. Sentinel node dissection is safe in the treatment of early-stage vulvar cancer. *Journal of clinical oncology,* 26**,** 884-889.

VAN HOLST PELLEKAAN, S. M., INGMAN, M., ROBERTS‐THOMSON, J. & HARDING, R. M. 2006. Mitochondrial genomics identifies major haplogroups in Aboriginal Australians. *American journal of physical anthropology,* 131**,** 282-294.

VANIN, K., SCURRY, J., THORNE, H., YUEN, K. & RAMSAY, R. G. 2002. Overexpression of wild-type p53 in lichen sclerosus adjacent to human papillomavirus-negative vulvar cancer. *Journal of investigative dermatology,* 119**,** 1027-1033.

VASCA, V., VASCA, E., FREIMAN, P., MARIAN, D., LUCE, A., MESOLELLA, M., CARAGLIA, M., RICCIARDIELLO, F. & DUMINICA, T. 2014. Keratin 5 expression in squamocellular carcinoma of the head and neck. *Oncology letters,* 8**,** 2501-2504.

VILLANUEVA, J., PHILIP, J., ENTENBERG, D., CHAPARRO, C. A., TANWAR, M. K., HOLLAND, E. C. & TEMPST, P. 2004. Serum peptide profiling by magnetic particle-assisted, automated sample processing and MALDI-TOF mass spectrometry. *Analytical chemistry,* 76**,** 1560-1570.

VITORINO, R., CALHEIROS‐LOBO, M. J., DUARTE, J. A., DOMINGUES, P. M. & AMADO, F. M. 2008. Peptide profile of human acquired enamel pellicle using MALDI tandem MS. *Journal of separation science,* 31**,** 523-537.

WAN, L., PANTEL, K. & KANG, Y. 2013. Tumor metastasis: moving new biological insights into the clinic. *Nature medicine,* 19**,** 1450-1464.

WANG, D.-Z., LI, C., XIE, Z.-X., DONG, H.-P., LIN, L. & HONG, H.-S. 2011. Homology-driven proteomics of dinoflagellates with unsequenced genomes using MALDI-TOF/TOF and automated de novo sequencing. *Evidence-based complementary and alternative medicine,* 2011.

WANG, W., EDDY, R. & CONDEELIS, J. 2007. The cofilin pathway in breast cancer invasion and metastasis. *Nature Reviews Cancer,* 7**,** 429.

WANG, X.-M., LI, J., YAN, M.-X., LIU, L., JIA, D.-S., GENG, Q., LIN, H.-C., HE, X.-H., LI, J.-J. & YAO, M. 2013. Integrative analyses identify osteopontin, LAMB3 and ITGB1 as critical pro-metastatic genes for lung cancer. *PLoS One,* 8**,** e55714.

WATKINS, J. C., HOWITT, B. E., HOROWITZ, N. S., RITTERHOUSE, L. L., DONG, F., MACCONAILL, L. E., GARCIA, E., LINDEMAN, N. I., LEE, L. J. & BERKOWITZ, R. S. 2016. Differentiated exophytic vulvar intraepithelial lesions are genetically distinct from keratinizing squamous cell carcinomas and contain mutations in PIK3CA. *Modern Pathology*.

WEBER, G. F. 2016. Metabolism in cancer metastasis. *International Journal of Cancer,* 138**,** 2061-2066.

WICKI, A., LEHEMBRE, F., WICK, N., HANTUSCH, B., KERJASCHKI, D. & CHRISTOFORI, G. 2006. Tumor invasion in the absence of epithelial-mesenchymal transition: podoplanin-mediated remodeling of the actin cytoskeleton. *Cancer cell,* 9**,** 261-272.

WINDERBAUM, L., KOCH, I., MITTAL, P. & HOFFMANN, P. 2016. Classification of MALDI‐MS imaging data of tissue microarrays using canonical correlation analysis‐based variable selection. *Proteomics,* 16**,** 1731-1735.

WINTER, S. C., BUFFA, F. M., SILVA, P., MILLER, C., VALENTINE, H. R., TURLEY, H., SHAH, K. A., COX, G. J., CORBRIDGE, R. J. & HOMER, J. J. 2007. Relation of a hypoxia metagene derived from head and neck cancer to prognosis of multiple cancers. *Cancer research,* 67**,** 3441-3449.

WISNIEWSKI, J. R., ZOUGMAN, A., NAGARAJ, N. & MANN, M. 2009. Universal sample preparation method for proteome analysis. *Nat Meth,* 6**,** 359-362.

WISZTORSKI, M., FRANCK, J., SALZET, M. & FOURNIER, I. 2010. MALDI direct analysis and imaging of frozen versus FFPE tissues: what strategy for which sample? *Mass Spectrometry Imaging: Principles and Protocols***,** 303-322.

WOELBER, L., EULENBURG, C., CHOSCHZICK, M., KRUELL, A., PETERSEN, C., GIESEKING, F., JAENICKE, F. & MAHNER, S. 2012. Prognostic role of lymph node metastases in vulvar cancer and implications for adjuvant treatment. *International Journal of Gynecological Cancer,* 22**,** 503-508.

WOELBER, L., KOCK, L., GIESEKING, F., PETERSEN, C., TRILLSCH, F., CHOSCHZICK, M., JAENICKE, F. & MAHNER, S. 2011. Clinical management of primary vulvar cancer. *European Journal of Cancer,* 47**,** 2315-2321.

WOJAKOWSKA, A., PIASECKA, A., GARCÍA-LÓPEZ, P. M., ZAMORA-NATERA, F., KRAJEWSKI, P., MARCZAK, Ł., KACHLICKI, P. & STOBIECKI, M. 2013. Structural analysis and profiling of phenolic secondary metabolites of Mexican lupine species using LC–MS techniques. *Phytochemistry,* 92**,** 71-86.

WONG, G., LEE, J., PARK, Y., KLEIN-SZANTO, A., WALDRON, T., CUKIERMAN, E., HERLYN, M., GIMOTTY, P., NAKAGAWA, H. & RUSTGI, A. 2013. Periostin cooperates with mutant p53 to mediate invasion through the induction of STAT1 signaling in the esophageal tumor microenvironment. *Oncogenesis,* 2**,** e59.

WRIGHT, H. L., MOOTS, R. J., BUCKNALL, R. C. & EDWARDS, S. W. 2010. Neutrophil function in inflammation and inflammatory diseases. *Rheumatology,* 49**,** 1618-1631.

WU, S.-L., KIM, J., BANDLE, R. W., LIOTTA, L., PETRICOIN, E. & KARGER, B. L. 2006. Dynamic profiling of the post-translational modifications and interaction partners of epidermal growth factor receptor signaling after stimulation by epidermal growth factor using Extended Range Proteomic Analysis (ERPA). *Molecular & Cellular Proteomics,* 5**,** 1610-1627.

XIAO, Y., GUO, L. & WANG, Y. 2014. A targeted quantitative proteomics strategy for global kinome profiling of cancer cells and tissues. *Molecular & Cellular Proteomics,* 13**,** 1065-1075.

YAMAGUCHI, H. & CONDEELIS, J. 2007. Regulation of the actin cytoskeleton in cancer cell migration and invasion. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research,* 1773**,** 642-652.

YAMAOKA, Y., KINUGASA, Y., SHIOMI, A., YAMAGUCHI, T., KAGAWA, H., YAMAKAWA, Y., FURUTANI, A. & NUMATA, M. 2016. Macroscopic Diagnosis of Lymph Node Metastasis by Size and Palpation in Colorectal Cancer. *Digestive Surgery,* 33**,** 85.

YAMAZAKI, D., KURISU, S. & TAKENAWA, T. 2005. Regulation of cancer cell motility through actin reorganization. *Cancer science,* 96**,** 379-386.

YUAN, M., BREITKOPF, S. B., YANG, X. & ASARA, J. M. 2012. A positive/negative ion-switching, targeted mass spectrometry-based metabolomics platform for bodily fluids, cells, and fresh and fixed tissue. *Nature protocols,* 7**,** 872-881.

ZANNONI, G. F., PRISCO, M. G., VELLONE, V. G., DE STEFANO, I., SCAMBIA, G. & GALLO, D. 2011. Changes in the expression of oestrogen receptors and E‐cadherin as molecular markers of progression from normal epithelium to invasive cancer in elderly patients with vulvar squamous cell carcinoma. *Histopathology,* 58**,** 265-275.

ZAVALIN, A., YANG, J., HAYDEN, K., VESTAL, M. & CAPRIOLI, R. M. 2015. Tissue protein imaging at 1 μm laser spot diameter for high spatial resolution and high imaging speed using transmission geometry MALDI TOF MS. *Analytical and bioanalytical chemistry,* 407**,** 2337-2342.

ZHANG, C., ARENTZ, G., WINDERBAUM, L., LOKMAN, N. A., KLINGLER-HOFFMANN, M., MITTAL, P., CARTER, C., OEHLER, M. K. & HOFFMANN, P. 2016a. MALDI mass spectrometry imaging reveals decreased CK5 levels in vulvar squamous cell carcinomas compared to the precursor lesion differentiated vulvar intraepithelial neoplasia. *International Journal of Molecular Sciences,* 17**,** 1088.

ZHANG, C., ARENTZ, G., WINDERBAUM, L., LOKMAN, N. A., KLINGLER-HOFFMANN, M., MITTAL, P., CARTER, C., OEHLER, M. K. & HOFFMANN, P. 2016b. MALDI Mass Spectrometry Imaging Reveals Decreased CK5 Levels in Vulvar Squamous Cell Carcinomas Compared to the Precursor Lesion Differentiated Vulvar Intraepithelial Neoplasia. *Int J Mol Sci,* 17.

ZHAO, G., CHEN, J., DENG, Y., GAO, F., ZHU, J., FENG, Z., LV, X. & ZHAO, Z. 2011. Identification of NDRG1-regulated genes associated with invasive potential in cervical and ovarian cancer cells. *Biochemical and biophysical research communications,* 408**,** 154-159.

ZHOU, W., FAN, M.-Y., WEI, Y.-X., HUANG, S., CHEN, J.-Y. & LIU, P. 2016. The expression of MYH9 in osteosarcoma and its effect on the migration and invasion abilities of tumor cell. *Asian Pacific journal of tropical medicine,* 9**,** 597-600.

ZHOU, X.-M., ZHANG, H. & HAN, X. 2014. Role of epithelial to mesenchymal transition proteins in gynecological cancers: pathological and therapeutic perspectives. *Tumor Biology,* 35**,** 9523-9530.

ZHU, H., WU, T. C., CHEN, W. Q., ZHOU, L. J., WU, Y., ZENG, L. & PEI, H. P. 2013. Roles of galectin‐7 and S100A9 in cervical squamous carcinoma: Clinicopathological and in vitro evidence. *International journal of cancer,* 132**,** 1051-1059.

ZHU, W., SMITH, J. W. & HUANG, C.-M. 2009. Mass spectrometry-based label-free quantitative proteomics. *BioMed research international,* 2010.

SUPPLEMENTARY

Supplementary files:

https://1drv.ms/f/s!Ar0ZNtdxtPpC-eh1t3I7XT0YaosX0g

Supplementary S2.1

Supplementary S2.2

Supplementary S2.2.4.1

Supplementary S2.2.4.2

Supplementary S2.2.4.3

Supplementary S2.2.4.4

Supplementary S2.2.4.5

Supplementary S2.2.4.6

Supplementary S2.2.4.7

Supplementary S2.2.4.8

Supplementary S3.1

Supplementary S3.2

Supplementary S3.3

Supplementary S3.4

Supplementary S4.1

Supplementary S4.2

Supplementary S4.3

Supplementary S4.4

Supplementary S4.5

Supplementary S4.6

Supplementary S4.7

Supplementary S4.8

Supplementary S5.1

Supplementary S5.2

Supplementary S5.3

Supplementary S5.4

Supplementary S5.5

Supplementary S5.6

APPENDIX

Appendix I: MALDI Mass Spectrometry Imaging Reveals Decreased CK5 Levels in Vulvar Squamous Cell Carcinomas Compared to the Precursor Lesion Differentiated Vulvar Intraepithelial Neoplasia

APPENDIX I: MALDI MASS SPECTROMETRY IMAGING REVEALS DECREASED CK5 LEVELS IN VULVAR SQUAMOUS CELL CARCINOMAS COMPARED TO THE PRECURSOR LESION DIFFERENTIATED VULVAR INTRAEPITHELIAL NEOPLASIA

Statement of Authorship

Title of Paper	MALDI Mass Spectrometry Imaging Reveals Decreased CK5 Levels in Vulvar Squamous Cell Carcinomas Compared to the Precursor Lesion Differentiated Vulvar Intraepithelial Neoplasia
Publication Status	☑ Published ☐ Accepted for Publication ☐ Submitted for Publication ☐ Unpublished and Unsubmitted w ork w ritten in manuscript style
Publication Details	Zhang, C., Arentz, G., Winderbaum, L., Lokman, N.A., Klingler-Hoffmann, M., Mittal, P., Carter, C., Oehler, M.K. and Hoffmann, P., 2016. MALDI mass spectrometry imaging reveals decreased CK5 levels in vulvar squamous cell carcinomas compared to the precursor lesion differentiated vulvar intraepithelial neoplasia. International journal of molecular sciences, 17(7), p.1088.

Principal Author

Name of Principal Author (Candidate)	Chao ZHANG
Contribution to the Paper	Performed MALDI and Immunochemistry analysis on all samples, interpreted data and wrote the first draft of the manuscript
Overall percentage (%)	30%
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.
Signature	Date: 5/12/2017

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Georgia Arentz		
Contribution to the Paper	Equally contributed author, LC-MS data acquisition, supervised development of work, data interpretation and manuscript evaluation		
Signature		Date	20/12/2017

Name of Co-Author	Lyron Winderbaum		
Contribution to the Paper	Equally contributed author, bioinformatics analysis, data interpretation and manuscript evaluation		
Signature		Date	12/12/2017

Name of Co-Author	Noor A. Lokman		
Contribution to the Paper	Immunohistochemstry support, data interpretation and manuscript evaluation		
Signature		Date	12/11/2017

Name of Co-Author	Manuela Klingler-Hoffmann		
Contribution to the Paper	Data interpretation and manuscript evaluation		
Signature		Date	21-12-17

Name of Co-Author	Parul Mittal
Contribution to the Paper	Participated in experimental design
Signature	Date 05/12/2017

Name of Co-Author	Christopher Carter
Contribution to the Paper	Provide tissue samples and histopathological analysis
Signature	Date 13/12/2017

Name of Co-Author	Martin K. Oehler
Contribution to the Paper	Provide tissue samples, ethics approval, pathological report, supervised development of work and manuscript evaluation
Signature	Date 6/2/18

Name of Co-Author	Peter Hoffmann
Contribution to the Paper	Supervised development of work, data interpretation, manuscript evaluation and acted as corresponding auther.
Signature	Date 05/12/17

International Journal of
Molecular Sciences

MDPI

Article

MALDI Mass Spectrometry Imaging Reveals Decreased CK5 Levels in Vulvar Squamous Cell Carcinomas Compared to the Precursor Lesion Differentiated Vulvar Intraepithelial Neoplasia

Chao Zhang [1,2,†], Georgia Arentz [1,2,†], Lyron Winderbaum [1,†], Noor A. Lokman [1,3], Manuela Klingler-Hoffmann [1,2], Parul Mittal [1,2], Christopher Carter [4], Martin K. Oehler [3,5] and Peter Hoffmann [1,2,*]

1 Adelaide Proteomics Centre, School of Biological Sciences, The University of Adelaide, Adelaide 5005, Australia; chao.zhang01@student.adelaide.edu.au (C.Z.); georgia.arentz@adelaide.edu.au (G.A.); lyron.winderbaum@student.adelaide.edu.au (L.W.); noor.lokman@adelaide.edu.au (N.A.L.); manuela.klinglerhoffmann@adelaide.edu.au (M.K.-H.); parul.mittal@adelaide.edu.au (P.M.)
2 Institute for Photonics and Advanced Sensing (IPAS), The University of Adelaide, Adelaide 5005, Australia
3 Discipline of Obstetrics and Gynaecology, School of Medicine, Research Centre for Reproductive Health, Robinson Institute, Adelaide 5005, Australia; martin.oehler@adelaide.edu.au
4 Department of Cytopathology, SA Pathology, Adelaide 5005, Australia; Christopher.Carter@sa.gov.au
5 Department of Gynaecological Oncology, Royal Adelaide Hospital, Adelaide 5005, Australia
* Correspondence: Peter.hoffmann@adelaide.edu.au; Tel.: +61-08-8313-4903; Fax: +61-08-0-8313-4362
† These authors contributed equally to this work.

Academic Editors: Jamal Zweit and Sundaresan Gobalakrishnan
Received: 10 May 2016; Accepted: 30 June 2016; Published: 8 July 2016

Abstract: Vulvar cancer is the fourth most common gynecological cancer worldwide. However, limited studies have been completed on the molecular characterization of vulvar squamous cell carcinoma resulting in a poor understanding of the disease initiation and progression. Analysis and early detection of the precursor lesion of HPV-independent vulvar squamous cell carcinoma (VSCC), differentiated vulvar intraepithelial neoplasia (dVIN), is of great importance given dVIN lesions have a high level of malignant potential. Here we present an examination of adjacent normal vulvar epithelium, dVIN, and VSCC from six patients by peptide Matrix-assisted laser desorption/ionization Mass Spectrometry Imaging (MALDI-MSI). The results reveal the differential expression of multiple peptides from the protein cytokeratin 5 (CK5) across the three vulvar tissue types. The difference observed in the relative abundance of CK5 by MALDI-MSI between the healthy epithelium, dVIN, and VSCC was further analyzed by immunohistochemistry (IHC) in tissue from eight VSCC patients. A decrease in CK5 immunostaining was observed in the VSCC compared to the healthy epithelium and dVIN. These results provide an insight into the molecular fingerprint of the vulvar intraepithelial neoplasia that appears to be more closely related to the healthy epithelium than the VSCC.

Keywords: Matrix-assisted laser desorption/ionization (MALDI); imaging mass spectrometry (MSI); differentiated vulvar intraepithelial neoplasia (dVIN); vulva squamous cell carcinomas (VSCC); Cytokeratin 5 (CK5)

1. Introduction

Vulvar cancer is the fourth most common gynecological cancer worldwide, constituting 5% of all gynecological cancers and 0.6% of all cancer cases in women [1]. In 2015, 5150 new cases were diagnosed in the United States with 1080 recorded deaths [1]. The majority of vulvar cancers are squamous cell carcinomas (VSCC), of which there are two etiological pathways: one linked to

human papilloma virus (HPV) infection, and the other HPV-independent with genetic alterations. The HPV related and non-HPV related forms of vulvar cancer have distinct precursor lesions, known as usual-type vulvar intraepithelial neoplasia and differentiated vulvar intraepithelial neoplasia, respectively. Mutations in the tumor suppressor gene TP16 are associated with HPV related VSCC, while mutations in TP53 have been shown to promote the development of non-HPV related VSCC from the inflammatory condition lichen sclerosus [2].

Recently, the incidence rates of non-HPV related invasive vulvar cancer and its precursor lesion, differentiated vulvar intraepithelial neoplasia (dVIN), have significantly increased amongst young women [3,4]. In Australia, the incidence rates of the disease are extremely high among Aboriginal women from Arnhem Land in the Northern Territory. Studies conducted within this cohort have found no relation between the disease and HPV infection or any underlying genomic irregularities that differentiate this group from other VSCC patients [3,4].

For many years the clinicopathological significance of vulvar intraepithelial neoplasia (VIN) as a precursor to vulvar squamous cell carcinoma (VSCC) was questioned [5] due to the low detection rates of dVIN as a solitary lesion [6]. However, in recent times, dVIN has been confirmed as the direct precursor lesion to HPV-independent vulvar carcinoma, which in the past may have been frequently misdiagnosed as lichen sclerosus or other forms of dermatoses [7,8]. Such misdiagnoses were in part due to numerous changes over the last few decades with regards to the terminology and criteria used to differentiate lichen sclerosus from dVIN and other inflammatory conditions such as squamous cell hyperplasia and lichen simplex chronicus [7,9,10].

The detection of dVIN as a solitary lesion is of great importance given the lesions have a high level of malignant potential [6], with studies showing the median time between the diagnoses of dVIN to VSCC development around 23 to 28 months [6,7]. Mutation of TP53 is a known early event in the development of dVIN [11] and the detection of dVIN adjacent to VSCC is high at around 80%. However, the detection of dVIN as a solitary lesion is scarce [6] and can be difficult to diagnose due to the subtle histological appearance of the lesion [12,13]. Currently the characterization of the dVIN at the molecular level has been poor due to sample rarity, hence the understanding of VSCC progression from healthy tissue to the precursor lesion is low compared to other well studied carcinomas.

MALDI-MSI allows the in situ analysis of tissue sections combining classical mass spectrometry with histological tissue analysis [14]. MALDI-MSI can identify peptides and other molecules from samples whilst retaining the important spatial information of the tissue. Previous studies in the gastrointestinal tract, lung, brain and other gynaecological cancers have shown that MALDI-MSI can acquire a more comprehensive proteomic picture than immunohistochemical testing [14–17]. Here we present an examination of adjacent normal vulvar epithelium, dVIN, and VSCC from six patients by peptide MALDI-MSI. All patients included in this study were HPV negative and lichen sclerosus positive. The results revealed the differential detection of multiple peptides from the protein cytokeratin 5 (CK5) across the three tissue types. The difference observed in the relative abundance of CK5 by MALDI-MSI between the healthy epithelium, dVIN, and VSCC was further investigated by IHC in eight tissues from VSCC patients. This revealed a decrease in CK5 immunostaining in the VSCC as compared to the healthy epithelium and dVIN.

2. Results

2.1. Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry Imaging (MALDI-MSI)

MALDI-MSI was carried out on tissue sections containing regions of healthy vulvar epithelium, dVIN, and VSCC from six late stage carcinoma patients. Peak groups, representing the imaged peptides across all of the analyzed tissue sections with the same m/z (assuming a mass accuracy of ± 0.02 Da), were detected using density-based clustering of peaks (DBSCAN* [18] with an epsilon of 0.02 Da and a minimum of 100 points) and the stringent criterion of requiring peak groups to contain a minimum of 10,000 peaks. This stringent criterion of a minimum of 10,000 peaks per peak group was initially

put in place to ensure that the peak groups analyzed could be reproducibly detected across a majority of spectra. With this criteria in place, 31 peak groups were detected (Table 1). When the minimum number of peaks per peak group was lowered to 1000, 316 peak groups were detected. For each peak group the abundance weighted mean (AWM) was calculated, representing the apex of the peak group and the m/z of the most intense peptide within the group. The widths of the 31 detected peak groups ranged from $\pm$0.075 to $\pm$0.9 Da. The maximum difference, d, in median log intensity between the three tissue regions was calculated for each of the 31 peak groups of interest across the patient cohort (Table 1). Of the 31 peak groups, 19 were found to have a difference in fold change intensity of $\geqslant$1.4-fold across the three tissue types.

Table 1. Matrix-assisted laser desorption/ionization mass spectrometry imaging peak groups ranked heuristically by the largest difference in median log intensity between the healthy epithelium, differentiated vulvar intraepithelial neoplasia, and vulvar squamous cell carcinoma.

Peptide AWM [M + H] of the Peak Group [1]	Da Range of the Peak Group	No. Spectra in Peak Group [2]	Maximum Difference Median Log Intensity [3]
1198.71	$\pm$0.62	23,099	0.75
1028.62	$\pm$0.56	15,354	0.66
1410.72 *	$\pm$0.82	18,784	0.58
1905.95	$\pm$0.52	10,956	0.56
1095.58	$\pm$0.82	21,704	0.55
944.54	$\pm$0.69	19,354	0.53
957.63	$\pm$0.65	10,464	0.52
857.60	$\pm$0.68	14,956	0.50
976.51	$\pm$0.65	15,534	0.49
878.56	$\pm$0.68	11,213	0.46
810.48 *	$\pm$0.76	17,095	0.45
958.59	$\pm$0.69	15,432	0.42
943.62	$\pm$0.82	20,030	0.41
865.42 *	$\pm$0.84	15,617	0.41
1217.66	$\pm$0.55	10,906	0.39
884.52 *	$\pm$0.66	14,337	0.37
856.67	$\pm$0.74	15,483	0.36
1143.65 *	$\pm$0.64	17,685	0.36
1045.62	$\pm$0.65	16,370	0.35
2147.19	$\pm$0.14	20,418	0.33
1105.59	$\pm$0.50	11,479	0.30
1296.68	$\pm$0.10	22,823	0.29
870.64	$\pm$0.85	22,749	0.29
842.60	$\pm$0.90	25,210	0.22
1267.65	$\pm$0.32	11,686	0.20
2932.59	$\pm$0.07	15,500	0.19
816.50	$\pm$0.73	10,983	0.19
971.56 *	$\pm$0.65	15,998	0.18
1111.58	$\pm$0.51	13,473	0.09
1570.70	$\pm$0.17	12,029	0.05
864.63	$\pm$0.70	11,114	0.05

[1] The abundance weighted means (AWM) m/z was calculated for the overall peak group; [2] Number of Spectra in each peak group compiled from all acquisition spectra across the patient cohort; [3] Peak groups were heuristically ranked based on the maximum difference, d, in median log intensity between the tissue regions of interest; * CK5 peptides.

2.2. Cytokeratin 5 (CK5) Identified as a Protein of Interest

In order to gain peptide identifications for the MALDI-MSI peak groups of interest nanoflow liquid chromatography tandem mass spectrometry (nano-LC-MS/MS) was performed on laser microdissected regions of the healthy epithelium, dVIN, and VSCC. Matching between the MALDI-MSI peak groups and nano-LC-MS/MS data was done by aligning the experimental m/z values of the

sequenced peptides that fell between the minimum and maximum m/z of each of the MALDI-MSI peak groups of interest. A table containing all of the matching results for the 31 peak groups of interest is provided in Table S2. Of the 31 MALDI-MSI peak groups of interest, six matched to sequenced peptides from Cytokeratin 5 (CK5) (Table 2), hence CK5 was selected for further analysis. One of the peak groups (AWM [M + H] of 1410.72) matched to two unique CK5 peptides that share the same mass to within 0.009 Da, SFSTASAITPSVSR (1409.7203) and TTAENEFVMLKK (1409.7295). A MALDI-MSI annotated ion intensity map for the peak group m/z 1410.67, which matches to the CK5 peptides SFSTASAITPSVSR and TTAENEFVMLKK (both with a m/z 1410.67), is shown in Figure 1. With regards to the remaining results in Table 1, only three other proteins were detected with more than one unique peptide match as shown in Table S2. Four peptides were detected from the protein Neuroblast differentiation-associated protein AHNAK, three peptides were detected from the protein Annexin A4, and two peptides were detected from the protein Nicotinate phosphoribosyltransferase.

Table 2. CK5 matrix-assisted laser desorption/ionization mass spectrometry imaging peak groups of interest with matched nanoflow liquid chromatography tandem mass spectrometry results.

Peptide AWM [M + H] [1]	*d* [2]	LC-MS/MS *m/z* [3]	LC-MS/MS MR [4]	Peptide Score [5]	Peptide Significance [6]	Peptide Sequence [7]
810.48	0.45	405.7087	809.4028	21.44	0.0072	QSSVSFR
865.42	0.41	433.1998	864.3851	26	0.0025	SGGGGGGGFGR
884.52	0.37	442.7286	883.4427	24.74	0.0034	TSFTSVSR
971.56	0.18	486.2429	970.4712	25.1	0.0031	FVSTTSSSR
1143.65	0.36	572.3175	1142.6205	39.59	0.00011	LAELEEALQK
1410.72	0.58	705.8674	1409.7203	115.28	3×10^{-12}	SFSTASAITPSVSR
1410.72	0.58	705.872	1409.7295	150.55	8.8×10^{-16}	TTAENEFVMLKK

[1] The peptide abundance weighted mean (AWM) m/z calculated for each peak group; [2] Heuristically ranked maximum difference, d, between the healthy epithelium, dVIN, and VSCC; [3] Experimental m/z of the CK5 matched nano-LC-MS/MS detected peptide; [4] Experimental mass of the CK5 matched nano-LC-MS/MS detected peptide; [5] Mascot Ions Score; [6] Mascot Expected Score; [7] CK5 matched peptide sequence.

2.3. Immunohistochemistry (IHC) Analysis of CK5 across the Healthy Epithelium, Differentiated Vulvar Intraepithelial Neoplasia and Vulvar Squamous Cell Carcinoma

The differential detection of CK5 across the three tissue types as determined by MALDI-MSI was further investigated by IHC in tissue from eight VSCC patients. Of these eight tissues, three had been previously analyzed in the MALDI-MSI experiments. IHC analysis of CK5 revealed a significant difference in the level of staining between the dVIN (mean 91.6 $\pm$ 6.1) and VSCC (mean 75.4 $\pm$ 10.7) tissues ($p = 1.758 \times 10^{-7}$). A similar result was observed for the healthy as compared to VSCC tissue ($p = 0.004$), with increased staining observed in the healthy epithelium (mean 86.7 $\pm$ 14.8). Overall, there was no significant difference in the level of staining between the dVIN and healthy epithelium ($p = 0.14$), although the mean staining intensity in the dVIN was slightly higher. Results for the IHC analysis and CK5 staining from two representative vulvar tissues are provided in Figure 2. A direct comparison of the CK5 peptides SFSTASAITPSVSR and TTAENEFVMLKK, both m/z 1410.72, as detected by MALDI-MSI to the detection of CK5 by IHC on consecutive vulvar tissue sections is shown in Figure 3.

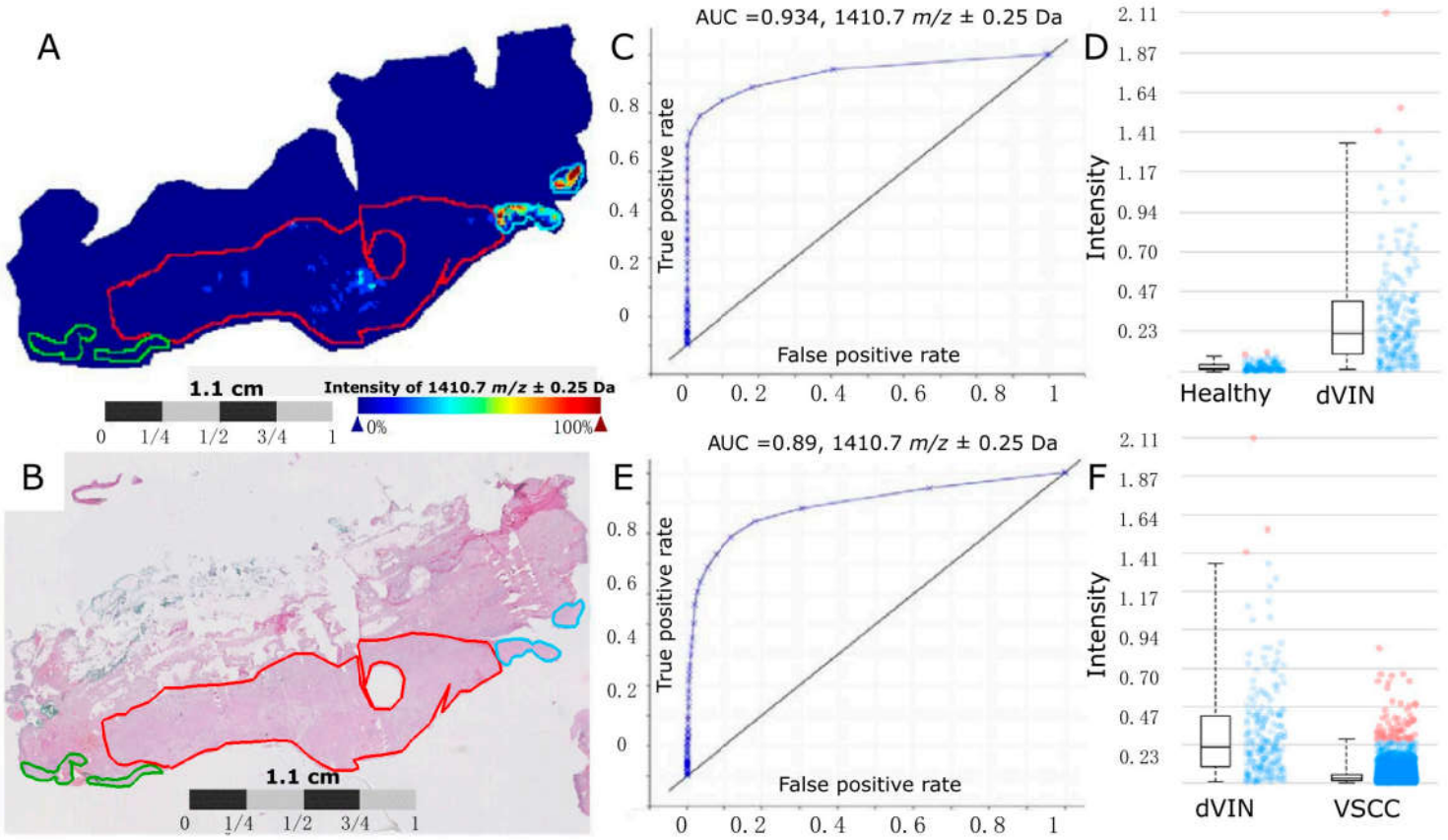


Figure 1. A MALDI-MSI annotated ion intensity map from a representative vulvar tissue section for the peak group *m*/*z* 1410.67. (**A**) An ion intensity map for *m*/*z* 1410.67, which matches to the CK5 peptides SFSTASAITPSVSR and TTAENEFVMLKK (both with a [M + H] *m*/*z* 1410.67). The regions of healthy epithelium are outlined in green, the dVIN is outlined in blue, and the VSCC is outlined in red; (**B**) H & E stain of Healthy (green), dVIN (blue) and VSCC (red) tissue post MALDI-MSI data acquisition; (**C**) A ROC curve and (**D**) boxplot analysis of the MALDI-MSI data for *m*/*z* 1410.67 comparing the annotated healthy epithelium to the dVIN region, AUC = 0.934; (**E**) A ROC curve and (**F**) boxplot analysis of the MALDI-MSI data for *m*/*z* 1410.67 comparing the annotated dVIN to the VSCC region, AUC = 0.89. Blue dots represent the *m*/*z* 1410.67 spectra intensity interval between the lower and upper quartiles. Red dots represent the spectra intensities outside of the intervals.

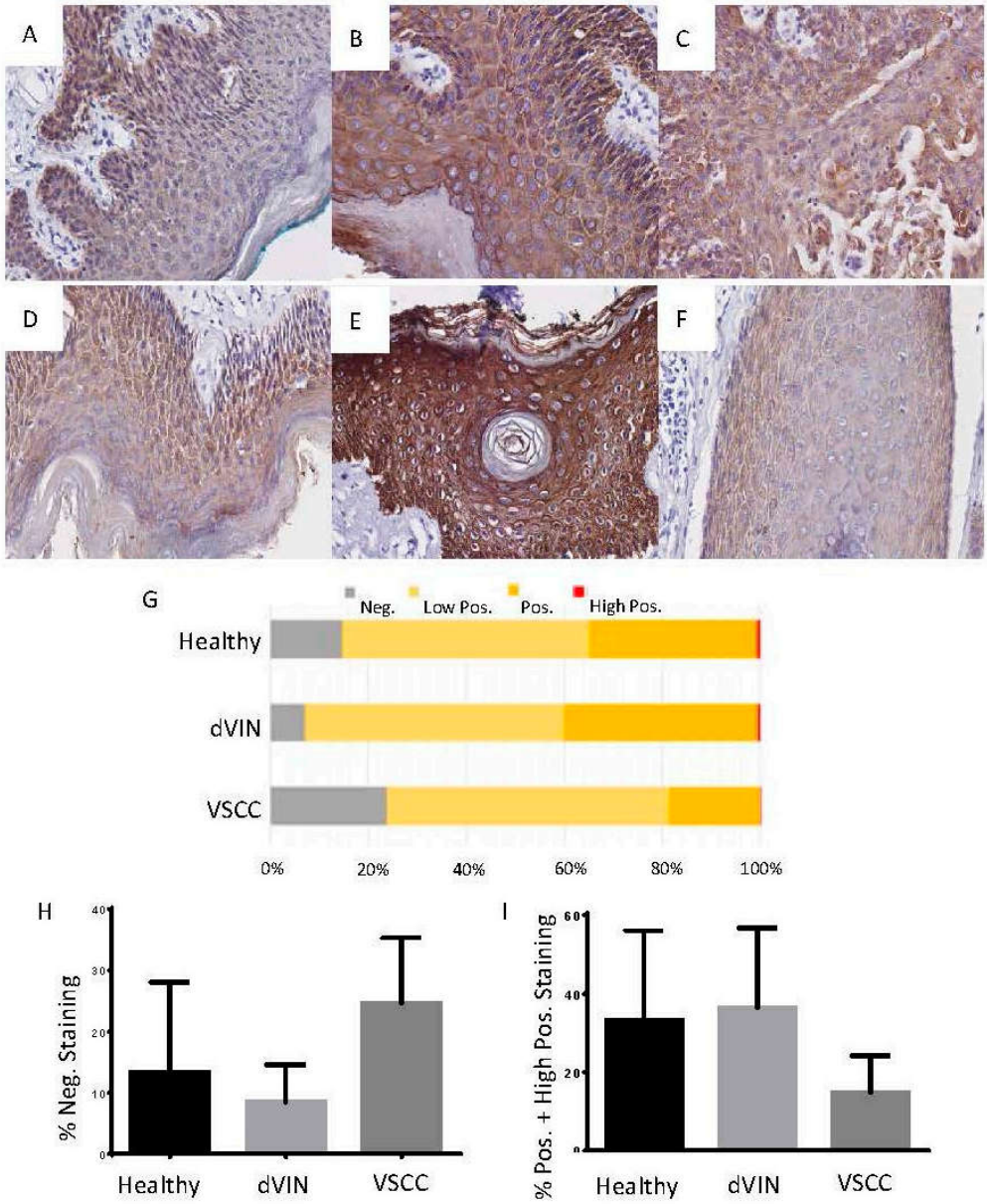


Figure 2. IHC analysis of CK5 on vulvar tissues containing regions of healthy epithelium, dVIN, and VSCC, all at 40× magnification. (**A**) CK5 staining of healthy epithelium; (**B**) dVIN; (**C**) VSCC from a single section; (**D**) CK5 staining of healthy epithelium; (**E**) dVIN; (**F**) VSCC from a single section of a different patient; (**G**) Quantitative analysis of CK5 staining across the tissue types was performed using IHC profiler-Image J. For each tissue section, three representative photo-micrographic images at 40× magnification were used and each image was assigned a staining score of negative, low positive, positive, and high positive; (**H**) The percent of negative staining was significantly different between the dVIN and the VSCC ($p < 0.0001$), but not between the dVIN and healthy epithelium ($p = 0.11$); (**I**) The percent of positive and high positive staining was significantly different between the dVIN and the VSCC ($p< 0.0001$), but not between the dVIN and healthy epithelium ($p = 0.59$).

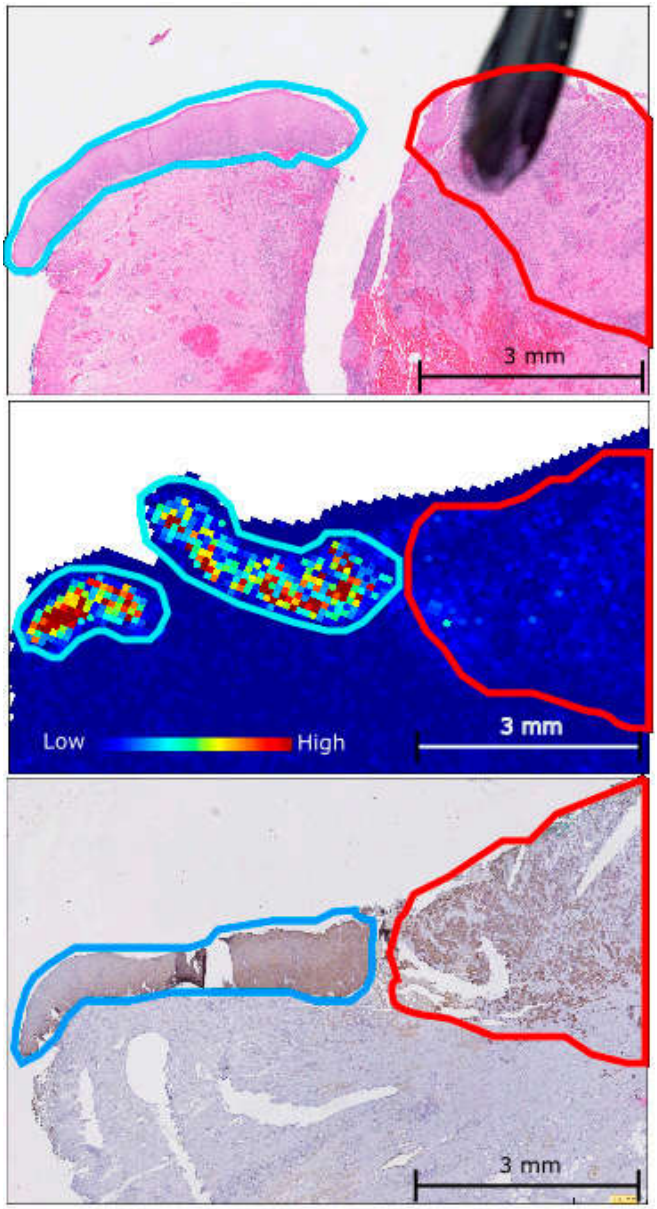


Figure 3. A comparison of the expression pattern observed for the CK5 peptides SFSTASAITPSVSR and TTAENEFVMLKK, both m/z 1410.72, in the dVIN (blue) and VSCC (red) using: MALDI-MSI ion intensity map (**middle**); H & E staining (**top**); and CK5 IHC staining (**bottom**) on consecutively cut vulvar tissue sections.

3. Discussion

Due to the relative rarity of vulvar cancer, characterization of the disease at the molecular level has been poor. A greater understanding of vulvar cancer progression is required, however, especially with regards to dVIN as the precursor lesion is known to have a high level of malignant potential [6] and is difficult to diagnose [12,13]. Generally, dVIN is characterized by thickened parakeratotic epithelium with elongation and anastomosing rete ridges, little to no atypia above the basal or parabasal layers of the epidermis, with the basal cell layer containing abnormal squamous cells undergoing atypical mitosis and high levels of epithelial cell differentiation in the upper epithelium [13,19,20]. The aim of this study was to in situ characterize the dVIN lesions at the peptide level using MALDI-MSI for

comparison to healthy vulvar epithelium and VSCC. To the best of our knowledge, this is the first report performing such an analysis from FFPE vulvar cancer tissues.

MALDI-MSI was carried out on tissue sections containing regions of healthy vulvar epithelium, dVIN, and VSCC from six late stage carcinoma patients. Thirty-one peak groups, representing the in situ imaged peptides, could be reproducibly detected across the majority of acquired spectra (Table 1). The median log intensity for each of the 31 peak groups was calculated for the regions of healthy epithelium, dVIN, and VSCC, and the maximum difference in intensity between the tissue types was calculated. This revealed 19 of the peak groups to have a change in intensity of $\geqslant$1.4-fold across the three tissue types, of which six matched to unique peptides from the protein CK5 (Table 2).

CK5 is a 58 kDa intermediate filament protein that dimerizes with cytokeratin 14 to form the cytoskeleton of basal epithelial cells [21]. CK5 is known to have high levels of expression in squamous cell carcinomas and can be used as a marker of differentiation, with weaker to no expression in adenocarcinomas [22]. Given squamous cell carcinomas account for over 90% of all vulvar cancers and that CK5 is a structural protein of basal epithelial cells [23], the detection of CK5 by both MALDI-MSI was to be expected. A study characterizing the expression of CK5 in serous gynecological carcinomas of the ovaries, endometrium, fallopian tube, primary peritoneum, and cervix observed the highest level of CK5 expression in the cervical and ovarian serous carcinomas, with little to no reactivity in the endometrial, fallopian tube, or primary peritoneum serous carcinomas [24].

In order to determine if there was a genuine difference in CK5 levels across the three tissue types, IHC staining was performed on vulvar tissue from eight VSCC patients. Of these eight tissues, three had been previously analyzed in the MALDI-MSI experiments, six were moderately differentiated, one was well differentiated, and one was poorly differentiated. A degree of keratinization was reported for all of the VSCC with the exception of the poorly differentiated tumor and one of the moderately differentiated tumors. Generally, the expression of CK5 appeared to be diffuse and cytosolic across the three tissue types, with no distinct tissue specific staining patterns observed. IHC analysis revealed a significant increase in the level of CK5 staining in the dVIN as compared to the VSCC legions, and in the healthy epithelium as compared to the VSCC legions (Figure 2). There was a slight increase in the level of CK5 staining detected in the dVIN as compared to the healthy epithelium, but the difference was not significant.

To verify the spatial expression pattern of the peptides detected by MALDI-MSI, consecutive tissue sections that had been H&E stained and analysed by CK5 IHC were compared to ion intensity maps generated for the CK5 peptides from the MSI experiments. A representative example of H&E staining, CK5 IHC, and an ion intensity map for the CK5 peptides SFSTASAITPSVSR and TTAENEFVMLKK, both m/z 1410.72, from consecutive vulvar tissue sections is shown in Figure 3. The intensity and spatial distribution of the CK5 peptides as detected by MALDI-MSI corroborated well with the CK5 immunostaining observed. These results demonstrate the power of MALDI-MSI in the ability to combine classical mass spectrometry with histological tissue analysis.

This study provides an insight into the molecular fingerprint of the VSCC precursor lesion dVIN, which appears to more closely resemble healthy epithelium than is does cancerous tissue, even in later stage patients. To our knowledge, this is the first MALDI imaging mass spectrometry analysis of vulva cancer.

4. Materials and Methods

4.1. Sample Cohort

Formalin-fixed paraffin-embedded (FFPE) samples diagnosed with advanced VSCC from 2001 to 2013 were retrieved from the archive of SA Pathology at the Royal Adelaide Hospital (RAH). Ethics approval for the study was granted by the RAH. The age of patients ranged from 37 to 83 years with median age of 66.5 years, all patients were HPV negative, and each tissue analyzed contained regions of healthy vulvar epithelium, dVIN, and invasive vulvar squamous carcinoma (VSCC). Nine of the

vulvar carcinomas analyzed were stage III and three were stage IB. Clinical information for the patients is provided in Table S1. All tissues analyzed in this study were annotated by a pathologist.

4.2. MALDI-MSI Preparation and Acquisition

MALDI-MSI of FFPE tissue sections was performed on tissue from 6 VSCC patients as previously described [25]. Briefly, tissues were sectioned 8 µm thick and mounted on the Indium-Tin-Oxide (ITO) coated glass slides (Bruker Daltonics, Bremen, Germany) by heating at 60 °C for 1 h. Tissues were deparaffinised in xylene for 5 min, following by two 2 min incubations in 100% ethanol, and two 5 min incubations in 100 mM NH_4HCO_3. The tissues were then subjected to heat induced citric acid antigen retrieval (CAAR) (10 mM citric acid, pH = 6) [26] followed by digestion with trypsin gold (Promega, Madison, WI, USA) using an ImagePrep station (Bruker Daltonics, Bremen, Germany) at 37 °C for 2 h. Peptide Internal calibrants (Angiotensin I, [Glu1]-Fibrinopeptide B, Dynorphin A and ACTH fragment [1–24]) and α-cyano-4-hydroxycinnamic acid (CHCA) matrix was overlaid onto the tissue sections using an ImagePrep station (Bruker Daltonics, Bremen, Germany) [27]. An ultrafleXtreme MALDI-TOF/TOF MS system (Bruker Daltonics, Bremen, Germany) was used for data acquisition in positive reflectron mode as monitored by flexControl (V3.0.1 Bruker Daltonics, Bremen, Germany) using following settings: 2 kHz, m/z 800–4000, with 100 µm spatial resolution. Following data acquisition slides were Heamatoxylin and Eosin (H & E) stained and scanned using a NanoZommer (Hamamatsu, Japan). The H & E stained images were then co-registered with the collected MALDI MSI spectra using the flexImaging software (V4.0.1 Bruker Daltonics, Bremen, Germany).

4.3. MALDI-MSI Data Analysis

Acquired spectra were recalibrated using the internal calibrants and pre-processed using flexAnalysis and flexImaging software (V4.0.1 Bruker Daltonics, Bremen, Germany) in order to produce peak lists. During the pre-processing the Snap Algorithm with a signal to noise ratio of 2 was used for peak detection, a baseline subtraction was performed using the TopHat method, and baseline smoothing was performed using the Gauss algorithm. Tissue specific regions of interest (ROI) (i.e., healthy epithelium, dVIN and VSCC) were selected and exported into flexImaging (Bruker, Germany). Peaklists from all ROI were combined and density based clustering (DBSCAN* [18], with an epsilon of 0.02 and minimum density of 100 peaks) was used to cluster peaks into peak groups based on their m/z values. Initially only peak groups containing at least 10,000 peaks were considered for further analysis. This strict criterion was put in place to ensure that all peak groups analyzed could be reproducibly detected in many spectra. The median log intensity in each of the 3 tissue types was calculated for each peak group and for each patient. Peak groups were heuristically ranked by the largest difference in median log intensity between tissue types. The raw data was also analyzed using the SCiLS lab software (SCiLS, GmbH, Bremen, Germany, 2015b) where the processing steps of baseline removal and normalization were carried out [28]. Ion intensity maps for the peptides of interest were generated using the SCiLS lab software.

4.4. Peptide Identification by Nanoflow Liquid Chromatography Tandem Mass Spectrometry (nano-LC-MS/MS)

Tissue areas of healthy, dVIN, and VSCC were collected using laser capture microdissection (LCM) for protein digestion and identification by nano-LC-MS/MS. The FFPE tissue was sectioned at 8 µm thickness, water bath mounted onto PEN membrane slides (Micro-Dissect, Herborn, Germany), and deparaffinised as described above. Areas of healthy epithelium, dVIN, and VSCC were dissected using a Leica AS LCM microscope (Leica Microsystems, Wetzlar, Germany) into 20 µL of 10 mM citric acid buffer (pH = 6) and subjected to heat induced antigen retrieval as described above. Samples were buffered with 10 mM of NH_4HCO_3 and digested with 100 ng of trypsin gold (Promega, Madison, WI, USA) overnight at 37 °C.

Nano-LC-MS/MS was performed using an Ultimate 3000 RSLC system (Thermo-Fisher Scientific, MA, USA) coupled to an Impact II™ QTOF mass spectrometer (Bruker Daltonics, Billerica, MA, USA)

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 min. 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 min, followed by a 20 min wash with 90% B, and a 20 min 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 DataAnalysis (Version 4.2, 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 was searched, the digestion enzyme was specified as trypsin with up to 2 missed cleavages, variable modification of oxidation of methionine, MS mass tolerance of 40 ppm, and MS/MS mass tolerance of 0.2 Da. In Mascot, the peptide false discovery rate was set to <1% using Percolator, the peptide ion score cut off was set to 20, and the peptide significance score was set to <0.05.

4.5. Matching the MALDI-MSI Peak Groups to the Nanoflow Liquid Chromatography Tandem Mass Spectrometry (Nano-LC-MS/MS) Results

Matching between the two data sets was completed manually by comparing the experimental m/z values of the nano-LC-MS/MS sequenced peptides that fell between the minimum and maximum m/z of each of the MALDI-MSI peak groups.

4.6. Immunohistochemistry

For the analysis of Cytokeratin 5 by IHC, tissue from 8 VSCC patients was analyzed as previously described [29]. Briefly, 8 µm tissue sections were placed on plain glass slides, de-waxed, rehydrated and subjected to microwave antigen retrieval in 10 mM citric acid buffer (pH = 6) for 10 min at 100 °C in a steam microwave (Sixth Sense, Whirlpool, VIC, Australia). Tissue sections were then incubated overnight at 4 °C with the Cytokeratin 5 antibody (rabbit monoclonal antibody at a 1/200 dilution, Abcam, Cambridge, MA, USA) in 5% goat serum blocking buffer, followed by incubation with biotinylated anti-rabbit immunoglobulin (1/400, Dako, North Sydney, NSW, Australia) and streptavidin-HRP (1/500, Dako, Australia). Diaminobenzidine (DAB)/H_2O_2 (Sigma Aldrich) substrate was added before counterstaining with haematoxylin (Sigma Aldrich), dehydrating and mounting in Pertex (Medite Medizintechnik, Germany). The CK5 stained sections were scanned using a NanoZommer (Hamamatsu, Japan) and viewed in NDP View (Hamamatsu, Japan). For each tissue section, three representative photo-micrographic images from areas of healthy epithelium, dVIN, and VSCC were captured at 40× magnification. Analysis was carried out using the IHC Profiler in ImageJ which determines areas of high positive, positive, low positive, and negative staining [30]. The levels of CK5 positive staining for each tissue type were summed and compared across patients using GraphPad Prism 6 v008 (GraphPad Software, La Jolla, CA, USA), where the mean, standard error of the mean, and significance (paired t-test) were calculated.

Supplementary Materials: Supplementary materials can be found at http://www.mdpi.com/1422-0067/17/7/1088/s1.

Acknowledgments: The authors gratefully acknowledge the financial support of Bioplatforms Australia and the Government of South Australia. The authors thank Chris Cursaro for outstanding technical support. We would like to acknowledge Johan O.R. Gustafsson for supervising the MALDI-MSI and experiments and providing the experiment protocol and Inge Koch (School of Mathematical Science, University of Adelaide) for mass spectra statistics analysis and classification.

Author Contributions: Martin K. Oehler, Christopher Carter, and Peter Hoffmann contributed to study design and sample collection; Chao Zhang, Noor A. Lokman, Parul Mittal, and Georgia Arentz contributed to the sample preparation and execution of experiments; Lyron Winderbaum performed the bioinformatics analysis; and Georgia Arentz, Martin K. Oehler, Peter Hoffmann and Manuela Klingler-Hoffmann wrote the paper. All authors read and approved the final manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Siegel, R.L.; Miller, K.D.; Jemal, A. Cancer statistics, 2015. *CA Cancer* **2015**, *65*, 5–29. [CrossRef] [PubMed]
2. Brooks, L.A.; Tidy, J.A.; Gusterson, B.; Hiller, L.; O'Nions, J.; Gasco, M.; Marin, M.C.; Farrell, P.J.; Kaelin, W.G.; Crook, T. Preferential retention of codon 72 arginine p53 in squamous cell carcinomas of the vulva occurs in cancers positive and negative for human papillomavirus. *Cancer Res.* **2000**, *60*, 6875–6877. [PubMed]
3. Condon, J.; Rumbold, A.; Thorn, J.; O'Brien, M.; Davy, M.; Zardawi, I. A cluster of vulvar cancer and vulvar intraepithelial neoplasia in young australian indigenous women. *Cancer Causes Control* **2009**, *20*, 67–74. [CrossRef] [PubMed]
4. Joura, E.A. Epidemiology, diagnosis and treatment of vulvar intraepithelial neoplasia. *Curr. Opin. Obstet. Gynecol.* **2002**, *14*, 39–43. [CrossRef] [PubMed]
5. Liegl, B.; Regauer, S. P53 immunostaining in lichen sclerosus is related to ischaemic stress and is not a marker of differentiated vulvar intraepithelial neoplasia (d-VIN). *Histopathology* **2006**, *48*, 268–274. [CrossRef] [PubMed]
6. Van de Nieuwenhof, H.P.; Massuger, L.F.; van der Avoort, I.A.; Bekkers, R.L.; Casparie, M.; Abma, W.; van Kempen, L.C.; de Hullu, J.A. Vulvar squamous cell carcinoma development after diagnosis of vin increases with age. *Eur. J. Cancer* **2009**, *45*, 851–856. [CrossRef] [PubMed]
7. Van de Nieuwenhof, H.P.; Bulten, J.; Hollema, H.; Dommerholt, R.G.; Massuger, L.F.; van der Zee, A.G.; de Hullu, J.A.; van Kempen, L.C. Differentiated vulvar intraepithelial neoplasia is often found in lesions, previously diagnosed as lichen sclerosus, which have progressed to vulvar squamous cell carcinoma. *Mod. Pathol.* **2011**, *24*, 297–305. [CrossRef] [PubMed]
8. Kokka, F.; Singh, N.; Faruqi, A.; Gibbon, K.; Rosenthal, A.N. Is differentiated vulval intraepithelial neoplasia the precursor lesion of human papillomavirus-negative vulval squamous cell carcinoma? *Int. J. Gynecol. Cancer* **2011**, *21*, 1297–1305. [CrossRef] [PubMed]
9. New Nomenclature for Vulvar Disease. Report of the committee on terminology of the international society for the study of vulvar disease. *J. Reprod. Med.* **1990**, *35*, 483–484.
10. Ridley, C.M.; Frankman, O.; Jones, I.S.; Pincus, S.H.; Wilkinson, E.J.; Fox, H.; Friedrich, E.G.J.R.; Kaufman, R.H.; Lynch, P.J. New nomenclature for vulvar disease: International society for the study of vulvar disease. *Hum. Pathol.* **1989**, *20*, 495–496. [CrossRef]
11. Pinto, A.P.; Miron, A.; Yassin, Y.; Monte, N.; Woo, T.Y.; Mehra, K.K.; Medeiros, F.; Crum, C.P. Differentiated vulvar intraepithelial neoplasia contains tp53 mutations and is genetically linked to vulvar squamous cell carcinoma. *Mod. Pathol.* **2010**, *23*, 404–412. [CrossRef] [PubMed]
12. Aide, S.; Lattario, F.R.; Almeida, G.; do Val, I.C.; Carvalho Mda, G. Promoter hypermethylation of death-associated protein kinase and p16 genes in vulvar lichen sclerosus. *J. Low. Genit. Tract Dis.* **2012**, *16*, 133–139. [CrossRef] [PubMed]
13. Fox, H.; Wells, M. Recent advances in the pathology of the vulva. *Histopathology* **2003**, *42*, 209–216. [CrossRef] [PubMed]
14. Gustafsson, J.O.R.; Oehler, M.K.; Ruszkiewicz, A.; McColl, S.R.; Hoffmann, P. Maldi imaging mass spectrometry (MALDI-IMS)—Application of spatial proteomics for ovarian cancer classification and diagnosis. *Int. J. Mol. Sci.* **2011**, *12*, 773–794. [CrossRef] [PubMed]
15. Balluff, B.; Rauser, S.; Meding, S.; Elsner, M.; Schöne, C.; Feuchtinger, A.; Schuhmacher, C.; Novotny, A.; Jütting, U.; Maccarrone, G. Maldi imaging identifies prognostic seven-protein signature of novel tissue markers in intestinal-type gastric cancer. *Am. J. Pathol.* **2011**, *179*, 2720–2729. [CrossRef] [PubMed]
16. Marko-Varga, G.; Fehniger, T.E.; Rezeli, M.; Döme, B.; Laurell, T.; Végvári, Á. Drug localization in different lung cancer phenotypes by maldi mass spectrometry imaging. *J. Proteom.* **2011**, *74*, 982–992. [CrossRef] [PubMed]

17. Crecelius, A.C.; Cornett, D.S.; Caprioli, R.M.; Williams, B.; Dawant, B.M.; Bodenheimer, B. Three-dimensional visualization of protein expression in mouse brain structures using imaging mass spectrometry. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 1093–1099. [CrossRef] [PubMed]
18. Campello, R.J.; Moulavi, D.; Sander, J. Density-based clustering based on hierarchical density estimates. In *Advances in Knowledge Discovery and Data Mining*; Springer: Berlin & Heidelberg, Germany, 2013; pp. 160–172.
19. Reyes, M.C.; Cooper, K. An update on vulvar intraepithelial neoplasia: Terminology and a practical approach to diagnosis. *J. Clin. Pathol.* **2014**, *67*, 290–294. [CrossRef] [PubMed]
20. Trietsch, M.D.; Nooij, L.S.; Gaarenstroom, K.N.; van Poelgeest, M.I. Genetic and epigenetic changes in vulvar squamous cell carcinoma and its precursor lesions: A review of the current literature. *Gynecol. Oncol.* **2015**, *136*, 143–157. [CrossRef] [PubMed]
21. Vasca, V.; Vasca, E.; Freiman, P.; Marian, D.; Luce, A.; Mesolella, M.; Caraglia, M.; Ricciardiello, F.; Duminica, T. Keratin 5 expression in squamocellular carcinoma of the head and neck. *Oncol. Lett.* **2014**, *8*, 2501–2504. [CrossRef] [PubMed]
22. Chu, P.G.; Weiss, L.M. Expression of cytokeratin 5/6 in epithelial neoplasms: An immunohistochemical study of 509 cases. *Mod. Pathol.* **2002**, *15*, 6–10. [CrossRef] [PubMed]
23. Khanna, N.; Rauh, L.A.; Lachiewicz, M.P.; Horowitz, I.R. Margins for cervical and vulvar cancer. *J. Surg. Oncol.* **2016**, *113*, 304–309. [CrossRef] [PubMed]
24. Nofech-Mozes, S.; Khalifa, M.A.; Ismiil, N.; Saad, R.S.; Hanna, W.M.; Covens, A.; Ghorab, Z. Immunophenotyping of serous carcinoma of the female genital tract. *Mod. Pathol.* **2008**, *21*, 1147–1155. [CrossRef] [PubMed]
25. Gustafsson, O.J.; Eddes, J.S.; Meding, S.; McColl, S.R.; Oehler, M.K.; Hoffmann, P. Matrix-assisted laser desorption/ionization imaging protocol for *in situ* characterization of tryptic peptide identity and distribution in formalin-fixed tissue. *Rapid Commun. Mass Spectrom.* **2013**, *27*, 655–670. [CrossRef] [PubMed]
26. Gustafsson, J.O.; Oehler, M.K.; McColl, S.R.; Hoffmann, P. Citric acid antigen retrieval (CAAR) for tryptic peptide imaging directly on archived formalin-fixed paraffin-embedded tissue. *J. Proteom. Res.* **2010**, *9*, 4315–4328. [CrossRef] [PubMed]
27. Meding, S.; Martin, K.; Gustafsson, O.J.; Eddes, J.S.; Hack, S.; Oehler, M.K.; Hoffmann, P. Tryptic peptide reference data sets for maldi imaging mass spectrometry on formalin-fixed ovarian cancer tissues. *J. Proteom. Res.* **2012**, *12*, 308–315. [CrossRef] [PubMed]
28. Thiele, H.; Heldmann, S.; Trede, D.; Strehlow, J.; Wirtz, S.; Dreher, W.; Berger, J.; Oetjen, J.; Kobarg, J.H.; Fischer, B. 2D and 3D MALDI-imaging: Conceptual strategies for visualization and data mining. *Biochim. Biophys. Acta* **2014**, *1844*, 117–137. [CrossRef] [PubMed]
29. Lokman, N.A.; Elder, A.S.; Ween, M.P.; Pyragius, C.E.; Hoffmann, P.; Oehler, M.K.; Ricciardelli, C. Annexin a2 is regulated by ovarian cancer-peritoneal cell interactions and promotes metastasis. *Oncotarget* **2013**, *4*, 1199–1211. [CrossRef] [PubMed]
30. 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. [CrossRef] [PubMed]



© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (http://creativecommons.org/licenses/by/4.0/).

Appendix II

Applications of Mass Spectrometry Imaging to Cancer

This page intentionally left blank

CHAPTER TWO

Applications of Mass Spectrometry Imaging to Cancer

G. Arentz*,†, P. Mittal*,†, C. Zhang*,†, Y.-Y. Ho*, M. Briggs*,†,‡, L. Winderbaum*, M.K. Hoffmann*,†, P. Hoffmann*,†,1

*Adelaide Proteomics Centre, School of Biological Sciences, University of Adelaide, Adelaide, SA, Australia
†Institute for Photonics and Advanced Sensing (IPAS), University of Adelaide, Adelaide, SA, Australia
‡ARC Centre for Nanoscale BioPhotonics (CNBP), University of Adelaide, Adelaide, SA, Australia
[1]Corresponding author: e-mail address: peter.hoffmann@adelaide.edu.au

Contents

1. Introduction 28
   1.1 The Advantages of MSI 29
   1.2 The Basic Principles of MSI 30
2. Protein MSI in Cancer Research 33
   2.1 Distinguishing Tissue Types by Peptide MSI 36
   2.2 Determining Tumor Margins by Peptide MSI 37
   2.3 Prediction of Metastasis by Peptide MSI 37
   2.4 Analysing Chemoresponse by Peptide MSI 39
   2.5 Identification of Diagnostic and Prognostic Markers by Peptide MSI 39
   2.6 Characterisation of Intra- and Intertumor Variability by Peptide MSI 41
   2.7 Practical Considerations for Proteolytic Peptide MSI: Sample Preparation 42
   2.8 Practical Considerations for Peptide MSI: Spatial Resolution 43
   2.9 Practical Considerations for Peptide MSI: Mass Analysers 44
   2.10 Practical Considerations for Peptide MSI: Identification 44
3. Lipid MSI in Cancer Research 45
   3.1 Profiling Lipids in Cancer by DESI-MSI 45
   3.2 Profiling Lipids in Cancer by SIMS-MSI 46
4. Glycan MSI in Cancer Research 47
5. Drug Imaging in Cancer Research 50
   5.1 MALDI-MSI on Tissue Sections 51
   5.2 MALDI-MSI on Whole Body Sections 53
   5.3 MALDI-MSI on 3D Tissue Cultures 53
6. Data Analysis 53
   6.1 Spatial Information 53
   6.2 Preprocessing: Peak Detection 55
   6.3 Classification of FFPE-TMAs and the Importance of Dimension Reduction 56
7. Concluding Remarks 57
   7.1 The Future of MSI: Molecular Pathology 57
References 57

Advances in Cancer Research, Volume 134
ISSN 0065-230X
http://dx.doi.org/10.1016/bs.acr.2016.11.002

© 2017 Elsevier Inc.
All rights reserved.

Abstract

Pathologists play an essential role in the diagnosis and prognosis of benign and cancerous tumors. Clinicians provide tissue samples, for example, from a biopsy, which are then processed and thin sections are placed onto glass slides, followed by staining of the tissue with visible dyes. Upon processing and microscopic examination, a pathology report is provided, which relies on the pathologist's interpretation of the phenotypical presentation of the tissue. Targeted analysis of single proteins provide further insight and together with clinical data these results influence clinical decision making. Recent developments in mass spectrometry facilitate the collection of molecular information about such tissue specimens. These relatively new techniques generate label-free mass spectra across tissue sections providing nonbiased, nontargeted molecular information. At each pixel with spatial coordinates (*x/y*) a mass spectrum is acquired. The acquired mass spectrums can be visualized as intensity maps displaying the distribution of single *m/z* values of interest. Based on the sample preparation, proteins, peptides, lipids, small molecules, or glycans can be analyzed. The generated intensity maps/images allow new insights into tumor tissues. The technique has the ability to detect and characterize tumor cells and their environment in a spatial context and combined with histological staining, can be used to aid pathologists and clinicians in the diagnosis and management of cancer. Moreover, such data may help classify patients to aid therapy decisions and predict outcomes. The novel complementary mass spectrometry-based methods described in this chapter will contribute to the transformation of pathology services around the world.

ABBREVIATIONS

CHCA α-cycano-4-hydroxycinnamic acid
DESI desorption electrospray ionization
DHB 2,5-dihydroxybenzoic acid
FFPE formalin-fixed paraffin embedded
FT-ICR Fourier transformed ion cyclotron resonance
IHC immunohistochemistry
LNM lymph node metastasis
MALDI matrix-assisted laser desorption/ionization
MSI mass spectrometry imaging
PCA principle component analysis
SIMS secondary ion mass spectrometry
TOF time-of-flight

1. INTRODUCTION

In order to optimize the treatment outcomes for cancer patients and even provide a cure, several issues regarding the clinical management of disease need to be addressed. Based on pathological and clinical parameters

patients are classified to receive the best possible therapy. The accuracy of this process would be better if a more complete understanding of the molecular events involved in the development and progression of cancer was available. Currently, mass spectrometry imaging (MSI) is being extensively applied to the in situ molecular analysis of cancerous cells and tissues with the aim of identifying tumor margins, classifying primary tumor tissues with regards to chemoresponse and metastatic status, identifying diagnostic and prognostic markers, and the analysis of drug response rates and resistance. Elucidation of the spatial location and abundance of peptides, proteins, lipids, glycans, and drug metabolites in relation to cancerous tissues and cells has the potential to significantly improve the diagnosis, staging, and treatment of disease.

1.1 The Advantages of MSI

Interpatient heterogeneity significantly impacts the success of diagnostic and prognostic tests, because it influences the way patients respond to the same treatment. Moreover, proteomic analysis is impacted by intratumor heterogeneity, as tumors are complex structures comprised of a number of different cell types, such as epithelial or endothelial, stromal, vascular, and inflammatory cells. This is particularly important when considering concepts such as innate chemoresistance, metastasis, and stem-cell-like cancer cells, which are believed to consist of a small subset of cells within the bulk of the tumor tissue. Analysis methods that require homogenisation of a sample lose important information about the spatial location of the molecules being studied. Moreover, the information of a potential difference in protein abundance of a small number of cells might not be reflected in the average abundance of the tissue. Certain analysis methods, such as immunohistochemistry (IHC) or fluorescence microscopy, allow for the quantification of compounds within a spatial context. However, these methods are targeted, requiring prior knowledge about the target and the sample in question, and only a small number of molecules can be analyzed simultaneously.

The rapidly evolving technique of MSI allows for the in situ analysis of biological samples, combining classical mass spectrometry (MS) with histological tissue analysis. MSI has the capacity to determine the relative intensity and spatial distribution of several hundreds of compounds from cells and tissue while retaining important spatial information (Schwamborn &

Caprioli, 2010). In the context of cancer research this affords the technique the ability to characterize the molecular features of a sample while maintaining morphology and requiring no prior knowledge of the molecular expression profile. The type of molecules detected, or ions in the context of MS, is dependent on the sample preparation, instrumentation, and acquisition protocols used.

1.2 The Basic Principles of MSI

MS analysis includes the processes of analyte ionization, mass analysis, and detection (de Hoffmann & Stroobant, 2007). During ionization analyte molecules become electrically charged and liberated into the gas phase. Once ionized the mass of the charged molecules is determined by analysis in the mass spectrometer; the analyte ions are separated based on their mass-to-charge ratio (m/z) and detected, commonly by a microchannel plate detector. In terms of MSI, there are two basic forms: microprobe mode imaging and microscope mode imaging (Bodzon-Kulakowska & Suder, 2016; Klerk, Maarten Altelaar, Froesch, McDonnell, & Heeren, 2009; Luxembourg, Mize, McDonnell, & Heeren, 2004). In microprobe mode MSI, where information is collected sequentially from each pixel (pixel sizes typically 5–100 μm), each mass spectrum contains information about the m/z and relative intensity of the detected ions at each location. Following data acquisition, the collected information is reconstructed into a molecular image. In microscope mode MSI, spectral information is collected over a relatively large sample area (100–300 μm) simultaneously; in this form of MSI, the ions produced over the analyzed area are measured using a position-sensitive two-dimensional detection system. The information collected by both microprobe and microscope mode can be displayed as ion maps which show the distribution of the detected m/z values and their intensities, thereby characterizing cancer cells or tissues within their environment in a spatial context.

Following MSI the sample morphology is frequently retained and can be histochemically analyzed postdata acquisition (Crecelius et al., 2005; Schwamborn et al., 2007), making it feasible to directly compare MSI results with histological staining. These methods are complementary and have a high potential for aiding pathologists and clinicians in their diagnosis and management of disease (Rodrigo et al., 2014). As indicated earlier, MSI usually requires very little material and, for example,

a single tissue section from an endoscopic biopsy is enough to perform a successful MSI experiment.

For MSI, the ionization methods matrix-assisted laser desorption/ionization (MALDI), secondary ion mass spectrometry (SIMS), and desorption electrospray ionization (DESI) are popular, and are frequently coupled to time-of-flight (TOF) or Fourier transform type mass analyzers. A very brief description of the commonly applied techniques and their capabilities is provided in the following sections.

1.2.1 MALDI-TOF

MALDI-TOF is commonly used for the analysis of proteins, peptides, lipids, metabolites, and glycans (Cho et al., 2015; Signor & Boeri Erba, 2013). The application of MALDI-TOF to imaging MS was first introduced in 1997 (Caprioli, Farmer, & Gile, 1997) and the technique has gained significant popularity over the last two decades. For analysis by MALDI-TOF-MSI, a sample is placed on a conductive surface such as a stainless steel plate or an indium tin oxide-coated glass slide. Commonly fresh frozen or formalin-fixed paraffin-embedded (FFPE) tissue sections 4–10 μm thick are analyzed by MALDI-TOF-MSI, with FFPE tissues requiring additional preparation steps of deparaffinization and antigen retrieval (Ronci et al., 2008). Once the tissue is placed on the slide, depending on the type of analysis being performed, a digestion step may be carried out. For proteolytic peptide analysis, the most common enzyme utilized is trypsin, for glycan analysis a peptide-*N*-glycosidase F (PNGase F) digestion may be employed to release N-linked glycan structures from proteins (Gustafsson et al., 2015). The analysis of lipids, metabolites, endogenous peptides, and intact proteins directly from tissue does not require enzyme treatment. Prior to analysis a small organic acid, referred to as the matrix, is suspended in an acidified solvent buffer and deposited over the sample (Goodwin, 2012; Kaletas et al., 2009). During this process the matrix incorporates the analyte molecules, forming crystals. A focussed laser beam can then be applied to the surface in order to induce analyte ionization. The small organic molecules of the matrix mixture are designed to absorb the energy of the focussed laser beam resulting in an explosive desorption of the matrix and analyte crystals. The protons present in the acidified matrix solution also facilitate the analyte ionization. MALDI is a "soft" ionization technique, meaning many analytes may be analyzed without significant fragmentation.

Following ionization the ions are rapidly accelerated though a strong electric field in the ion source, the ions then enter the TOF tube or "drift" region, allowing separation based on velocity and thus m/z. The charged analyte molecules strike the detector and the number of events within a time period are recorded, providing intensity information. By comparing the measured information to calibration standards, the mass of the detected analyte ion can be calculated. The spatial resolution of MALDI-TOF imaging sits at around 20 μm and can accommodate a molecular mass range of over 100 kDa.

1.2.2 MALDI-FT-ICR

MALDI-FT-ICR is routinely applied to the imaging analysis of lipids, drugs, and metabolites. MALDI is carried out as described earlier with the Fourier transform ion cyclotron resonance (FT-ICR) acting as the mass analyser, where the m/z of analyte ions is determined by the cyclotron frequency of the ions within the fixed magnetic field of the instrument (Scigelova, Hornshaw, Giannakopulos, & Makarov, 2011).

Following ionization the charged molecules are trapped in a Penning trap, which consists of a series of electrodes that use a homogeneous axial magnetic field to capture the molecules radially, and an inhomogeneous quadrupole electric field to confine the molecules axially. Within the Penning trap, a RF electric field is applied in the plane of ion cyclotron motion which excites the ions to move in a larger, synchronous, cyclotron radius. Once the excitation field is removed, the ions are left rotating at their cyclotron frequencies in phase. The ions induce a charge which is detected as an image current by electrode detectors in close proximity. The recorded signal consists of a superposition of sine waves; this information can be used to calculate the masses of the detected ions using a Fourier transform equation. In FT-ICR a superconducting magnet is used to generate a highly stable magnetic field resulting in very high levels of resolution.

1.2.3 SIMS-TOF

SIMS utilizes beams of primary monoatomic (e.g., $Bi+$, $Au+$) or polyatomic ($C+_{60}$, $Bi+_3$, $Cs+_n$) ions in a vacuum for ionization (Fletcher, Vickerman, & Winograd, 2011). During SIMS the sample surface is bombarded with a beam of primary ions, causing molecules on the surface of the sample to be released, a process referred to as sputtering. During sputtering energy from the primary ions is transferred to the analyte molecules. Around 1% of the sputtered analyte molecules possess an electric

charge; these ions can be detected by a mass analyzer, such as a TOF (as described earlier). The energy of the primary ion beam is typically high compared to the covalent bond energies of the analyte molecules, resulting in their fragmentation. This makes SIMS a "hard" ionization technique, although polyatomic ion beams offer softer ionization than monoatomic ion beams (Bodzon-Kulakowska & Suder, 2016). SIMS requires no matrix and offers higher depth and spatial resolution (<10 μm) as compared to MALDI. However, MALDI is more sensitive as it has access to higher yields of analyte and is able to detect larger molecules (Fletcher, Lockyer, & Vickerman, 2011), with the highest detectable mass range of SIMS sitting at around 1 kDa.

1.2.4 DESI

DESI uses an electrospray source of highly charged aqueous spray droplets under ambient conditions to gently desorb and ionize analyte molecules from the sample surface (Takáts, Wiseman, Gologan, & Cooks, 2004). An advantage of DESI is that it allows for the analysis of a sample in its native state. In the ion source an electrospray emitter is used to generate charged microdroplets from a solvent, which create a thin film on the sample surface that dissolves the analyte molecules. Secondary microdroplets are formed from the kinetic impact of the primary droplets, which contain the highly charged solvent and dissolved analyte. These secondary microdroplets are ionized from the sample surface to the mass spectrometer inlet. A DESI source can be coupled to most standard electrospray mass spectrometers (Hsu & Dorrestein, 2015) where following ionization, the analytes are sent via an ion transfer tube to the mass analyser for detection. During ionization little analyte fragmentation occurs, hence DESI is considered a "soft" ionization technique with a spatial resolution of around 50–100 μm and an upper mass range detection limit of around 2 kDa.

2. PROTEIN MSI IN CANCER RESEARCH

Changes in protein function play a crucial role in the development of cancer. Such alterations can be caused by number of events including altered localization, posttranslational modifications (PTMs), and/or abundance levels. Furthermore, the biology of a tumor depends not only on the cancerous cells but also on their interaction with surrounding stroma, blood vessels, and the immune system (Kriegsmann, Kriegsmann, & Casadonte, 2015). The ability to spatially resolve different proteins while simultaneously

acquiring information concerning their relative abundance is particularly important in heterogeneous diseases such as cancer. Therefore, an approach combining high-resolution MS with in situ spatial analysis of the diseased area has the potential to directly identify and quantify proteins and peptides.

When MSI first entered the field of tissue-based research, it was used to analyze the spatial distribution of proteins (Stoeckli, Chaurand, Hallahan, & Caprioli, 2001). Following this ground-breaking research, different imaging methods have been established and applied to a wide range of biological problems. Among the several MSI techniques, MALDI-MSI is one of the most commonly used. MALDI-MSI has the capability of detecting a broad range of compounds present in a tissue section, provided the compound can be ionized and desorbed into gaseous phase during the MALDI process. High-resolution MS, such as MALDI-FT-ICR, has greater strength in the field of low-molecular weight compounds, while MALDI-TOF is used for the analysis of peptides and small proteins up to 25 kDa (Aichler & Walch, 2015). MALDI-MSI of thin tissue sections can result in the detection of over 500 individual protein signals within the mass range of 2–70 kDa that can be directly correlated with specific morphological regions of the tissue (Chaurand, Schwartz, & Caprioli, 2004).

Protocols have been developed for the preparation of tissue sections and matrix application that provide a high level of reproducibility (Gustafsson et al., 2013). The methodology behind protein MSI, via the analysis of their proteolytic peptides, is straightforward; a digestion enzyme is uniformly applied over the entire tissue surface either by a spotting or spray deposition method. Thereafter, using the same application method, internal calibrants and matrix are applied. Data acquisition is then carried out by a MS (Gustafsson et al., 2013), where analyte ionization, mass analysis, and detection occur. This results in the subsequent generation of mass spectrums at regular series of points across the tissue section. The mass spectra produced are then visualized as ion intensity maps; generating two-dimensional distribution maps of the detected mass spectra containing information about the location and the relative intensity of the detected peptides. This visualization is similar in nature to images generated by histological or IHC analysis, but with the advantage that multiple molecules can be examined within a single measurement (Aichler & Walch, 2015) (Fig. 1).

Three-dimensional MALDI-MS images can also be generated by analyzing a number of consecutively cut tissue sections, and following data acquisition, combining the results to generate three-dimensional maps (Andersson, Groseclose, Deutch, & Caprioli, 2008; Sinha et al., 2008).

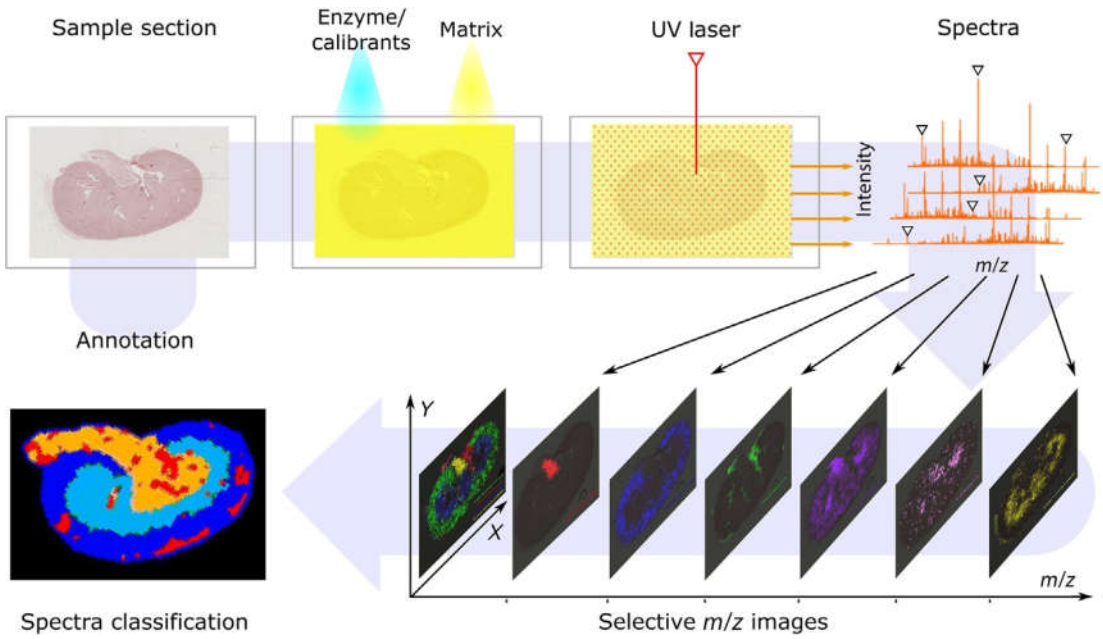


Fig. 1 Work flow of MALDI mass spectrometry imaging.

Recently, methods were developed for the integration of three-dimensional volume reconstructions of spatially resolved MALDI-MSI ion intensity images taken of whole mouse heads with corresponding in vivo data provided by magnetic resonance imaging (MRI) (Sinha et al., 2008). The added spatial dimension provides contextual information to the MALDI-MSI data that allow the interrogation of proteomic relationships within a tissue volume. The work displayed a correlation between the in vivo MRI and proteomic profiles obtained by MALDI-MSI, with both techniques able to distinguish tumor from healthy tissues (Sinha et al., 2008).

Until recently, MALDI-MSI had been carried out almost exclusively on fresh frozen samples as it was believed proteins were inaccessible from FFPE tissue due to the cross-linking caused by the formalin fixation. However, this limitation has been overcome and equal number of proteins can be identified from both fresh frozen and FFPE tissues (Diehl et al., 2015). The use of FFPE tissues in cancer research is a huge advantage, given that the samples are well preserved by the formalin fixation process and can be stored for long periods of time at room temperature. Another attractive feature is that the samples are often well documented as they have been processed by clinical pathology departments, meaning information on patient treatment response, disease progression, and other relevant clinical data have been cataloged (Gorzolka & Walch, 2014). Additionally, FFPE tissues are well suited for the construction of tissue microarrays (TMAs), where needle core biopsies from multiple tissue samples are assembled into a paraffin block. However,

fresh frozen TMAs have also been studied in a limited context (Fejzo & Slamon, 2010).

One of the main advantages of MALDI-MSI is the technique's ability to detect changes in disease-associated protein expression prior to histological transformation, making it a useful tool in the identification of early stage disease biomarkers (McDonnell et al., 2010). There have been number of cancer studies that have successfully applied peptide imaging to the discovery of potential biomarkers with the capacity to differentiate tumors of various subtypes, stages, and/or degree of metastasis.

2.1 Distinguishing Tissue Types by Peptide MSI

The ability of peptide MSI to distinguish specific types of tissue, such as regions of healthy, cancer, stroma, and vasculature, has been shown multiple times (Mittal, Klingler-Hoffmann, Arentz, Zhang, et al., 2016). Schwamborn et al. performed MSI on prostate cancer specimens and was able to classify the cancerous tissue from normal tissue with 85% sensitivity and 90% specificity, with an overall cross validation accuracy of 88% (Schwamborn et al., 2007). Moreover, MSI of prostate cancer TMAs on a large scale allowed for the immediate prioritization of MSI signals based on associations with clinic-pathological and molecular data (Steurer et al., 2013). A comparison of these signals with clinic-pathological features revealed statistical association with a favorable phenotype such as grade, stage, and relationship with prolonged time to prostate-specific antigen recurrence (Steurer et al., 2013). Peptide MSI experiments have been able to discriminate normal, preinvasive, and invasive lung tumor tissues with an accuracy of 90% (Rahman et al., 2011), to differentiate normal tissue from oral squamous cell carcinoma (Patel et al., 2009), to discriminate normal tissue from gastric cancer tissues (Kim et al., 2010), and have identified mitogen-activated protein kinase/extracellular signal-regulated kinase kinase 2 (MEKK2) as a marker with the capacity to accurately discern cancerous prostate from normal tissue (Cazares et al., 2009).

One of the most challenging aspects of neurochemistry is the detection of endogenous neuropeptides due to their low in vivo concentrations ranging from pico- to femto-molar levels, which makes it difficult to obtain enough material for quantitative analysis (Andersson, Andren, & Caprioli, 2010; Strand, 2003). MALDI-MSI helps overcome this issue as it allows the direct detection of neuropeptides in discrete regions of a brain sections, enabling the study of the physiological and disease-related

metabolic processing of neuropeptides (Andersson et al., 2010). Recently, Andersson et al. has been able to sequence and identify neuropeptides by in situ tandem mass spectrometry (MS/MS) directly off rat brain sections using a MALDI-QTOF mass spectrometer (Andersson et al., 2008).

Furthermore, peptide MALDI-MSI has been employed to differentiate benign regions from malignant in ovarian cancer (Lemaire, Ait Menguellet, et al., 2007) and in thyroid cancer (Pagni et al., 2015). Likewise, biomarkers have been identified that could distinguish Hodgkin's lymphoma from lymphadenitis (Schwamborn et al., 2010).

2.2 Determining Tumor Margins by Peptide MSI

One of the major aims of clinical oncology is to ensure a tumor has been completely removed during surgery to minimize the possibility of disease recurrence (Han et al., 2011). Defining tumor margins also plays a pivotal role in cancer staging, which subsequently influences the administration of treatment. Moreover, the molecular characterization of tumor margins will assist in understanding the process of invasion into surrounding tissues and may help in the clinical management of disease (Agar et al., 2010). MALDI-MSI has been proposed as a valuable tool for gaining information with regards to tumor margins and heterogeneity (Chaurand, Sanders, Jensen, & Caprioli, 2004). Peptide MSI has interestingly revealed histologically normal tissue adjacent to renal cell carcinoma regions, share many molecular characteristics with that of the tumor tissue, providing unprecedented insight into cancer development (Oppenheimer, Mi, Sanders, & Caprioli, 2010). The technique has also been used to identify protein biomarkers for the differentiation of delineated hepatocellular carcinoma from adjacent cirrhotic tissue (Le Faouder et al., 2011). Analysis of ovarian cancer interface zones, the regions between tumor and normal tissues, by MSI revealed a unique peptide profile from both the tumor and normal tissue, and detected plastin 2 and peroxiredoxin 1 as interface markers specific to ovarian cancer (Kang et al., 2010).

2.3 Prediction of Metastasis by Peptide MSI

Metastatic status is often the defining feature in determining if a patient will receive chemotherapy and is currently resolved by the removal and analysis of local lymph nodes, which frequently results in noncancer-related health complications. Identification of primary tumor markers of metastasis

would prevent the unnecessary removal of lymph nodes, thereby reducing patient suffering and associated treatment costs, and allow the introduction of more personalized cancer therapies (Casadonte et al., 2014). However, identification of metastasis based on analysis of the primary tumor tissues is challenging, especially when the primary carcinoma site is unknown (Pavlidis & Pentheroudakis, 2016). A statistical classification model based on peptide MSI data has been developed for determining the metastatic status of primary cancers of pancreatic origin and their metastasis to secondary breast cancer with an overall accuracy of 83% (Casadonte et al., 2014). In another study the HER2 status of gastric cancers could be predicted with 90% accuracy using peptide expression patterns originating from breast cancers (Balluff et al., 2010). These studies display the potential of peptide MSI in the profiling of cancers, independent of their site of origin.

The likelihood of lymph node metastasis (LNM) is one of the most important factors to consider while determining the appropriate treatment for a cancer patient. Several studies have applied MALDI-MSI to the prediction of LNM status. Yanagisawa et al. was able to classify primary lung tumors with and without LNM with an accuracy of 85% (Yanagisawa et al., 2003). Different peptide signatures have been identified with the ability to distinguish melanoma (LNM) from control lymph nodes (Hardesty, Kelley, Mi, Low, & Caprioli, 2011). Similarly, the overexpression of S100A10, thioredoxin, and S100A6, have been shown to be highly associated with LNM in papillary thyroid carcinoma (PTC) (Nipp et al., 2011), and it was suggested the three proteins may be used for risk stratification regarding metastatic potential in PTC. Other protein biomarkers have been suggested for the prediction of LNM in several cancers from peptide MSI experiments, such as S100A8 and S100A9 for gastric adenocarcinoma (Choi et al., 2012), COX7A2, TAGLN2, and S100-A10 for monitoring the Barrett's adenocarcinoma development as well as for predicting regional (LNM) and disease outcome (Elsner et al., 2012).

Recently, it has been shown that the LNM status of primary endometrial tumors could be predicted with an accuracy of 88% using a classification system based on tissue peptide markers detected by MALDI-MSI (Mittal, Klingler-Hoffmann, Arentz, Winderbaum, et al., 2016; Winderbaum, Koch, Mittal, & Hoffmann, 2016). In the study FFPE-TMAs comprised of duplicate tumor tissue cores from endometrial cancer patients diagnosed with and without LNM were analyzed by MALDI-TOF-MSI. An overview of the sample preparation and approach used for the proteomic marker

identification by MALDI-MSI-(top panel) and LC-MS/MS (bottom panel) is outlined in Fig. 2.

2.4 Analysing Chemoresponse by Peptide MSI

Innate and acquired chemoresistance significantly impact the survival rates of late stage cancer patients. Administration of chemotherapy to patients with specific cancer types has traditionally been standardized, despite the knowledge that response can differ significantly among a cohort of patients due to genetic diversity and tumor heterogeneity. Identification of markers or tests that predict resistance to standard chemotherapy regimens would help save patients from damaging, ineffective treatments, and may help determine if an alternative treatment course may be more beneficial.

IHC is one of the most widely applied proteomic techniques in clinical practice. However, IHC is often unable to distinguish the different isoforms of the same protein that may play different role in the disease development. Recent studies using MSI have been able to distinguish two isoforms of the protein defensin. This is of particular interest given the 2 defensin protein isoforms respond differently in breast cancer tumors when exposed to chemotherapy (Bauer et al., 2010). The finding was discovered during the analysis of a breast cancer patient receiving neoadjuvant paclitaxel and radiation therapy, and was one of the first biomarker discovery studies to utilize MALDI-MSI (Bauer et al., 2010). Mitochondrial defects have also been found and linked to chemotherapy response in oesophageal adenocarcinoma patients using peptide MSI (Aichler et al., 2013), displaying the techniques potential to change classifications of tumors and chemoresponse.

2.5 Identification of Diagnostic and Prognostic Markers by Peptide MSI

A high proportion of cancer patients are diagnosed in the later stages of disease, when the cancer has already metastasized to distant lymph nodes or organs. Detecting cancer in the earlier stages of disease, prior to metastasis, would help increase survival rates and reduce the treatment burden for patients. In a genetically engineered mouse model, peptide MSI was used to detected several discriminative peptide masses that could distinguish between intraepithelial neoplasia, intraductal papillary mucinous neoplasm, and normal pancreatic tissue (Grüner et al., 2012). Using MALDI-MSI to analyze TMAs, significant differences between papillary and clear cell renal cell cancer were found (Steurer, Seddiqi, et al., 2014). Analysis of bladder

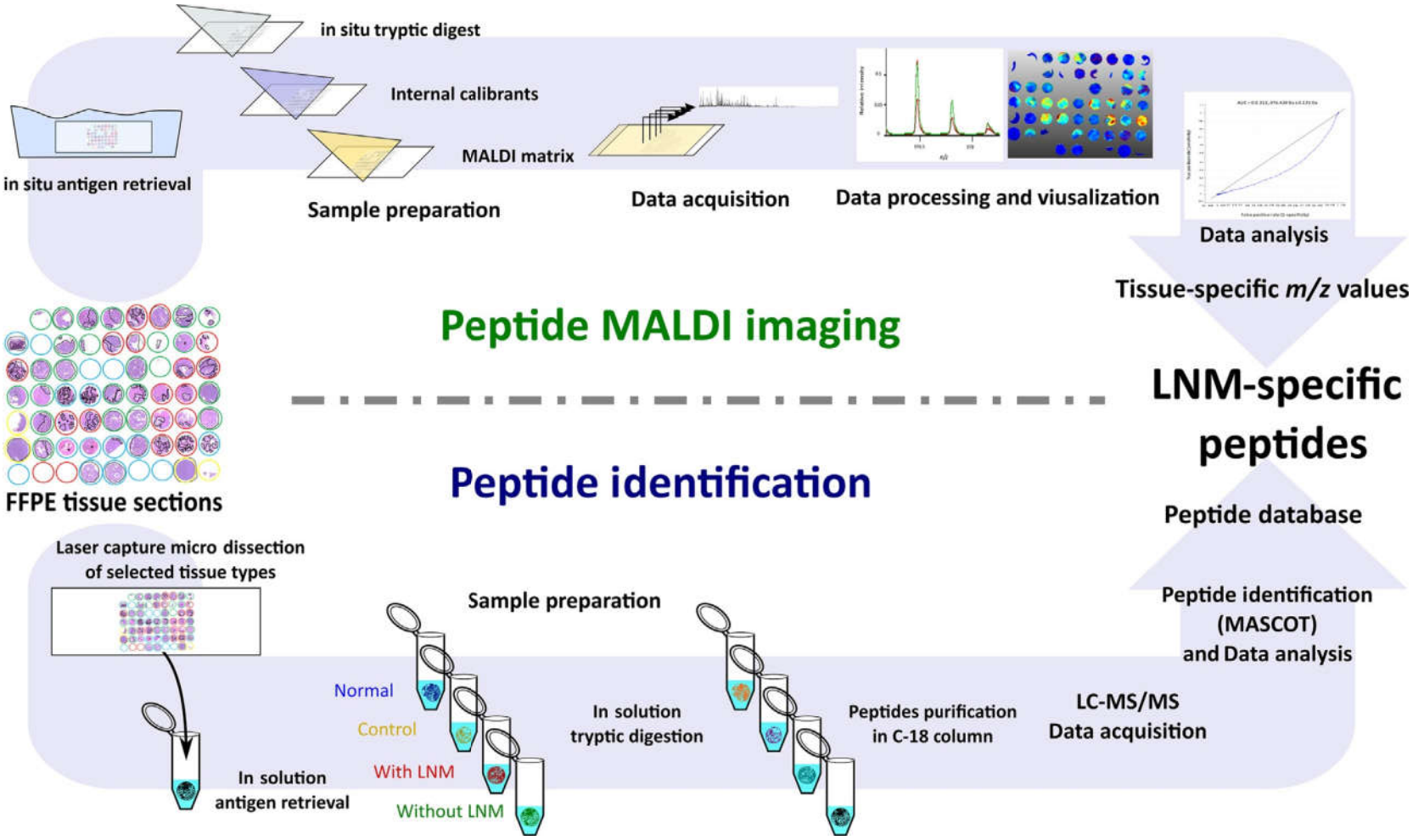


Fig. 2 Work flow for proteomic marker identification by peptide MALDI-MSI and label-free quantification.

cancer TMAs by MSI lead to the identification of multiple signals in the mass spectra associated with unfavorable tumor phenotype and clinical characteristics, such as tumor recurrence, progression, and patient survival time (Steurer, Singer, et al., 2014).

MALDI-MSI has also been used to identify a protein signature that can accurately determine the HER2 status in breast cancer with sensitivity of 83%, for specificity of 92%, and overall accuracy of 89% (Rauser et al., 2010). A seven-protein signature of novel tissue markers was identified as an independent indicator of unfavorable overall survival in intestinal type gastric cancer (Balluff, Rauser, et al., 2011). Other investigators identified histone 4 expression as specific to poorly differentiated gastric cancer tissues by peptide MSI (Morita et al., 2010). A comparison of genomically stable (diploid) and unstable (aneuploid) colon cancer tissues against normal mucosa by MSI identified thymosin beta 4 as an independent prognostic marker for colorectal cancer (Gemoll, Strohkamp, Schillo, Thorns, & Habermann, 2015). Analysis of lung cancer TMAs containing 112 needle core biopsies were used to generate a support vector machine model utilizing MSI information on 73 peptides that could classify adenocarcinomas with an accuracy of 97% and squamous cell carcinomas with an accuracy of 98% (Groseclose, Massion, Chaurand, & Caprioli, 2008).

2.6 Characterisation of Intra- and Intertumor Variability by Peptide MSI

Understanding the spatiotemporal patterns and dynamics of intratumor heterogeneity is crucial for clinical management and designing personalized targeted regimens for cancer patients (Renovanz & Kim, 2014). Balluff et al. showed for the first time that MALDI-MSI of tumor tissues can reveal microscopically indistinct tumor populations that might have adverse impacts in clinical outcome (Balluff et al., 2015). In the proof-of-principal study tumor-specific mass spectra were grouped using an advanced statistically clustering algorithm in order to obtain segmentation maps of molecularly distinct regions. The distinct regions were then statistically compared with patient clinical data to identify *phenotypic* tumor subpopulations (Balluff et al., 2015). This approach revealed that several of the detected tumor subpopulations are associated with different overall survival characteristics of gastric cancer patients ($p = 0.025$) and the presence of regional metastasis in patients with breast cancer ($p = 0.036$). This approach enables researchers to gain deeper insights into the underlying biological

process and changes of tumor subpopulations on a genetic, metabolic, and proteomic level, which may result in novel targeted therapies (Balluff et al., 2015).

The ability of MALDI-MSI to reveal inter- and intratumoral biomolecular heterogeneity of histologically contiguous tumors has been shown, specifically in the capacity to distinguish different types and grades of myxoid sarcomas (Willems et al., 2010). By applying hierarchical clustering to a gastric cancer MALDI-MSI dataset, it has been demonstrated the different histological regions could be distinguished solely on the basis of their protein and peptide expression profiles (Deininger, Ebert, Fütterer, Gerhard, & Röcken, 2008). Ion mobility separation combined with MALDI-MSI has been used to classify prostate cancer TMAs and has identified a number of proteins, including type I collagen and tumor necrosis factor receptor, directly from the tissue that could discriminate different tumor classes (Djidja et al., 2010). Low grade urothelial neoplasia could be separated from high grade urothelial neoplasia with an overall cross validation rate of 97.18% upon analysis of peptide MSI data (Oezdemir et al., 2012). Likewise, based on protein profiling, a grading system for meningioma was developed (Agar et al., 2010). Proteomic signatures associated with breast cancer-activated stromal tissues have been identified by MSI, and hierarchical clustering of the proteomic signals could be used to distinguish the activated intratumoral stroma from the quiescent extratumoral stroma (Dekker et al., 2014). High grade and low grade glioma can also be distinguished by MALDI-MSI (Chaurand et al., 2004).

2.7 Practical Considerations for Proteolytic Peptide MSI: Sample Preparation

MSI of proteolytic peptides in cancer is frequently performed on fresh frozen or FFPE tissues; each sample type has unique advantages and limitations with regards to MSI. The protein component of fresh tissues is easily accessible, especially when compared to FFPE tissues, due to a lack of chemical fixation. Hence, no retrieval steps are required prior to the proteolyic digestion of fresh frozen samples. A disadvantage of fresh frozen tissues, however, is that they require rapid freezing upon collection and storage at extremely low temperatures to inhibit endogenous enzymatic degradation (Fentz, Zornig, Juhl, & David, 2004). Fresh tissues also require solvent washing to remove lipids and other biomolecules that may compete for ionization and interrupt the detection of peptides. Washing steps must be performed carefully, as extensive washing may cause the loss of soluble proteins

(McDonnell et al., 2010; Seeley, Oppenheimer, Mi, Chaurand, & Caprioli, 2008). Two significant advantages of FFPE samples are that they are widely used in modern clinical practice, meaning very large sample banks have often been developed, and that the fixation process allows for storage of the samples at room temperature for indefinite periods of time without loss of morphological information (Wisztorski et al., 2007). The downfall is that formalin fixation results in the formation of methylene bridges between amino residues that causes protein cross-linking (Metz et al., 2004). This cross-linking makes the proteins inaccessible to proteolysis and must be reversed prior to any further preparation steps. Heat-induced antigen retrieval prior to in situ proteolytic digestion is a widely used method for reversing protein cross-links, which allows for the peptide profiling of FFPE tissues that is comparable to fresh frozen tissue analysis (Casadonte & Caprioli, 2011).

2.8 Practical Considerations for Peptide MSI: Spatial Resolution

Currently, MALDI is the most popular ionization method applied in the field of MSI. The success of analyte ionization in a MALDI experiment relies significantly on the matrix used; the type of matrix applied will impact on the mass range of the detectable ions, the intensity of detection, and the minimum spatial resolution achievable (Lemaire et al., 2006). For peptide MALDI-MSI commonly used matrices include α-cycano-4-hydroxycinnamic acid (CHCA) and 2,5-dihydroxybenzoic acid (DHB) (van Hove, Erika, Smith, & Heeren, 2010). The homogeneity of matrix crystallization is a major limiting factor with regard to the spatial resolution that can be achieved by MALDI-MSI, as larger and nonuniform matrix crystals result in lower resolutions. DHB matrix is prone to producing large, needle-like crystals, making CHCA a more suitable matrix as it commonly forms small, homogeneous crystals resulting in higher image resolution (Schwartz, Reyzer, & Caprioli, 2003). The way in which the matrix is applied also has a significant impact on the spatial resolution, with the application of larger size droplets increasing the likely hood of analyte displacement. Instrumentation has been developed that allows matrix solutions to be sprayed or spotted onto the sample surface in a more homogeneous manner compared to manual application (Aerni, Cornett, & Caprioli, 2006). For example, using a matrix spray device CHCA can be applied to form uniform crystals that allow for spatial resolution peptide imaging of $<25\ \mu m$ (McDonnell et al., 2010; Schuerenberg & Deininger, 2010).

Laser specifications also have a significant impact on the spatial resolution achievable. As opposed to the limitations of matrix crystallization, laser beam resolutions can easily reach less than 10 μm, theoretically allowing for peptide imaging at the single cell level (Balluff, Schöne, Höfler, & Walch, 2011). However, reducing the size of the laser beam concurrently reduces the yield of ionized analytes. Higher frequency lasers and overlapping laser shots can counteract the issue of lower ion yields during detection (Jurchen, Rubakhin, & Sweedler, 2005). Thus, currently matrix is the limiting factor for the spatial resolution achievable in MALDI-MSI.

2.9 Practical Considerations for Peptide MSI: Mass Analysers

The type of mass analyser used for MSI will largely determine the mass resolution and mass range of the detectable ions. TOF mass analysers are commonly coupled to MALDI sources and have high acquisition speeds but typically lower mass resolution (Caprioli et al., 1997). Higher mass accuracy can be achieved through the use of internal or external calibrants. Other commonly used mass analysers include FT-ICR and Orbitraps; these instruments have significantly higher mass accuracy and resolution (Taban et al., 2007).

Internal or external calibrants can be used to improve the mass accuracy of acquired spectra during or after data acquisition (Gobom et al., 2002). Calibrants are often comprised of a small number of specific analytes with known masses that span a specific mass range. Within this mass range, the use of calibrants can significantly increase the mass accuracy of the detected analytes. However, the mass error of any analyte molecules that fall outside the mass range of the calibrants cannot be corrected (Coombes, Koomen, Baggerly, Morris, & Kobayashi, 2005).

2.10 Practical Considerations for Peptide MSI: Identification

Currently, MALDI-MSI enables the discrimination of heterogeneous tissue types, but does not allow for the direct identification of the peptide sequences. During data acquisition the intact mass, or the m/z, of the detected peptides is recorded. This allows peptide mass fingerprinting (PMF) to be performed, a process where the intact peptide masses are matched back to a database containing the in silico digested proteome or proteomes of specific species (Pappin, Hojrup, & Bleasby, 1993). Due to a lack of peptide sequence information, often PMF can provide a list of potential peptide identifications but is not able to provide definitive

identifications for those peptides with similar masses. However, the higher the mass accuracy of the instrument used the more likely it is a definitive identification will be gained (Horn, Peters, Klock, Meyers, & Brock, 2004). The use of complementary MS techniques, such as matching data to peptides sequenced by traditional liquid chromatography electrospray ionization tandem mass spectrometry (LC-ESI-MS/MS) or by in situ tandem mass spectrometry (MALDI-MS/MS), can help clarify peptide sequence identifications (Lemaire, Desmons, et al., 2007). However, there can be significant difficulties associated with these options also. Multiple sequence matches are often encountered when matching MSI *m*/z values back to LC-ESI-MS/MS data obtained from the same sample. Generating high quality in situ MS/MS spectrums is often challenging due to the issue that acquisition of very complex peptide mass spectra directly off tissue results in poor fragmentation and high MS/MS complexity, both of which limit direct peptide sequencing from tissues (Gustafsson et al., 2012).

3. LIPID MSI IN CANCER RESEARCH

3.1 Profiling Lipids in Cancer by DESI-MSI

The soft ionization technique of DESI has been developed for the MSI of small molecules and lipids. DESI uses highly charged solvent droplets traveling at high velocities for analyte ionization in an ambient atmosphere. An advantage of DESI for the analysis of lipids and small molecules is that no sample preparation steps are required, unlike MALDI where matrix deposition is a necessity. A further advantage is that a DESI source can be coupled to most standard mass analysers, such as a FT-ICR, Orbitrap, Ion Trap, or a TOF.

In practice, DESI-MSI has been applied to the study of human cancers for the discrimination of tumor subtypes and grades, as well as for the identification of tumor margins (Calligaris et al., 2014; Eberlin et al., 2014). One particular application used DESI-MSI to generate lipid profiles for the classification of human brain tumors (Eberlin et al., 2012). This study was able to classify the subtype, grade, and tumor concentration of 36 human glioma samples via lipidomics imaging in agreement with the histopathology diagnoses (Eberlin et al., 2012). Lipidomic analysis by DESI-MSI has identified the phospholipids phosphatidylserine (PS), phosphatidylinositol (PI), and phosphatidylethanolamine as biomarkers for distinguishing healthy tissue

from benign and malignant tumor tissues (Abbassi-Ghadi et al., 2014; Dill et al., 2011, 2009; Eberlin et al., 2012).

Discriminatory lipid signatures between cancerous and normal breast tissue have been detected by DESI-MSI (Calligaris et al., 2014), with the delineation of tumor margins possible through the analysis of PI (18:0/20:4). PI (18:0/20:4) has been shown to be abundant at the tumor center and tumor edge, whereas it was absent or weak in normal tissues. This lipid sharply stratifies the tumor margin in agreement the histopathological staining. Using DESI-MSI, another study found PS (20:4/18:0) and PI (18:0/20:4) are significantly lower in concentration in primary tumor tissues and LNMs when compared to normal lymph node tissues (Abbassi-Ghadi et al., 2014). The same lipid, PI (18:0/20:4), was found to be a biomarker for distinguishing benign and malignant breast tumors (Yang et al., 2015). Collectively, the discussed research shows MSI lipid profiling of cancer tissues can be used for classifying tumor types, with high specificity in differentiating histopathologic grade, and in the mapping of tumor margins.

3.2 Profiling Lipids in Cancer by SIMS-MSI

A less popular ionization method for MSI in the lipidomics field is static SIMS. Generally, a static SIMS source is coupled to a TOF (or QTOF) instrument, and is particularly suitable for the analysis of intact molecules such as lipids and metabolites. It uses a beam of energetic primary ions with either monatomic source, such as Cs^+ and Ga^+ for limited mass range of low hundreds Daltons, or cluster ions such as SF_5^+, Bi_3^+, and Au_3^+ that have extended the mass range to a couple of thousands Daltons (Lanni, Rubakhin, & Sweedler, 2012). The potential of SIMS-TOF imaging for lipidomics has been shown in rat brain sections, a well-established model system for MSI (Benabdellah et al., 2010; Passarelli & Winograd, 2011; Sjovall, Lausmaa, & Johansson, 2004; Sjovall, Lausmaa, Nygren, Carlsson, & Malmberg, 2003). In the field of cancer lipidomics, it is frequently used in cellular imaging, where single cells or a population of cells are isolated after drug treatment for the characterisation of molecules on the cell surface. For example, lipid composition analysis of individual human breast cancer stem cells (CSCs) reported significantly lower expression levels of palmitoleic acids FA (16:1) as compared to nonstem cancer cells (NSCCs), and is thought to have been successfully characterized from complex clinical specimens using TOF-SIMS (Waki et al., 2014).

4. GLYCAN MSI IN CANCER RESEARCH

Glycosylation is the enzymatic process where oligosaccharides, polysaccharides, and carbohydrates (i.e., glycans) are attached to proteins, lipids, or other organic molecules (Ohtsubo & Marth, 2006). In recent years, the word "glycomics" has emerged to describe the study of glycan structures from an organism's "glycome." The mammalian glycome repertoire is comprised of 10 known monosaccharides. These monosaccharides include xylose (Xyl), glucose (Glc), galactose (Gal), mannose (Man), *N*-acetylglucosamine (GlcNAc), *N*-acetylgalactosamine (GalNAc), fucose (Fuc), glucuronic acid (GlcA), iduronic acid (IdoA), and *N*-acetylneuraminic acid (NeuAc or Neu5Ac), a sialic acid (Wuhrer, Deelder, & van der Burgt, 2011). Glycan complexity arises from variation in expression of specific glycosyltransferases, resulting in a diversity of monosaccharide arrangements. Furthermore, modifications such as sulfation, acetylation, or phosphorylation increase the complexity of glycosylation.

The major type of glycosylation is protein glycosylation, where glycans are attached to asparagine residues by their amine group (N-linked) or at serine/threonine residues by their hydroxyl group (O-linked) (North, Hitchen, Haslam, & Dell, 2009). Protein glycosylation is an important PTM which has relevance in many biological processes such as cell signaling, immune responses, extracellular interaction, and cell adhesion (Varki, 1993). In the tumor microenvironment, aberrant protein glycosylation such as the expression of truncated glycans have been well described in various cancers (Dube & Bertozzi, 2005). These structures may resemble glycomic "fingerprints" which discriminate between healthy and cancerous tissues or potentially discriminate between various cancer subtypes (Abbott et al., 2010).

One of the most common glycosylation changes in various cancers is an increase in branching of N-glycans, where large tetra-antennary structures are formed by the expression of *N*-acetylglucosaminyltransferase V (GnT-V) (Lau & Dennis, 2008). The presence of increased branching antennas creates additional sites for the addition of NeuAc by sialyltransferases in various cancers (Bull et al., 2013). Likewise, terminal modifications on glycan structures such as fucosylation on the cancer cell surface can also give rise to the presence of Lewis and sialyl Lewis antigens (sialyl Lea; [Neu5Acα2-3Galβ1-3(Fucα1-4)GlcNAcβ] and sialyl Lex; [Neu5Acα2-3Galβ1-4(Fucα1-3)

GlcNAcβ]) which have been shown to correlate with tumor progression and metastasis (Nakagoe et al., 2000). It is evident that aberrant glycosylation is indeed a key event in metastasis and therefore the discovery of cancer-specific protein glycosylation markers would be beneficial in the staging and treatment of patients (Drake et al., 2010).

As glycomics gains recognition in this postgenomics era, MS-based methodologies are becoming indispensable and routinely used for the reliable profiling of glycans from clinical samples. Glycan MSI is a recent MS-based development which allows the visualization of the tissue-specific spatial distribution of glycans. N-glycans have been the most comprehensively studied type of protein glycosylation, as they can be easily released from proteins by enzymatic cleavage using PNGase F (Jensen, Karlsson, Kolarich, & Packer, 2012). The majority of these studies have applied optimized N-glycan MALDI-MSI methods to FFPE tissues, with the most recent focus on various cancer tissues.

One of the first N-glycan MALDI-MSI methods to emerge was published in 2013 (Powers et al., 2013). This proof-of-principle paper described the measurement of N-glycans by MALDI-MSI across fresh frozen mouse brain tissue. The m/z values from the sum spectrum were assigned to N-glycans based on further MS structural analysis, for example, high-performance liquid chromatography (HPLC)-MS. The authors concluded that this new approach could lead to novel disease-related targets for biomarker and therapeutic applications. Subsequently, N-glycan MALDI-MSI has been successfully applied to the analysis of glioblastoma tumor xenografts compared to normal brain tissues (Toghi Eshghi et al., 2014), where it was found specific N-glycans associated with distinct tissue regions. N-glycan imaging of FFPE pancreatic and prostate cancers, and a human hepatocellular carcinomas in TMAs found tumor and nontumor regions could be distinguished in the TMA format, enabling high-throughput analysis (Powers et al., 2014).

The use of complimentary MS methods, such as liquid chromatography tandem MS (LC-MS/MS) with MALDI-TOF-MSI, have been developed to enable the mapping of tissue-specific N-glycans at high resolution (Gustafsson et al., 2015). The application of these techniques has resulted in several interesting findings. A panel of over 30N-glycans has been detected from FFPE human hepatocellular carcinoma tissues (Powers, Holst, Wuhrer, Mehta, & Drake, 2015), from which the N-glycans were extracted and the sialylated species were stabilized by an off-tissue ethylation reaction. N-glycan sialyation linkage-specificity was further found to

correlate with the different tissue types analyzed. A novel method for the analysis of N-glycan sialylated species (i.e., α 2–3 and α 2–6) by MALDI-MSI in colon carcinoma found sialylated N-glycans with α 2–3 linkage were observed in stroma, tumor, and necrotic regions, while sialylated N-glycans with α 2–6 linkage were observed in necrotic, collagen-rich, and red blood cell regions (Holst et al., 2016).

The analysis of FFPE stage III ovarian cancer tissues found high mannose structure (Man5 + Man3GlcNAc2) in the tumor regions, while triantennary complex structures (Hex3HexNAc3Deoxyhexose1 + Man3GlcNAc2) were observed in the stromal regions (Fig. 3). This observation was consistent between not only the patients investigated, but also between each N-glycan family. The tissue analyzed in the study had been annotated by a pathologist, and it was found specific glycan structures were differentially detected in specific tissue regions, such as tumor, stroma, and adipose tissue.

In summary, N-glycan MALDI-MSI has developed immensely in recent times, with improvements in areas such as sensitivity and mass accuracy,

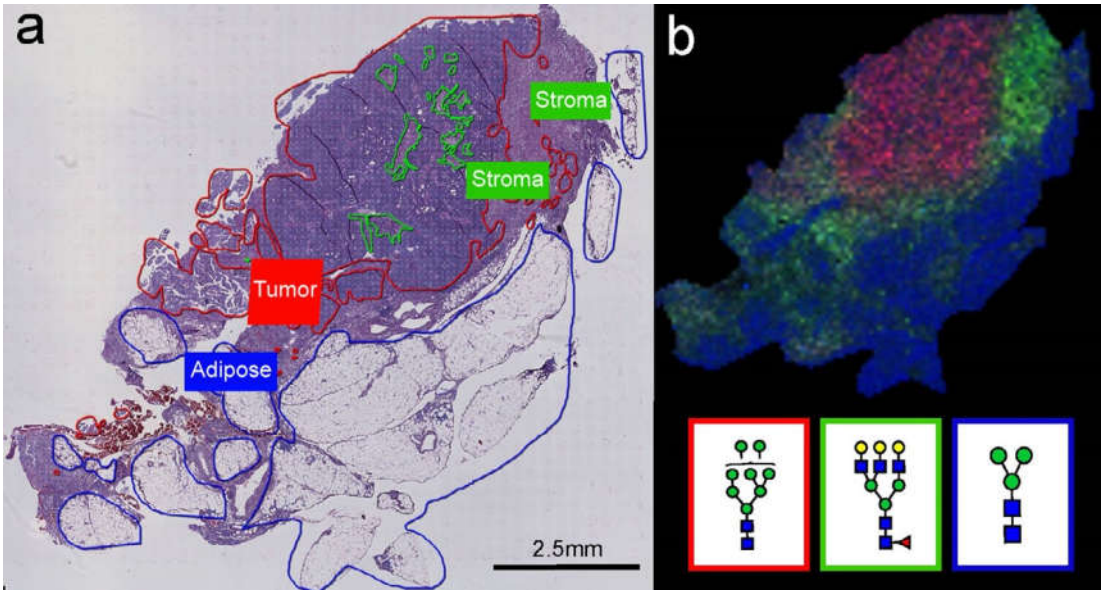


Fig. 3 Glycan analysis of ovarian cancer. The formalin-fixed paraffin-embedded stage III ovarian cancer section was treated with antigen retrieval prior to printing of 15 nL/spot dialyzed PNGase F with 250 μm spacing. 2,5-DHB (20 mg/mL) was sprayed onto the sections and MS spectra were acquired by oversampling at 100 μm intervals using a MALDI-TOF/TOF MS instrument. (A) Haematoxylin and eosin stain of the ovarian cancer section with tumor (*red*), stroma (*green*), and adipose (*blue*) regions annotated by a pathologist. (B) An ion intensity map of *m/z* 1743.7 (*red*), *m/z* 2174.9 (*green*), and *m/z* 933.3 (*blue*) on the ovarian cancer section. *Yellow circle*, galactose (Gal); *green circle*, mannose (Man); *blue square*, N-acetylglucosamine (GlcNAc); *red triangle*, fucose (Fuc).

compositional analysis by complementary techniques (i.e., LC-MS/MS and lectin histochemistry staining), and identification of linkage-specific sialylated species by in situ derivatization methods. However, further work in the field is required to properly answer clinical questions and discover cancer-specific markers. TMAs have not yet been fully utilized in N-glycan MSI experiments, but offer promise in providing researchers and clinicians with high-throughput data for analysis.

5. DRUG IMAGING IN CANCER RESEARCH

Quantification of drugs and their metabolites is crucial in oncology where drug distribution within the tumor tissue is thought to play a pivotal role in response to therapy and could partly explain the variable response rates observed among patients with similar tumor types (Minchinton & Tannock, 2006). Traditionally, autoradiography has been used to visualize the localization of a drug. However, this technique suffers from some significant limitations. One of them is that the synthesis of the radiolabeled drug is an expensive and time-consuming process. Second, the technique does not provide information on the molecular structure. Hence, metabolites either cannot be distinguished from the parent drug or are not detected if the radiolabel is lost in a metabolic process (Kertesz et al., 2008). The other clinically used noninvasive drug imaging techniques are positron emission tomography and MRI. These techniques can provide information on the drug distribution but suffers from the limitation of low spatial resolution. Due to the relatively simple sample preparation steps required for MSI the technique has rapidly emerged as an alternative to the conventional methods of analysis. MSI allows detailed distribution analysis of the parent drug and its metabolites in a single experiment and without any labeling of the targeted compound in contrast to conventional immunohistochemical methods. Such drug characterization in situ, by both spatial and temporal behaviors within tissue compartments, provides a new understanding of the dynamic processes impacting drug uptake and metabolism at the local sites targeted by a drug therapy (Kwon et al., 2015). Multiple MSI techniques including MALDI-MSI, SIMS, and nanostructure initiator mass spectrometry (NIMS) have been used to analyze the distribution of the drug and metabolites. However, MALDI-MSI is the most frequently used method for looking at anticancer drug distribution in tumor samples.

5.1 MALDI-MSI on Tissue Sections

MALDI-MSI was used for the first time in 2003 to directly analyze and image pharmaceutical compounds in intact tissues (Reyzer, Hsieh, Ng, Korfmacher, & Caprioli, 2003). From then, the technique has been applied to the imaging of a number of drugs. Prideaux et al. applied MALDI-MRM-MSI to visualize the distribution of the second line tuberculosis drug moxifloxacin at a range of time points after dosing in tuberculosis-infected rabbits, which was further validated by quantitative LC-MS/MS of lung and granuloma extracts from adjacent biopsies taken from the same animals (Prideaux et al., 2011). Drug distribution within the granulomas was observed to be inhomogeneous, and very low levels were observed in the caseum in comparison to the cellular granuloma regions (Prideaux et al., 2011). Fehniger et al. measured the occurrence of inhaled bronchodilator, ipratropium, within human bronchial biopsies obtained by fiber optic bronchoscopy shortly after dosing exposure and showed the drug is rapidly absorbed into the airway wall (Fehniger et al., 2011). Nilson et al. provided the first evidence that compounds administrated by inhaled delivery at standard pharmacological dosage can be quantitatively detected by MALDI-MSI with accuracy and precision (Nilsson et al., 2010). The distribution of the inhaled drug tiotropium was tracked and quantified in the lungs of dosed rats and a concentration gradient (80 fmol–5 pmol) away from the central airways into the lung parenchyma and pleura was observed (Nilsson et al., 2010). Hsieh et al. visualized the spatial distribution of astemizole (withdrawn in most countries due to its side effects) and its primary metabolite in rat brain tissues and showed that the drug alone is likely to be responsible for the associated central nervous system side effects upon elevated exposure (Hsieh, Li, & Korfmacher, 2010). Using MALDI-MSI the specific distribution of unlabeled chloroquine was examined in the retinas of rats and was found to be similar to autoradiograms results reported previously (Yamada, Hidefumi, Shion, Oshikata, & Haramaki, 2011).

Atkinson et al. showed the distribution of the bioreductive anticancer drug AQ4N (banoxatrone), its active metabolite AQ4, and ATP in treated human tumor xenografts (Atkinson, Loadman, Sutton, Patterson, & Clench, 2007). The distribution of ATP was found similar to that of AQ4N, i.e., in regions of abundant ATP there was no evidence of conversion of AQ4N into AQ4, indicating that the cytotoxic metabolite AQ4 is confined to hypoxic regions of the tumor as intended (Atkinson et al., 2007). Using

MALDI-FT-ICR, Cornett et al. imaged the antitumor drug imatinib and two of its metabolites from a mouse brain glioma and showed the metabolites to be more abundant in tumor region compared to normal, which is consistent with other imaging studies of the drug (Cornett, Frappier, & Caprioli, 2008). In order to better understand the penetration of the anticancer drug oxaliplatin, MALDI-MSI of tissue sections from treated rat kidneys was performed and it was observed that the drug and its metabolites were localized exclusively in the kidney cortex, suggesting the drug did not penetrate deeply into the organ (Bouslimani, Bec, Glueckmann, Hirtz, & Larroque, 2010). Sugihara et al. demonstrated the first data on the localization of personalized medicine (vemurafenib) within tumor compartments of malignant melanoma (MM). In a proof-of-concept in vitro study the overexpression and localization of the drug in the MM was shown using MS fragment ion signatures (Sugihara et al., 2014). Using MALDI-MSI, it has been shown the drug irinotecan reaches tumors mainly through microvessels whereas the conversion of the drug into its active form SN-38 in tumors of colorectal cancer was poor compared to normal tissue (Buck et al., 2015). The distribution of tamoxifen in both ER-positive and ER-negative breast cancer tumor tissues as analyzed by MALDI-MSI detected the drug at significantly lower intensities in tumor cells compared with stroma in ER-negative samples (Végvári et al., 2016). Aikawa et al. employed a quantitative MSI method (a combination of MALDI-MSI and LC-MS/MS) in a preclinical mouse model to evaluate the intrabrain distribution of an anticancer drug (Aikawa et al., 2016). While no differences were observed between the mice for the drug concentration, diffuse drug distribution was found in the brain of the knockout mice vs wild mice (Aikawa et al., 2016).

To remove the background signals from matrix degradation observed in MALDI-MSI, the method nanoparticle-assisted laser desorption ionization (n-PALDI) was developed which utilises nanoparticles as matrixes. It was found the protocol could be used to investigate the distribution of anticancer agents in primary tumors and in metastasis, to ascertain whether resistance is related to inadequate drug penetration in poorly vascularized parts of the tumor (Morosi et al., 2013). Using the n-PALDI protocol, the different distribution of paclitaxel in tumor and normal tissues was visualized and related to the dosage-schedules and pathological features of the tumors (Morosi et al., 2013).

5.2 MALDI-MSI on Whole Body Sections

Rohner et al. first demonstrated the MALDI-MSI of drugs in whole body animal sections in 2005 (Rohner, Staab, & Stoeckli, 2005). Since then, a wide range of drugs have been imaged in whole body animal sections, allowing the label-free tracking of both endogenous and exogenous compounds with spatial resolution and molecular specificity (Khatib-Shahidi, Andersson, Herman, Gillespie, & Caprioli, 2006). MALDI-MSI has been used to simultaneously detect drugs and their individual metabolite distributions at various time points across whold body tissue sections (Khatib-Shahidi et al., 2006). Stoeckli et al. measured the distribution of a ^{14}C labeled compound (dosed intratracheally) in whole rat tissue sections using MALDI-MSI and whole body autoradiography (WBA) which displayed good quantitative agreement between both techniques (Stoeckli, Staab, & Schweitzer, 2007). Trim et al. compared the distribution of vinblastine within whole body sections using MALDI-MSI and WBA which displayed MALDI-MSI to be advantageous by separating the drug from an endogenous isobaric lipid (Trim et al., 2008).

5.3 MALDI-MSI on 3D Tissue Cultures

Liu et al. applied MALDI-MSI to the analysis of 3D spheroids in order to assess the distribution of pharmaceuticals and their metabolites. As a proof-of-concept study, MALDI-MSI was applied to in HCT 116 colon carcinoma multicellular spheroids to analyze the distribution of irinotecan, showing the time-dependent penetration of the drug and three metabolites (Liu, Weaver, & Hummon, 2013).

6. DATA ANALYSIS

6.1 Spatial Information

The main advantage of MSI over other methods is the access it provides to spatial information. Possibly the most common approach to leveraging the spatial information in MSI data is simply in displaying results as spatial plots. Fig. 4 shows the results of several analyses of such data. These analyses are blind to the spatial information in the data but their results have spatial patterns that can be easily seen when plotted. Plotting such results next to a stained image of the tissue allows for the molecular information in the MSI data

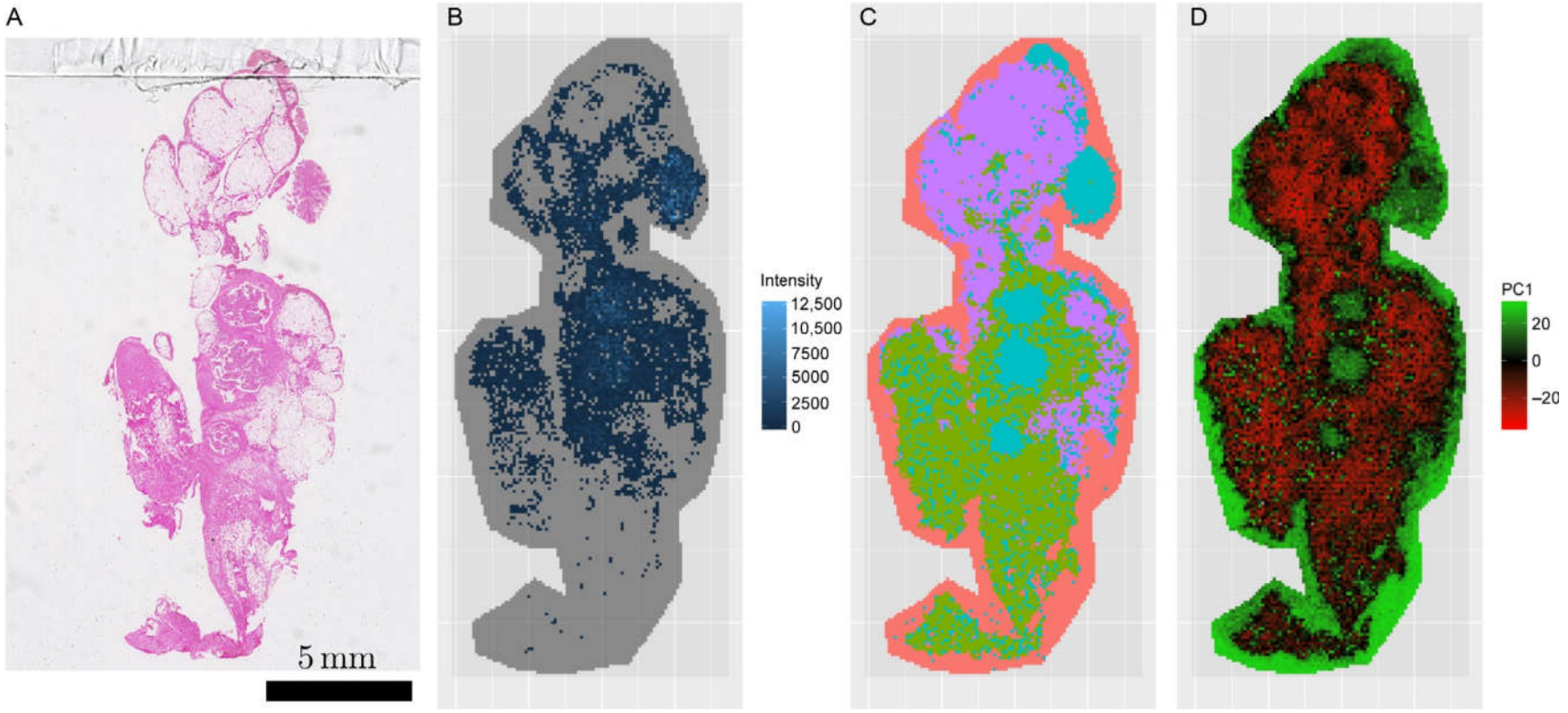


Fig. 4 Results of various analyses of tryptic peptide MALDI-MSI data collected from the same section of tissue. (A) H&E stained section of several primary serous ovarian tumors embedded in surrounding stromal tissue. (B) Intensity of a single peptide peak at $m/z = 1628.8$ confirmed as belong to heterogeneous nuclear ribonucleoprotein A1 which has been shown to be of interest in the past (Chen, Zhang, & Manley, 2010; Lee et al., 2010). (C) Results of 4-means clustering on the log-intensities including all measured *m/z* values. (D) Results of the first principle component including all measured *m/z* values on the log-intensities.

to be correlated to the histology of the tissue, thereby allowing biologically relevant use of the spatial information. Such results can even be overlayed directly with stained images (Walch, Rauser, Deininger, & Höfler, 2008), and used to partially automate and aid in histological annotations in an unbiased manner (Cornett et al., 2006). A popular alternative to the k-means clustering shown in Fig. 4 is the so-called semisupervised hierarchical clustering approach applied by Deininger et al. (2008). These semisupervised methods allow the user to manipulate the clustering results until they agree with the observed histology. Jones et al. (2011) demonstrated how multiple approaches to multivariate analysis, such as k-means and principle component analysis (PCA) (Fig. 4), could be combined in an automated manner to produce more reliable partitioning of spatial regions based on the multivariate MSI data.

Spatial information can also be applied to the partitioning of data in a histology driven fashion. To use Fig. 4 as an example, one could use the histology shown in the H&E stain to distinguish different tissue types such as tumor and surrounding stroma, and then by matching the image to the acquired region, partition the spectra into groups representing specific tissue. These groups of spectra could then be compared in an attempt to detect molecular differences between tissue types—this is also a very common approach used in the analysis of MSI data. Using this technique Oppenheimer et al. (2010) demonstrated that histologically healthy tissues within a margin of tumor tissues share molecular characteristics, and hypothesized that this phenomena could be involved in tumor recurrence postresection. Such information could be used by surgeons to identify an appropriate margin to excise the tissue surrounding a tumor.

6.2 Preprocessing: Peak Detection

MSI data have a uniquely complicated structure due to the spatial information it contains, and additionally can be computationally challenging due to the shear amount of information recorded during acquisition. The data therefore require novel approaches to analysis when compared to other, more traditional forms of mass spectrometry (Jones, Deininger, Hogendoorn, Deelder, & McDonnell, 2012). Furthermore, the appropriate approach will vary depending on the application. One approach to addressing the quantity of data is to perform some, preferably computationally fast, data reduction early in the data analysis workflow and then work

with the reduced data. This allows for analyses to be performed quickly, but also means that all further analyses will be affected by the choices made in the data reduction step. The most simplistic approach to data reduction is centroiding—essentially reducing spectra to their local maxima. Centroiding is perhaps the most simplistic approach to peak detection. Essentially centroiding involves reducing the acquired spectra to their local maxima, allowing peaks in the spectra to be identified, information relevant to these peaks to be stored and the remainder of the "noise" to be removed. Yang, He, and Yu (2009) provide a good overview of peak detection in the context of MALDI-MS. Wavelet approaches to peak detection, such as those suggested by Du, Kibbe, and Lin (2006), Lange, Gropl, Reinert, Kohlbacher, and Hildebrandt (2006), show some promise but have not yet made it into mainstream software packages, and so do not currently see much use. Regardless, peak detection prior to further analysis is a useful tool as it can reduce the volume of data by as much as three orders of magnitude. However, care must be taken when interpreting final results; it is important to be aware that all analyses will be affected by any peak detection steps and that a quality control analysis of the data following peak detection should be implemented to ensure its performance.

6.3 Classification of FFPE-TMAs and the Importance of Dimension Reduction

The application of MSI to FFPE-TMAs for classification of various diagnostic and prognostic clinical factors is a particularly promising field of research. The ability to use FFPE tissue allows access to large archives including tissue samples and associated patient meta data (Hood et al., 2005). The use of TMAs with MSI facilitates the collection of label-free and untargeted mass data from large numbers of patients (Groseclose et al., 2008), which in turn allows for classification problems to be considered. One of the key obstacles to the classification of such data is that, due to the untargeted nature of MSI, the resulting data are of high-dimensionality. The natural approach to address the dimensionality of such data is to perform some dimension reduction prior to classification. Mascini et al. (2015) proposed the use PCA for dimension reduction prior to classification. PCA is a classical multivariate approach that preserves the maximum amount of variability in the dimension-reduced data. Winderbaum et al. (2016) proposed using a cross correlation analysis (CCA) to rank MSI variables based on their contribution to the highest multivariate correlation to the clinical variable and to reduce the dimension of the data by restricting to only the most highly ranked MSI

variables. Either way, dimension reduction prior to classification is crucial. In order to apply these ideas in clinical practice, large cohort studies and/or clinical trials would need to be carried out to validate the performance of these MS-based classifiers. To date most studies have been exploratory, using small samples of patients as a proof of principle, with a study by Steurer et al. (2013) being a notable exception. Dependent on large scale validation studies, classification of MSI data could potentially be used in the individualisation of treatments, predicting factors such as LNM or potentially even chemotherapy response.

7. CONCLUDING REMARKS

7.1 The Future of MSI: Molecular Pathology

For over 100 years clinical pathology has relied upon a description of morphology by trained specialists for tumor classification and grading. The information that can be gained from such an analyses is limited, however, and is often purely descriptive. During the last decade mass spectrometers have found their way from physics and biochemistry research laboratories into clinical pathology laboratories, bringing together researchers from a wide spectrum of disciplines, such as physics, biochemistry, medicine, pathology, and mathematics. A key feature of MSI is the ability to discern molecular signatures of disease with the additional bonus that the features can be directly related back to the tissue morphology. In the context of clinical oncology, the application of MSI has the potential to shift traditional pathology from descriptive morphological analyses to detailed molecular analyses. As such, the introduction of MSI equipment into pathology laboratories will extend their capabilities and revolutionize management and treatment of cancer.

REFERENCES

Abbassi-Ghadi, N., Veselkov, K., Kumar, S., Huang, J., Jones, E., Strittmatter, N., et al. (2014). Discrimination of lymph node metastases using desorption electrospray ionisation-mass spectrometry imaging. *Chemical Communications (Cambridge, England)*, *50*(28), 3661–3664.

Abbott, K. L., Lim, J. M., Wells, L., Benigno, B. B., McDonald, J. F., & Pierce, M. (2010). Identification of candidate biomarkers with cancer-specific glycosylation in the tissue and serum of endometrioid ovarian cancer patients by glycoproteomic analysis. *Proteomics*, *10*(3), 470–481.

Aerni, H.-R., Cornett, D. S., & Caprioli, R. M. (2006). Automated acoustic matrix deposition for MALDI sample preparation. *Analytical Chemistry*, *78*(3), 827–834.

Agar, N. Y., Malcolm, J. G., Mohan, V., Yang, H. W., Johnson, M. D., Tannenbaum, A., et al. (2010). Imaging of meningioma progression by matrix-assisted laser desorption ionization time-of-flight mass spectrometry. *Analytical Chemistry*, *82*(7), 2621–2625.

Aichler, M., Elsner, M., Ludyga, N., Feuchtinger, A., Zangen, V., Maier, S. K., et al. (2013). Clinical response to chemotherapy in oesophageal adenocarcinoma patients is linked to defects in mitochondria. *The Journal of Pathology*, *230*(4), 410–419.

Aichler, M., & Walch, A. (2015). MALDI Imaging mass spectrometry: Current frontiers and perspectives in pathology research and practice. *Laboratory Investigation*, *95*(4), 422–431.

Aikawa, H., Hayashi, M., Ryu, S., Yamashita, M., Ohtsuka, N., Nishidate, M., et al. (2016). Visualizing spatial distribution of alectinib in murine brain using quantitative mass spectrometry imaging. *Scientific Reports*, *6*, 23749.

Andersson, M., Andren, P., & Caprioli, R. M. (2010). MALDI imaging and profiling mass spectrometry in neuroproteomics. In O. Alzate (Ed.), *Neuroproteomics*. Boca Raton, FL: CRC Press/Taylor & Francis.

Andersson, M., Groseclose, M. R., Deutch, A. Y., & Caprioli, R. M. (2008). Imaging mass spectrometry of proteins and peptides: 3D volume reconstruction. *Nature Methods*, *5*(1), 101–108.

Atkinson, S. J., Loadman, P. M., Sutton, C., Patterson, L. H., & Clench, M. R. (2007). Examination of the distribution of the bioreductive drug AQ4N and its active metabolite AQ4 in solid tumours by imaging matrix-assisted laser desorption/ionisation mass spectrometry. *Rapid Communications in Mass Spectrometry*, *21*(7), 1271–1276.

Balluff, B., Elsner, M., Kowarsch, A., Rauser, S., Meding, S., Schuhmacher, C., et al. (2010). Classification of HER2/neu status in gastric cancer using a breast-cancer derived proteome classifier. *Journal of Proteome Research*, *9*(12), 6317–6322.

Balluff, B., Frese, C. K., Maier, S. K., Schone, C., Kuster, B., Schmitt, M., et al. (2015). De novo discovery of phenotypic intratumour heterogeneity using imaging mass spectrometry. *The Journal of Pathology*, *235*(1), 3–13.

Balluff, B., Rauser, S., Meding, S., Elsner, M., Schone, C., Feuchtinger, A., et al. (2011). MALDI imaging identifies prognostic seven-protein signature of novel tissue markers in intestinal-type gastric cancer. *The American Journal of Pathology*, *179*(6), 2720–2729.

Balluff, B., Schöne, C., Höfler, H., & Walch, A. (2011). MALDI imaging mass spectrometry for direct tissue analysis: Technological advancements and recent applications. *Histochemistry and Cell Biology*, *136*(3), 227–244.

Bauer, J. A., Chakravarthy, A. B., Rosenbluth, J. M., Mi, D., Seeley, E. H., De Matos Granja-Ingram, N., et al. (2010). Identification of markers of taxane sensitivity using proteomic and genomic analyses of breast tumors from patients receiving neoadjuvant paclitaxel and radiation. *Clinical Cancer Research*, *16*(2), 681–690.

Benabdellah, F., Seyer, A., Quinton, L., Touboul, D., Brunelle, A., & Laprevote, O. (2010). Mass spectrometry imaging of rat brain sections: Nanomolar sensitivity with MALDI versus nanometer resolution by TOF-SIMS. *Analytical and Bioanalytical Chemistry*, *396*(1), 151–162.

Bodzon-Kulakowska, A., & Suder, P. (2016). Imaging mass spectrometry: Instrumentation, applications, and combination with other visualization techniques. *Mass Spectrometry Reviews*, *35*(1), 147–169.

Bouslimani, A., Bec, N., Glueckmann, M., Hirtz, C., & Larroque, C. (2010). Matrix-assisted laser desorption/ionization imaging mass spectrometry of oxaliplatin derivatives in heated intraoperative chemotherapy (HIPEC)-like treated rat kidney. *Rapid Communications in Mass Spectrometry*, *24*(4), 415–421.

Buck, A., Halbritter, S., Spath, C., Feuchtinger, A., Aichler, M., Zitzelsberger, H., et al. (2015). Distribution and quantification of irinotecan and its active metabolite SN-38

in colon cancer murine model systems using MALDI MSI. *Analytical and Bioanalytical Chemistry*, *407*(8), 2107–2116.

Büll, C., Boltje, T. J., Wassink, M., de Graaf, A. M., van Delft, F. L., den Brok, M. H., et al. (2013). Targeting aberrant sialylation in cancer cells using a fluorinated sialic acid analog impairs adhesion, migration, and in vivo tumor growth. *Molecular Cancer Therapeutics*, *12*(10), 1935–1946.

Calligaris, D., Caragacianu, D., Liu, X., Norton, I., Thompson, C. J., Richardson, A. L., et al. (2014). Application of desorption electrospray ionization mass spectrometry imaging in breast cancer margin analysis. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(42), 15184–15189.

Caprioli, R. M., Farmer, T. B., & Gile, J. (1997). Molecular imaging of biological samples: Localization of peptides and proteins using MALDI-TOF MS. *Analytical Chemistry*, *69*(23), 4751–4760.

Casadonte, R., & Caprioli, R. M. (2011). Proteomic analysis of formalin-fixed paraffin-embedded tissue by MALDI imaging mass spectrometry. *Nature Protocols*, *6*(11), 1695–1709.

Casadonte, R., Kriegsmann, M., Zweynert, F., Friedrich, K., Bretton, G., Otto, M., et al. (2014). Imaging mass spectrometry to discriminate breast from pancreatic cancer metastasis in formalin-fixed paraffin-embedded tissues. *Proteomics*, *14*(7–8), 956–964.

Cazares, L. H., Troyer, D., Mendrinos, S., Lance, R. A., Nyalwidhe, J. O., Beydoun, H. A., et al. (2009). Imaging mass spectrometry of a specific fragment of mitogen-activated protein kinase/extracellular signal-regulated kinase kinase kinase 2 discriminates cancer from uninvolved prostate tissue. *Clinical Cancer Research*, *15*(17), 5541–5551.

Chaurand, P., Sanders, M. E., Jensen, R. A., & Caprioli, R. M. (2004). Proteomics in diagnostic pathology: Profiling and imaging proteins directly in tissue sections. *The American Journal of Pathology*, *165*(4), 1057–1068.

Chaurand, P., Schwartz, S. A., & Caprioli, R. M. (2004). Assessing protein patterns in disease using imaging mass spectrometry. *Journal of Proteome Research*, *3*(2), 245–252.

Chen, M., Zhang, J., & Manley, J. L. (2010). Turning on a fuel switch of cancer: hnRNP proteins regulate alternative splicing of pyruvate kinase mRNA. *Cancer Research*, *70*(22), 8977–8980.

Cho, Y. T., Su, H., Wu, W. J., Wu, D. C., Hou, M. F., Kuo, C. H., et al. (2015). Biomarker characterization by MALDI-TOF/MS. *Advances in Clinical Chemistry*, *69*, 209–254.

Choi, J. H., Shin, N. R., Moon, H. J., Kwon, C. H., Kim, G. H., Song, G. A., et al. (2012). Identification of S100A8 and S100A9 as negative regulators for lymph node metastasis of gastric adenocarcinoma. *Histology and Histopathology*, *27*(11), 1439–1448.

Coombes, K. R., Koomen, J. M., Baggerly, K. A., Morris, J. S., & Kobayashi, R. (2005). Understanding the characteristics of mass spectrometry data through the use of simulation. *Cancer Informatics*, *1*, 41–52.

Cornett, D. S., Frappier, S. L., & Caprioli, R. M. (2008). MALDI-FTICR imaging mass spectrometry of drugs and metabolites in tissue. *Analytical Chemistry*, *80*(14), 5648–5653.

Cornett, D. S., Mobley, J. A., Dias, E. C., Andersson, M., Arteaga, C. L., Sanders, M. E., et al. (2006). A novel histology-directed strategy for MALDI-MS tissue profiling that improves throughput and cellular specificity in human breast cancer. *Molecular & Cellular Proteomics*, *5*(10), 1975–1983.

Crecelius, A. C., Cornett, D. S., Caprioli, R. M., Williams, B., Dawant, B. M., & Bodenheimer, B. (2005). Three-dimensional visualization of protein expression in mouse brain structures using imaging mass spectrometry. *Journal of the American Society for Mass Spectrometry*, *16*(7), 1093–1099.

de Hoffmann, E., & Stroobant, V. (2007). *Mass spectrometry: Principles and applications* (3rd ed.). England: Wiley.

Deininger, S.-O., Ebert, M. P., Fütterer, A., Gerhard, M., & Röcken, C. (2008). MALDI imaging combined with hierarchical clustering as a new tool for the interpretation of complex human cancers. *Journal of Proteome Research*, 7(12), 5230–5236.

Dekker, T. J. A., Balluff, B. D., Jones, E. A., Schöne, C. D., Schmitt, M., Aubele, M., et al. (2014). Multicenter matrix-assisted laser desorption/ionization mass spectrometry imaging (MALDI MSI) identifies proteomic differences in breast-cancer-associated stroma. *Journal of Proteome Research*, *13*(11), 4730–4738.

Diehl, H. C., Beine, B., Elm, J., Trede, D., Ahrens, M., Eisenacher, M., et al. (2015). The challenge of on-tissue digestion for MALDI MSI—A comparison of different protocols to improve imaging experiments. *Analytical and Bioanalytical Chemistry*, *407*(8), 2223–2243.

Dill, A. L., Eberlin, L. S., Costa, A. B., Zheng, C., Ifa, D. R., Cheng, L., et al. (2011). Multivariate statistical identification of human bladder carcinomas using ambient ionization imaging mass spectrometry. *Chemistry*, *17*(10), 2897–2902.

Dill, A. L., Ifa, D. R., Manicke, N. E., Costa, A. B., Ramos-Vara, J. A., Knapp, D. W., et al. (2009). Lipid profiles of canine invasive transitional cell carcinoma of the urinary bladder and adjacent normal tissue by desorption electrospray ionization imaging mass spectrometry. *Analytical Chemistry*, *81*(21), 8758–8764.

Djidja, M.-C., Claude, E., Snel, M. F., Francese, S., Scriven, P., Carolan, V., et al. (2010). Novel molecular tumour classification using MALDI–mass spectrometry imaging of tissue micro-array. *Analytical and Bioanalytical Chemistry*, *397*(2), 587–601.

Drake, P. M., Cho, W., Li, B., Prakobphol, A., Johansen, E., Anderson, N. L., et al. (2010). Sweetening the pot: adding glycosylation to the biomarker discovery equation. *Clinical Chemistry*, *56*(2), 223–236.

Du, P., Kibbe, W. A., & Lin, S. M. (2006). Improved peak detection in mass spectrum by incorporating continuous wavelet transform-based pattern matching. *Bioinformatics*, *22*(17), 2059–2065.

Dube, D. H., & Bertozzi, C. R. (2005). Glycans in cancer and inflammation—Potential for therapeutics and diagnostics. *Nature Reviews. Drug Discovery*, *4*(6), 477–488.

Eberlin, L. S., Norton, I., Dill, A. L., Golby, A. J., Ligon, K. L., Santagata, S., et al. (2012). Classifying human brain tumors by lipid imaging with mass spectrometry. *Cancer Research*, *72*(3), 645–654.

Eberlin, L. S., Tibshirani, R. J., Zhang, J., Longacre, T. A., Berry, G. J., Bingham, D. B., et al. (2014). Molecular assessment of surgical-resection margins of gastric cancer by mass-spectrometric imaging. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(7), 2436–2441.

Elsner, M., Rauser, S., Maier, S., Schöne, C., Balluff, B., Meding, S., et al. (2012). MALDI imaging mass spectrometry reveals COX7A2, TAGLN2 and S100-A10 as novel prognostic markers in Barrett's adenocarcinoma. *Journal of Proteomics*, *75*(15), 4693–4704.

Fehniger, T. E., Vegvari, A., Rezeli, M., Prikk, K., Ross, P., Dahlback, M., et al. (2011). Direct demonstration of tissue uptake of an inhaled drug: Proof-of-principle study using matrix-assisted laser desorption ionization mass spectrometry imaging. *Analytical Chemistry*, *83*(21), 8329–8336.

Fejzo, M. S., & Slamon, D. J. (2010). Tissue microarrays from frozen tissues-OCT technique. *Methods in Molecular Biology*, *664*, 73–80.

Fentz, J. S., Zornig, C., Juhl, H. H., & David, K. A. (2004). Tissue ischemia time affects gene and protein expression patterns within minutes following surgical tumor excision. *Biotechniques*, *36*(6), 1030–1037.

Fletcher, J. S., Lockyer, N. P., & Vickerman, J. C. (2011). Developments in molecular SIMS depth profiling and 3D imaging of biological systems using polyatomic primary ions. *Mass Spectrometry Reviews*, *30*(1), 142–174.

Fletcher, J. S., Vickerman, J. C., & Winograd, N. (2011). Label free biochemical 2D and 3D imaging using secondary ion mass spectrometry. *Current Opinion in Chemical Biology*, *15*(5), 733–740.

Gemoll, T., Strohkamp, S., Schillo, K., Thorns, C., & Habermann, J. K. (2015). MALDI-imaging reveals thymosin beta-4 as an independent prognostic marker for colorectal cancer. *Oncotarget*, *6*(41), 43869–43880.

Gobom, J., Mueller, M., Egelhofer, V., Theiss, D., Lehrach, H., & Nordhoff, E. (2002). A calibration method that simplifies and improves accurate determination of peptide molecular masses by MALDI-TOF MS. *Analytical Chemistry*, *74*(15), 3915–3923.

Goodwin, R. J. (2012). Sample preparation for mass spectrometry imaging: Small mistakes can lead to big consequences. *Journal of Proteomics*, *75*(16), 4893–4911.

Gorzolka, K., & Walch, A. (2014). MALDI mass spectrometry imaging of formalin-fixed paraffin-embedded tissues in clinical research. *Histology and Histopathology*, *29*(11), 1365–1376.

Groseclose, M. R., Massion, P. P., Chaurand, P., & Caprioli, R. M. (2008). High-throughput proteomic analysis of formalin-fixed paraffin-embedded tissue microarrays using MALDI imaging mass spectrometry. *Proteomics*, *8*(18), 3715–3724.

Grüner, B. M., Hahne, H., Mazur, P. K., Trajkovic-Arsic, M., Maier, S., Esposito, I., et al. (2012). MALDI imaging mass spectrometry for in situ proteomic analysis of preneoplastic lesions in pancreatic cancer. *PLoS One*, *7*(6), e39424.

Gustafsson, O. J., Briggs, M. T., Condina, M. R., Winderbaum, L. J., Pelzing, M., McColl, S. R., et al. (2015). MALDI imaging mass spectrometry of N-linked glycans on formalin-fixed paraffin-embedded murine kidney. *Analytical and Bioanalytical Chemistry*, *407*(8), 2127–2139.

Gustafsson, J. O., Eddes, J. S., Meding, S., Koudelka, T., Oehler, M. K., McColl, S. R., et al. (2012). Internal calibrants allow high accuracy peptide matching between MALDI imaging MS and LC-MS/MS. *Journal of Proteomics*, *75*(16), 5093–5105.

Gustafsson, O. J., Eddes, J. S., Meding, S., McColl, S. R., Oehler, M. K., & Hoffmann, P. (2013). 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*, *27*(6), 655–670.

Han, E. C., Lee, Y.-S., Liao, W.-S., Liu, Y.-C., Liao, H.-Y., & Jeng, L.-B. (2011). Direct tissue analysis by MALDI-TOF mass spectrometry in human hepatocellular carcinoma. *Clinica Chimica Acta*, *412*(3–4), 230–239.

Hardesty, W. M., Kelley, M. C., Mi, D., Low, R. L., & Caprioli, R. M. (2011). Protein signatures for survival and recurrence in metastatic melanoma. *Journal of Proteomics*, *74*(7), 1002–1014.

Holst, S., Heijs, B., de Haan, N., van Zeijl, R. J., Briaire-de Bruijn, I. H., van Pelt, G. W., et al. (2016). Linkage-specific in situ sialic acid derivatization for N-glycan mass spectrometry imaging of formalin-fixed paraffin-embedded tissues. *Analytical Chemistry*, *88*(11), 5904–5913.

Hood, B. L., Darfler, M. M., Guiel, T. G., Furusato, B., Lucas, D. A., Ringeisen, B. R., et al. (2005). Proteomic analysis of formalin-fixed prostate cancer tissue. *Molecular & Cellular Proteomics*, *4*(11), 1741–1753.

Horn, D. M., Peters, E. C., Klock, H., Meyers, A., & Brock, A. (2004). Improved protein identification using automated high mass measurement accuracy MALDI FT-ICR MS peptide mass fingerprinting. *International Journal of Mass Spectrometry*, *238*(2), 189–196.

Hsieh, Y., Li, F., & Korfmacher, W. A. (2010). Mapping pharmaceuticals in rat brain sections using MALDI imaging mass spectrometry. *Methods in Molecular Biology*, *656*, 147–158.

Hsu, C. C., & Dorrestein, P. C. (2015). Visualizing life with ambient mass spectrometry. *Current Opinion in Biotechnology*, *31*, 24–34.

Jensen, P. H., Karlsson, N. G., Kolarich, D., & Packer, N. H. (2012). Structural analysis of N- and O-glycans released from glycoproteins. *Nature Protocols*, *7*(7), 1299–1310.

Jones, E. A., Deininger, S.-O., Hogendoorn, P. C., Deelder, A. M., & McDonnell, L. A. (2012). Imaging mass spectrometry statistical analysis. *Journal of Proteomics*, *75*(16), 4962–4989.

Jones, E. A., van Remoortere, A., van Zeijl, R. J. M., Hogendoorn, P. C. W., Bovée, J. V. M. G., Deelder, A. M., et al. (2011). Multiple statistical analysis techniques corroborate intratumor heterogeneity in imaging mass spectrometry datasets of myxofibrosarcoma. *PLoS One*, *6*(9), e24913.

Jurchen, J. C., Rubakhin, S. S., & Sweedler, J. V. (2005). MALDI-MS imaging of features smaller than the size of the laser beam. *Journal of the American Society for Mass Spectrometry*, *16*(10), 1654–1659.

Kaletas, B. K., van der Wiel, I. M., Stauber, J., Guzel, C., Kros, J. M., Luider, T. M., et al. (2009). Sample preparation issues for tissue imaging by imaging MS. *Proteomics*, *9*(10), 2622–2633.

Kang, S., Shim, H. S., Lee, J. S., Kim, D. S., Kim, H. Y., Hong, S. H., et al. (2010). Molecular proteomics imaging of tumor interfaces by mass spectrometry. *Journal of Proteome Research*, *9*(2), 1157–1164.

Kertesz, V., Berkel, V., Gary, J., Vavrek, M., Koeplinger, K. A., Schneider, B. B., et al. (2008). Comparison of drug distribution images from whole-body thin tissue sections obtained using desorption electrospray ionization tandem mass spectrometry and autoradiography. *Analytical Chemistry*, *80*(13), 5168–5177.

Khatib-Shahidi, S., Andersson, M., Herman, J. L., Gillespie, T. A., & Caprioli, R. M. (2006). Direct molecular analysis of whole-body animal tissue sections by imaging MALDI mass spectrometry. *Analytical Chemistry*, *78*(18), 6448–6456.

Kim, H. K., Reyzer, M. L., Choi, I. J., Kim, C. G., Kim, H. S., Oshima, A., et al. (2010). Gastric cancer-specific protein profile identified using endoscopic biopsy samples via MALDI mass spectrometry. *Journal of Proteome Research*, *9*(8), 4123–4130.

Klerk, L. A., Maarten Altelaar, A. F., Froesch, M., McDonnell, L. A., & Heeren, R. M. A. (2009). Fast and automated large-area imaging MALDI mass spectrometry in microprobe and microscope mode. *International Journal of Mass Spectrometry*, *285*(1–2), 19–25.

Kriegsmann, J., Kriegsmann, M., & Casadonte, R. (2015). MALDI TOF imaging mass spectrometry in clinical pathology: A valuable tool for cancer diagnostics (review). *International Journal of Oncology*, *46*(3), 893–906.

Kwon, H. J., Kim, Y., Sugihara, Y., Baldetorp, B., Welinder, C., Watanabe, K., et al. (2015). Drug compound characterization by mass spectrometry imaging in cancer tissue. *Archives of Pharmacal Research*, *38*(9), 1718–1727.

Lange, E., Gropl, C., Reinert, K., Kohlbacher, O., & Hildebrandt, A. (2006). High-accuracy peak picking of proteomics data using wavelet techniques. *Pacific Symposium on Biocomputing*, *11*, 243–254. Paper presented at the Pacific Symposium on Biocomputing.

Lanni, E. J., Rubakhin, S. S., & Sweedler, J. V. (2012). Mass spectrometry imaging and profiling of single cells. *Journal of Proteomics*, *75*(16), 5036–5051.

Lau, K. S., & Dennis, J. W. (2008). N-Glycans in cancer progression. *Glycobiology*, *18*(10), 750–760.

Le Faouder, J., Laouirem, S., Chapelle, M., Albuquerque, M., Belghiti, J., Degos, F., et al. (2011). Imaging mass spectrometry provides fingerprints for distinguishing hepatocellular carcinoma from cirrhosis. *Journal of Proteome Research*, *10*(8), 3755–3765.

Lee, D. H., Chung, K., Song, J.-A., Kim, T.-h., Kang, H., Huh, J. H., et al. (2010). Proteomic identification of paclitaxel-resistance associated hnRNP A2 and GDI 2 proteins in human ovarian cancer cells. *Journal of Proteome Research*, *9*(11), 5668–5676.

Lemaire, R., Ait Menguellet, S., Stauber, J., Marchaudon, V., Lucot, J.-P., Collinet, P., et al. (2007). Specific MALDI imaging and profiling for biomarker hunting and validation: Fragment of the 11S proteasome activator complex, Reg alpha fragment, is a new potential ovary cancer biomarker. *Journal of Proteome Research*, *6*(11), 4127–4134.

Lemaire, R., Desmons, A., Tabet, J. C., Day, R., Salzet, M., & Fournier, I. (2007). Direct analysis and MALDI imaging of formalin-fixed, paraffin-embedded tissue sections. *Journal of Proteome Research*, *6*(4), 1295–1305.

Lemaire, R., Tabet, J. C., Ducoroy, P., Hendra, J. B., Salzet, M., & Fournier, I. (2006). Solid ionic matrixes for direct tissue analysis and MALDI imaging. *Analytical Chemistry*, *78*(3), 809–819.

Liu, X., Weaver, E. M., & Hummon, A. B. (2013). Evaluation of therapeutics in three-dimensional cell culture systems by MALDI imaging mass spectrometry. *Analytical Chemistry*, *85*(13), 6295–6302.

Luxembourg, S. L., Mize, T. H., McDonnell, L. A., & Heeren, R. M. (2004). High-spatial resolution mass spectrometric imaging of peptide and protein distributions on a surface. *Analytical Chemistry*, *76*(18), 5339–5344.

Mascini, N. E., Eijkel, G. B., ter Brugge, P., Jonkers, J., Wesseling, J., & Heeren, R. M. A. (2015). The use of mass spectrometry imaging to predict treatment response of patient-derived xenograft models of triple-negative breast cancer. *Journal of Proteome Research*, *14*(2), 1069–1075.

McDonnell, L. A., Corthals, G. L., Willems, S. M., van Remoortere, A., van Zeijl, R. J. M., & Deelder, A. M. (2010). Peptide and protein imaging mass spectrometry in cancer research. *Journal of Proteomics*, *73*(10), 1921–1944.

Metz, B., Kersten, G. F. A., Hoogerhout, P., Brugghe, H. F., Timmermans, H. A. M., De Jong, A. D., et al. (2004). Identification of formaldehyde-induced modifications in proteins reactions with model peptides. *Journal of Biological Chemistry*, *279*(8), 6235–6243.

Minchinton, A. I., & Tannock, I. F. (2006). Drug penetration in solid tumours. *Nature Reviews. Cancer*, *6*(8), 583–592. http://dx.doi.org/10.1038/nrc1893.

Mittal, P., Klingler-Hoffmann, M., Arentz, G., Winderbaum, L., Lokman, N. A., Zhang, C., et al. (2016). Maldi imaging of primary endometrial cancers reveals proteins associated with lymph node metastasis. *Proteomics*, *16*, 1793–1801.

Mittal, P., Klingler-Hoffmann, M., Arentz, G., Zhang, C., Kaur, G., Oehler, M. K., et al. (2016). Proteomics of endometrial cancer diagnosis, treatment, and prognosis. *Proteomics. Clinical Applications*, *10*(3), 217–229.

Morita, Y., Ikegami, K., Goto-Inoue, N., Hayasaka, T., Zaima, N., Tanaka, H., et al. (2010). Imaging mass spectrometry of gastric carcinoma in formalin-fixed paraffin-embedded tissue microarray. *Cancer Science*, *101*(1), 267–273.

Morosi, L., Spinelli, P., Zucchetti, M., Pretto, F., Carrà, A., D'Incalci, M., et al. (2013). Determination of paclitaxel distribution in solid tumors by nano-particle assisted laser desorption ionization mass spectrometry imaging. *PLoS One*, *8*(8), e72532.

Nakagoe, T., Fukushima, K., Nanashima, A., Sawai, T., Tsuji, T., Jibiki, M., et al. (2000). Expression of Lewis(a), sialyl Lewis(a), Lewis(x) and sialyl Lewis(x) antigens as prognostic factors in patients with colorectal cancer. *Canadian Journal of Gastroenterology*, *14*(9), 753–760.

Nilsson, A., Fehniger, T. E., Gustavsson, L., Andersson, M., Kenne, K., Marko-Varga, G., et al. (2010). Fine mapping the spatial distribution and concentration of unlabeled drugs within tissue micro-compartments using imaging mass spectrometry. *PLoS One*, *5*(7), e11411.

Nipp, M., Elsner, M., Balluff, B., Meding, S., Sarioglu, H., Ueffing, M., et al. (2011). S100-A10, thioredoxin, and S100-A6 as biomarkers of papillary thyroid carcinoma with lymph node metastasis identified by MALDI imaging. *Journal of Molecular Medicine*, *90*(2), 163–174.

North, S. J., Hitchen, P. G., Haslam, S. M., & Dell, A. (2009). Mass spectrometry in the analysis of N-linked and O-linked glycans. *Current Opinion in Structural Biology*, *19*(5), 498–506.

Oezdemir, R. F., Gaisa, N. T., Lindemann-Docter, K., Gostek, S., Weiskirchen, R., Ahrens, M., et al. (2012). Proteomic tissue profiling for the improvement of

grading of noninvasive papillary urothelial neoplasia. *Clinical Biochemistry*, *45*(1–2), 7–11.
Ohtsubo, K., & Marth, J. D. (2006). Glycosylation in cellular mechanisms of health and disease. *Cell*, *126*(5), 855–867.
Oppenheimer, S. R., Mi, D., Sanders, M. E., & Caprioli, R. M. (2010). Molecular analysis of tumor margins by MALDI mass spectrometry in renal carcinoma. *Journal of Proteome Research*, *9*(5), 2182–2190.
Pagni, F., Mainini, V., Garancini, M., Bono, F., Vanzati, A., Giardini, V., et al. (2015). Proteomics for the diagnosis of thyroid lesions: Preliminary report. *Cytopathology*, *26*(5), 318–324.
Pappin, D. J. C., Hojrup, P., & Bleasby, A. J. (1993). Rapid identification of proteins by peptide-mass fingerprinting. *Current Biology*, *3*(6), 327–332.
Passarelli, M. K., & Winograd, N. (2011). Lipid imaging with time-of-flight secondary ion mass spectrometry (ToF-SIMS). *Biochimica et Biophysica Acta*, *1811*(11), 976–990.
Patel, S. A., Barnes, A., Loftus, N., Martin, R., Sloan, P., Thakker, N., et al. (2009). Imaging mass spectrometry using chemical inkjet printing reveals differential protein expression in human oral squamous cell carcinoma. *Analyst*, *134*(2), 301–307. http://dx.doi.org/10.1039/B812533C.
Pavlidis, N., & Pentheroudakis, G. (2016). Cancer of unknown primary site. *The Lancet*, *379*(9824), 1428–1435.
Powers, T. W., Holst, S., Wuhrer, M., Mehta, A. S., & Drake, R. R. (2015). Two-dimensional N-glycan distribution mapping of hepatocellular carcinoma tissues by MALDI-imaging mass spectrometry. *Biomolecules*, *4*, 2554–2572.
Powers, T. W., Jones, E. E., Betesh, L. R., Romano, P. R., Gao, P., Copland, J. A., et al. (2013). Matrix assisted laser desorption ionization imaging mass spectrometry workflow for spatial profiling analysis of N-linked glycan expression in tissues. *Analytical Chemistry*, *85*(20), 9799–9806.
Powers, T. W., Neely, B. A., Shao, Y., Tang, H., Troyer, D. A., Mehta, A. S., et al. (2014). MALDI imaging mass spectrometry profiling of N-glycans in formalin-fixed paraffin embedded clinical tissue blocks and tissue microarrays. *PLoS One*, *9*(9), e106255.
Prideaux, B., Dartois, V., Staab, D., Weiner, D. M., Goh, A., Via, L. E., et al. (2011). High-sensitivity MALDI-MRM-MS imaging of moxifloxacin distribution in tuberculosis-infected rabbit lungs and granulomatous lesions. *Analytical Chemistry*, *83*(6), 2112–2118.
Rahman, S. M., Gonzalez, A. L., Li, M., Seeley, E. H., Zimmerman, L. J., Zhang, X. J., et al. (2011). Lung cancer diagnosis from proteomic analysis of preinvasive lesions. *Cancer Research*, *71*(8), 3009–3017.
Rauser, S., Marquardt, C., Balluff, B., Deininger, S.-O., Albers, C., Belau, E., et al. (2010). Classification of HER2 receptor status in breast cancer tissues by MALDI imaging mass spectrometry. *Journal of Proteome Research*, *9*(4), 1854–1863.
Renovanz, M., & Kim, E. L. (2014). Intratumoral heterogeneity, its contribution to therapy resistance and methodological caveats to assessment. *Frontiers in Oncology*, *4*, 142.
Reyzer, M. L., Hsieh, Y., Ng, K., Korfmacher, W. A., & Caprioli, R. M. (2003). Direct analysis of drug candidates in tissue by matrix-assisted laser desorption/ionization mass spectrometry. *Journal of Mass Spectrometry*, *38*(10), 1081–1092.
Rodrigo, M. A., Zitka, O., Krizkova, S., Moulick, A., Adam, V., & Kizek, R. (2014). MALDI-TOF MS as evolving cancer diagnostic tool: A review. *Journal of Pharmaceutical and Biomedical Analysis*, *95*, 245–255.
Rohner, T. C., Staab, D., & Stoeckli, M. (2005). MALDI mass spectrometric imaging of biological tissue sections. *Mechanisms of Ageing and Development*, *126*(1), 177–185.

Ronci, M., Bonanno, E., Colantoni, A., Pieroni, L., Di Ilio, C., Spagnoli, L. G., et al. (2008). Protein unlocking procedures of formalin-fixed paraffin-embedded tissues: Application to MALDI-TOF imaging MS investigations. *Proteomics*, *8*(18), 3702–3714.

Schuerenberg, M., & Deininger, S.-O. (2010). Matrix application with ImagePrep. *Imaging mass spectrometry* (pp. 87–91). New York, USA: Springer.

Schwamborn, K., & Caprioli, R. M. (2010). Molecular imaging by mass spectrometry—Looking beyond classical histology. *Nature Reviews. Cancer*, *10*(9), 639–646. http://dx.doi.org/10.1038/nrc2917.

Schwamborn, K., Krieg, R. C., Jirak, P., Ott, G., Knuchel, R., Rosenwald, A., et al. (2010). Application of MALDI imaging for the diagnosis of classical Hodgkin lymphoma. *Journal of Cancer Research and Clinical Oncology*, *136*(11), 1651–1655.

Schwamborn, K., Krieg, R. C., Reska, M., Jakse, G., Knuechel, R., & Wellmann, A. (2007). Identifying prostate carcinoma by MALDI-Imaging. *International Journal of Molecular Medicine*, *20*(2), 155–159.

Schwartz, S. A., Reyzer, M. L., & Caprioli, R. M. (2003). Direct tissue analysis using matrix-assisted laser desorption/ionization mass spectrometry: Practical aspects of sample preparation. *Journal of Mass Spectrometry*, *38*(7), 699–708.

Scigelova, M., Hornshaw, M., Giannakopulos, A., & Makarov, A. (2011). Fourier transform mass spectrometry. *Molecular & Cellular Proteomics*, *10*(7). M111.009431.

Seeley, E. H., Oppenheimer, S. R., Mi, D., Chaurand, P., & Caprioli, R. M. (2008). Enhancement of protein sensitivity for MALDI imaging mass spectrometry after chemical treatment of tissue sections. *Journal of the American Society for Mass Spectrometry*, *19*(8), 1069–1077.

Signor, L., & Boeri Erba, E. (2013). Matrix-assisted laser desorption/ionization time of flight (MALDI-TOF) mass spectrometric analysis of intact proteins larger than 100 kDa. *Journal of Visualized Experiments*, *79*, e506351.

Sinha, T. K., Khatib-Shahidi, S., Yankeelov, T. E., Mapara, K., Ehtesham, M., Cornett, D. S., et al. (2008). Integrating spatially resolved three-dimensional MALDI IMS with in vivo magnetic resonance imaging. *Nature Methods*, *5*(1), 57–59. http://dx.doi.org/10.1038/nmeth1147.

Sjovall, P., Lausmaa, J., & Johansson, B. (2004). Mass spectrometric imaging of lipids in brain tissue. *Analytical Chemistry*, *76*(15), 4271–4278.

Sjovall, P., Lausmaa, J., Nygren, H., Carlsson, L., & Malmberg, P. (2003). Imaging of membrane lipids in single cells by imprint-imaging time-of-flight secondary ion mass spectrometry. *Analytical Chemistry*, *75*(14), 3429–3434.

Steurer, S., Borkowski, C., Odinga, S., Buchholz, M., Koop, C., Huland, H., et al. (2013). MALDI mass spectrometric imaging based identification of clinically relevant signals in prostate cancer using large-scale tissue microarrays. *International Journal of Cancer*, *133*(4), 920–928.

Steurer, S., Seddiqi, A. S., Singer, J. M., Bahar, A. S., Eichelberg, C., Rink, M., et al. (2014). MALDI imaging on tissue microarrays identifies molecular features associated with renal cell cancer phenotype. *Anticancer Research*, *34*(5), 2255–2261.

Steurer, S., Singer, J. M., Rink, M., Chun, F., Dahlem, R., Simon, R., et al. (2014). MALDI imaging-based identification of prognostically relevant signals in bladder cancer using large-scale tissue microarrays. *Urologic Oncology*, *32*(8), 1225–1233.

Stoeckli, M., Chaurand, P., Hallahan, D. E., & Caprioli, R. M. (2001). Imaging mass spectrometry: A new technology for the analysis of protein expression in mammalian tissues. *Nature Medicine*, *7*(4), 493–496.

Stoeckli, M., Staab, D., & Schweitzer, A. (2007). Compound and metabolite distribution measured by MALDI mass spectrometric imaging in whole-body tissue sections. *International Journal of Mass Spectrometry*, *260*(2–3), 195–202.

Strand, F. L. (2003). Neuropeptides: General characteristics and neuropharmaceutical potential in treating CNS disorders. *Progress in Drug Research*, *61*, 1–37.

Sugihara, Y., Végvári, Á., Welinder, C., Jönsson, G., Ingvar, C., Lundgren, L., et al. (2014). A new look at drugs targeting malignant melanoma—An application for mass spectrometry imaging. *Proteomics*, *14*(17–18), 1963–1970.

Taban, I. M., Altelaar, A. F., van der Burgt, Y. E., McDonnell, L. A., Heeren, R. M., Fuchser, J., et al. (2007). Imaging of peptides in the rat brain using MALDI-FTICR mass spectrometry. *Journal of the American Society for Mass Spectrometry*, *18*(1), 145–151.

Takáts, Z., Wiseman, J. M., Gologan, B., & Cooks, R. G. (2004). Mass spectrometry sampling under ambient conditions with desorption electrospray ionization. *Science*, *306*(5695), 471–473.

Toghi Eshghi, S., Yang, S., Wang, X., Shah, P., Li, X., & Zhang, H. (2014). Imaging of N-linked glycans from formalin-fixed paraffin-embedded tissue sections using MALDI mass spectrometry. *ACS Chemical Biology*, *9*(9), 2149–2156.

Trim, P. J., Henson, C. M., Avery, J. L., McEwen, A., Snel, M. F., Claude, E., et al. (2008). Matrix-assisted laser desorption/ionization-ion mobility separation-mass spectrometry imaging of vinblastine in whole body tissue sections. *Analytical Chemistry*, *80*(22), 8628–8634.

van Hove, A., Erika, R., Smith, D. F., & Heeren, R. M. A. (2010). A concise review of mass spectrometry imaging. *Journal of Chromatography. A*, *1217*(25), 3946–3954.

Varki, A. (1993). Biological roles of oligosaccharides: All of the theories are correct. *Glycobiology*, *3*(2), 97–130.

Végvári, Á., Shavkunov, A. S., Fehniger, T. E., Grabau, D., Niméus, E., & Marko-Varga, G. (2016). Localization of tamoxifen in human breast cancer tumors by MALDI mass spectrometry imaging. *Clinical and Translational Medicine*, *5*(1), 1–8.

Waki, M., Ide, Y., Ishizaki, I., Nagata, Y., Masaki, N., Sugiyama, E., et al. (2014). Single-cell time-of-flight secondary ion mass spectrometry reveals that human breast cancer stem cells have significantly lower content of palmitoleic acid compared to their counterpart non-stem cancer cells. *Biochimie*, *107*(Pt. A), 73–77.

Walch, A., Rauser, S., Deininger, S.-O., & Höfler, H. (2008). MALDI imaging mass spectrometry for direct tissue analysis: A new frontier for molecular histology. *Histochemistry and Cell Biology*, *130*(3), 421–434.

Willems, S. M., van Remoortere, A., van Zeijl, R., Deelder, A. M., McDonnell, L. A., & Hogendoorn, P. C. W. (2010). Imaging mass spectrometry of myxoid sarcomas identifies proteins and lipids specific to tumour type and grade, and reveals biochemical intratumour heterogeneity. *The Journal of Pathology*, *222*(4), 400–409.

Winderbaum, L., Koch, I., Mittal, P., & Hoffmann, P. (2016). Classification of MALDI-MS imaging data of tissue microarrays using canonical correlation analysis based variable selection. *Proteomics*, *16*, 1731–1735.

Wisztorski, M., Lemaire, R., Stauber, J., Menguelet, S. A., Croix, D., Mathé, O. J., et al. (2007). New developments in MALDI imaging for pathology proteomic studies. *Current Pharmaceutical Design*, *13*(32), 3317–3324.

Wuhrer, M., Deelder, A. M., & van der Burgt, Y. E. (2011). Mass spectrometric glycan rearrangements. *Mass Spectrometry Reviews*, *30*(4), 664–680.

Yamada, Y., Hidefumi, K., Shion, H., Oshikata, M., & Haramaki, Y. (2011). Distribution of chloroquine in ocular tissue of pigmented rat using matrix-assisted laser desorption/ionization imaging quadrupole time-of-flight tandem mass spectrometry. *Rapid Communications in Mass Spectrometry*, *25*(11), 1600–1608.

Yanagisawa, K., Yu, S., Xu, B. J., Massion, P. P., Larsen, P. H., White, B. C., et al. (2003). Proteomic patterns of tumour subsets in non-small-cell lung cancer. *The Lancet*, *362*(9382), 433–439.

Yang, L., Cui, X., Zhang, N., Li, M., Bai, Y., Han, X., et al. (2015). Comprehensive lipid profiling of plasma in patients with benign breast tumor and breast cancer reveals novel biomarkers. *Analytical and Bioanalytical Chemistry*, *407*(17), 5065–5077.

Yang, C., He, Z., & Yu, W. (2009). Comparison of public peak detection algorithms for MALDI mass spectrometry data analysis. *BMC Bioinformatics*, *10*(1), 4.

Appendix III

Proteomics of Endometrial Cancer Diagnosis, Treatment and Prognosis

This page intentionally left blank

REVIEW

Proteomics of endometrial cancer diagnosis, treatment, and prognosis

Parul Mittal[1,2], Manuela Klingler-Hoffmann[1,2], Georgia Arentz[1,2], Chao Zhang[1,2], Gurjeet Kaur[3], Martin K. Oehler[4] and Peter Hoffmann[1]

[1] Adelaide Proteomics Centre, School of Biological Sciences, The University of Adelaide, Adelaide, Australia
[2] Institute for Photonics and Advanced Sensing (IPAS), The University of Adelaide, Adelaide, Australia
[3] Institute for Research in Molecular Medicine, Universiti Sains Malaysia, Minden, Pulau Pinang, Malaysia
[4] Department of Gynaecological Oncology, Royal Adelaide Hospital, North Terrace, Adelaide, Australia

This review discusses the current status of proteomics technology in endometrial cancer diagnosis, treatment and prognosis. The first part of this review focuses on recently identified biomarkers for endometrial cancer, their importance in clinical use as well as the proteomic methods used in their discovery. The second part highlights some of the emerging mass spectrometry based proteomic technologies that promise to contribute to a better understanding of endometrial cancer by comparing the abundance of hundreds or thousands of proteins simultaneously.

Received: May 8, 2015
Revised: August 13, 2015
Accepted: November 2, 2015

Keywords:
Biomarker / Endometrial cancer

1 Introduction

Endometrial cancer (EC) is the most common malignant tumor of the female reproductive tract. According to the American Cancer Society, an estimated 52 630 new EC cases were diagnosed and 8590 patients died from the disease in 2014 in the United States. From 2006 to 2010, incidence rates of EC increased by 1.5% per year among women younger than 50 years and by 2.6% per year among women 50 years and older [1]. Based on the classification system introduced by Bokhman in 1983, EC is divided into two histological subtypes: endometroid adenocarcinomas (Type I) and nonendometroid (Type II) carcinomas [2]. Type I adenocarcinomas account for 90% of all EC. These cancers are usually low grade, diagnosed at an early stage and have a good prognosis. They are associated with oestrogen excess, obesity, and atypical hyperplasia. In contrast, type II nonendometroid tumors are high grade tumors with more aggressive biological behavior compared to type I disease. Clinically, type II cancers are often diagnosed at advanced stage when prognosis is poor. Type II tumors include serous and clear cell carcinomas, as well as carcinosarcomas.

Proteomics, more precisely the comparative quantitation of protein subsets, holds great promise in improving outcome of patients with EC as it provides unique tools for discovery of new biomarkers and therapeutic targets. To date, biomarkers have proven tremendous clinical value in early diagnosis [3, 4], categorizing different subtypes of malignancies [5, 6], and in monitoring patient's response to therapy [7, 8]. Regardless of the numerous proteomic studies that have contributed to the standardization of experimental protocols for digestion [9], separation [10], enrichment [11], identification [12], and quantification [13] of low abundant proteins by highly efficient mass spectrometric techniques, proteomic research is still restricted by both technology and bioinformatics tools. Many protein and peptide peaks have been reported to bear significant diagnostic [14], prognostic [15], or predictive value [16] for EC; however, the candidate biomarkers have not yet been validated for use in clinical patient care [17]. Some authors have speculated that this might be due to studies using a single proteomics approach, which is not sufficient to gain an in depth understanding of the protein function and does not eliminate false-negative and false-positive results [18]. Nevertheless, implementing "omics" integration approaches

Correspondence: Professor Peter Hoffmann, Adelaide Proteomics Centre, University of Adelaide, Gate 8 Victoria Drive, Adelaide 5005, South Australia
E-mail: peter.hoffmann@adelaide.edu.au
Fax: +61-08-0-8313-4362

Abbreviations: **ANXA2**, annexin A2; **COX**, cyclooxygenase; **DDA**, data-dependent acquisition; **DIA**, data-independent acquisition; **EC**, endometrial cancer; **FFPE**, formalin-fixed paraffin embedded; **MSI**, microsatellite instability; **ROC**, receiver operating characteristic

Colour Online: See the article online to view Figs. 1 and 4 in colour.

© 2015 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

and the use of the rapidly emerging MS-based proteomic technologies will aid in the venture to elucidate protein markers and their function, ultimately providing more reliable, sensitive, and specific biomarkers for EC.

In the first part of this review article, we will summarize the various MS-based proteomic approaches that have been used in EC studies. In the second part, we will introduce emerging proteomic technologies that may not only be used as discovery tools but also allow the implementation of MS in the diagnosis and prognosis of EC.

2 Clinical considerations of endometrial cancer

The clinical management of EC remains a challenge with patients presenting with a full spectrum of disease ranging from those with excellent prognosis and high curability to aggressive disease with poor outcome [19]. EC is staged according to the International Federation of Gynaecology and Obstetrics (FIGO) system (Table 1) [20]. The classification of EC into low and high risk disease is dependent on a number of parameters, only two of which are assessable pre-operatively [http://wiki.cancer.org.au/australiawiki/index.php?oldid=86 682] after a biopsy has been obtained, namely the histological type and grade. However, Jacques et al. have shown a large percentage of cancer will be classified as a higher grade tumor after definitive surgery [21]. Therefore, the exact type, grade, and stage of EC can only be determined by surgery and subsequent histopathological assessment.

Table 1. Classification of carcinoma of the endometrium, International Federation of Gynaecology and Obstetrics (FIGO 2010), table modified from Creasman et al. [100]

FIGO stage	Tumor localization
Stage I	Tumor confined to the corpus uteri
IA	No or less than half myometrial invasion
IB	Invasion equal or more than half of the myometrium
Stage II	Tumor invades cervical stroma, but does not extend beyond the uterus
Stage III	Local and/or regional spread of the tumor
IIIA	Tumor invades the serosa of the corpus uteri and/or adnexa
IIIB	Vaginal and/or parametrial involvement
IIIC	Metastasis to pelvic and/or para-aortic lymph nodes
IIIC1	Positive pelvic nodes
IIIC2	Positive para-aortic lymph nodes with or without positive pelvic lymph nodes
Stage IV	Tumor invades bladder and/or bowel mucosa, and/or distant metastasis
IVA	Tumor invades bladder and/or bowel mucosa
IVB	Distant metastasis, including intraabdominal metastasis and/or inguinal lymph nodes

2.1 Diagnosis

EC is frequently diagnosed at an early stage, as it regularly presents with symptoms such as postmenopausal bleeding that usually develops early in the disease process. Nevertheless, the discovery of serum biomarkers for early detection of EC has become a high priority [22]. A number of serum biomarkers have been identified so far, with the most commonly used serum biomarker in gynecological oncology being Carbohydrate Antigen 125 (CA 125). In EC elevated serum CA 125 level have been detected in 11–43% of the cases and is shown to correlate with advanced stage and with the presence of extrauterine disease [23–25]. However, due to lack of specificity and sensitivity, CA 125 has limited significance in the diagnosis of EC. Human Epididymis protein 4 (HE4) has recently emerged as a promising biomarker for EC [26]. Brennan et al. highlighted the utility of serum HE4 using ELISA for preoperative risk stratification to identify high-risk patients within low-grade endometrioid EC patients who might benefit from lymphadenectomy [27]. For stage I EC, HE4 showed a 17% improvement in sensitivity when compared to CA 125 [28].

Hareyama et al. investigated the immunohistochemical CA72-4 expression in EC. The authors reported an elevated level of serum CA72-4 in 22–32% of the cases, that was associated with depth of myometrial invasion, adnexal metastasis, lymphovascular space invasion, and pelvic and para-aortic lymph node metastasis [29]. Konno et al. demonstrated the level of serum soluble Fas (sFas) is significantly higher in EC patients with advanced cancer when compared to localized cancer ($p < 0.0001$) [30]. A study by Sawada et al. showed raised serum levels of immunosuppressive acidic protein in 55–76% of EC cases where the level was found to increase with the stage of the disease [31]. Hakala et al. reported the serum level of macrophage colony-stimulating factor 1 (mcsf 1), correlated significantly with tumor grade and poor prognosis in 25–73% of EC cases [32,33]. A recent study by Kang-Wai Mu et al. reported differential levels of zinc alpha-2 glycoprotein, alpha 1-acid glycoprotein, and CD59 in the urine of EC patients when compared to urine from healthy controls using 2DIGE and *O*-Glycan binging lectin [34].

2.2 Treatment

EC is usually treated by surgery consisting of a hysterectomy, bilateral salpingo-oophorectomy, and pelvic lymph node dissection and has a very good prognosis when tumour mass is confined to the uterus [28]. In contrast, EC with metastasis to the lymph nodes has a high mortality rate. Lymph node metastasis is therefore a crucial factor in the prognosis and choice of treatment of EC patients. Clinicians presently face the challenge that conventional surgical-pathological variables (e.g. tumor size, depth of invasion, and grade of disease) and radiological imaging are unreliable in determining if a gynecological cancer has spread. Consequently, although

© 2015 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

only 5% of patients suffer from metastasic disease, the majority undergo radical treatment including removal of the lymph nodes. Lymph node dissection, however, is associated with significant complications such as lower extremity lymphedema [35]. Predictive tissue markers for lymph node metastasis are therefore warranted to determine the optimal treatment strategy and to avoid morbidity in gynecological cancer.

2.3 Prognosis

The prognosis of EC depends on various factors including the histological subtype, grade, and stage of the disease (which is determined by depths of tumor invasion, lymph node metastasis, and spread of the disease to other organs) [36]. The depth of myometrial invasion and histological grade strongly correlate with the prevalence of lymph node metastasis and patient survival [37].

Various protein biomarkers have been described in EC that could be of benefit for the prediction of disease outcome. According to Cocco et al. serum amyloid A may represent a novel biomarker for EC to monitor disease recurrence and response to therapy [38]. Lo et al. demonstrated that increased levels of CA-125, CA 15-3, and CA 19-9 were significantly associated with poor prognostic clinical parameters [39]. Lambropoulou et al. evaluated the prognostic significance of survivin, c-erbB2, and cyclooxygenase (COX-2) levels in EC and stated that in a 10-year follow-up, patients with tumors expressing more of these three antigens had significantly lower survival rate that those with smaller expression score [40]. Yilmaz et al. have also shown higher nuclear expression of survivin in type I when compared to type II EC ($p = 0.040$); but no difference for cytoplasmic survivin and matrix metalloproteinase-2 expressions between type I and type II EC [41]. Zeimet et al., identified L1 cell adhesion molecule (L1CAM) as the best variable for predicting recurrence (sensitivity = 0.74; specificity = 0.91) and death (sensitivity = 0.77; specificity = 0.89) [42]. Although, the above-mentioned biomarkers are able to identify high risk patients with low grade EC, for clinical application additional biomarkers are required. New prognostic EC biomarkers will help distinguish patients who are at a low-risk of developing metastasis as compared to those who are at a high risk. Ideally, novel biomarkers will also be able to distinguish patients who have already developed metastasis. This would consequently allow patients to selectively undergo radical surgery while avoiding those who will not benefit from it and subsequently decreasing the risk of postsurgical morbidity.

3 Molecular genetics of endometrial cancer

Aside from the morphologic and clinical features, EC can be further distinguished into type I and type II EC on the basis of genetic alterations [43]. Type I and type II ECs are associated with mutations of independent gene sets [44]. Type I endometrial adenocarcinomas are characterized by mutations in the PTEN, K-ras, and β catenin, as well as DNA mismatch repair genes [35] while type II nonendometrioid EC frequently show aneuploidy and p53 mutations [35].

PTEN has been reported to be the most frequent genetic alteration of type I EC. PTEN, a tumor suppressor gene has been shown to be altered in up to 83% of endometrial adenocarcinomas and 55% in precancerous lesions [45]. Microsatellite instability (MSI) has been demonstrated in 20% of type I EC [46]. According to Bilbao et al. both PTEN mutations and MSI represent early events in endometrial carcinogenesis [47]. Other genetic alterations that occur in type I EC include mutations in K-ras [48] and β catenin [49]. PTEN, MSIS, and K-Ras mutations often coexist with each other, whereas mutations in β-catenin are usually observed alone [50].

The most common genetic alteration in type II EC is within p53, the tumor suppressor gene, with an occurrence of 93% [48, 51]. Other frequent genetic alteration in type II ECs are inactivation of the tumor suppressor gene p16 and overexpression of the oncogene, HER-2/neu [52].

4 Metastasis and protein biomarkers

Metastatic cancer cells proliferate, lose contact with neighboring cells, migrate through interstitial matrix, invade blood and lymph vessels, and colonize lymph nodes or distant organs [53]. Most cancer cells fail to undergo metastasis due to deficiency in one of the required steps like invasion, detachment or survival [54]. Discovery of EC biomarkers would significantly aid gynecological oncologists who currently face the challenge that radiological imaging and conventional surgical-pathological variables such as tumor size, depth of invasion, and grade of disease are unreliable in determining if an EC has metastasized.

The working model of metastasis implies that primary tumor cells acquire genetic alterations over time, which enables these cells to metastasize and form new solid tumors at distant sites [55]. It has been proposed that the gene expression program of metastasis is actually present in the bulk of primary tumor cells [56]. Ramaswamy et al. showed that some primary tumors are preconfigured to metastasize, and their susceptibility is detectable at the time of diagnosis. They defined a 17-gene signature pattern associated with metastasis, in which eight genes are upregulated, while nine genes are downregulated. They also concluded that metastasis is not dependent on only a single gene but on the complete 17 gene sequence.

Yi et al., have shown the expression of COX-2 to play an important role in metastasis of EC [57]. They identified the COX-2 inhibitor NS-398 to inhibit proliferation, viability, and invasion of the EC cell line RL95–2 [57]. Maxwell et al. confirmed annexin A2 (ANXA2) and peroxiredoxin were both being overexpressed in stage 1 EC when compared to normal endometrium [58].

© 2015 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

5 MS-based proteomic approaches in endometrial cancer

The use of proteomic technologies are now enabling the identification and relative quantification of multiple proteins simultaneously from a single experiment to identify disease related and specific biomarkers. Proteomics in general deals with the large-scale determination of gene and cellular function directly at the protein level [59]. The proteome has also been defined as the protein complement expressed by a genome [60] but in reality, due to PTM and alternative splicing, it is estimated that about 22 000 protein coding genes [61] code for more than 500 000 proteins [62] in the human proteome.

In serum it is estimated that cancer derived proteins are 10 million times less abundant than common high abundant proteins secreted by normal cells. Therefore, it is challenging to correctly identify and quantify tumor-derived proteins from the whole serum proteome [63]. Alternatively, it is obvious that cancer derived proteins will be present in higher concentration in native tissue, organs, and their proximal fluids from where the tumor has originated as compared to the distant sites in which the tumor derived proteins may be secreted or leaked. Therefore, targeting those regional sites will dramatically increase the possibility of isolating and identifying tumor-specific biomarkers [64].

Proteomic analysis principally relies on MS for protein identification. The uses of MS-based techniques have made it possible to build comprehensive profiles of near complete proteomes, comparing the expressions of individual proteins that may serve as biomarkers. Generally, MS-based proteomic techniques can be classified as "gel based" and "gel free" (Fig. 1). In gel-based approaches such as 1D-PAGE, 2D-PAGE, and 2D-DIGE, proteins are separated on polyacrylamide gels via electrophoresis in one or two dimensions, creating unique proteomic patterns based on the mass and/or charge/pH of the proteins.

Gel-free MS approaches can be either label based or label free. Label-based proteomic techniques are common for MS approaches where proteins are tagged either with isotopes such as ICAT, iTRAQ or labeled chemically such as SILAC. Labeling facilitates the quantification of proteins during MS analysis and is especially useful when acquiring data in the standard form of data-dependent acquisition (DDA), as labeling allows for the mixing and analysis of multiple samples at once. This alleviates the stochastic nature of DDA that can otherwise lead to sampling biases and poor reproducibility. Label-free techniques in the form of data-independent acquisition (DIA) help resolve the issue of sampling biases and eliminates any problems encountered due to artifacts from labeling that can interfere with the identity of the protein, leading to false conclusions [65,66]. DIA is addressed in more depth later in this article.

MS-based proteomics techniques have developed substantially over the past several years. Previously, protein analysis was reliant on 1D and 2D gel electrophoresis followed by sequencing; finally resulting in the identification of tens to hundreds of proteins. This limitation has been overcome with novel mass spectrometric approaches, which can identify many more proteins at once. Identification and quantification of proteins/peptides by MS can be done either by "bottom-up" or "top-down" approaches. The term "bottom-up" refers to the reconstruction of the protein sequence after proteolytic digestion, and thereby, identification and quantification of the peptide fragments using MS and appropriate databases [67]. The second approach "top-down," refers to the identification of the whole proteins directly without enzymatic digestion [68].

To date, the most widely used proteomic technologies applied to the identification of the EC biomarkers includes

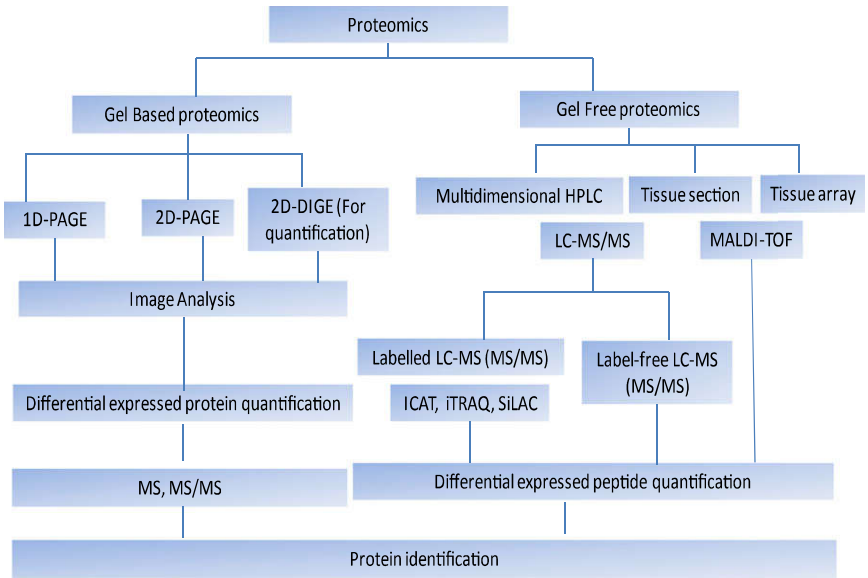


Figure 1. Proteomic approaches used in the EC biomarker discovery: Gel based and Gel free. In Gel-based proteomic approaches, proteins are separated on the basis of their charge/pH or mass by isoelectric focussing and electrophoresis. Gel-free techniques include MS-based technologies.

© 2015 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

LC-MS. This approach is generally used with "bottom-up" methods. LC-MS has been used to identify and quantify differential change in protein abundance in healthy versus disease state. Identification of proteins by LC-MS/MS can be broken down into a few essential components. First, complex protein samples are digested with an enzyme (protease) into peptides, which are separated by a high resolution chromatographic technique coupled directly to a MS. Following chromatographic separation samples are sprayed into the MS while being ionized, a process called ESI. Here, the intact peptide mass is measured, followed by collision with an inert gas to induce fragmentation, upon which the mass of the fragment ions are measured, which is commonly referred as MS/MS or tandem MS. The acquired data is then processed to gain peptide sequences and protein identifications, and examined using mathematical algorithms, functions to detect and identify differences within the sample and/or between the samples. In traditional label-free approaches, peptide separation is compared across chromatographic profiles, and quantitation is determined either by spectral counting or by peak intensity during MS [69].

A recent study by Alconada et al. has identified and validated ANXA2 as a reliable biomarker of recurrent disease in EC by 2D DIGE followed by traditional LC-MS/MS. They demonstrated in vitro and with an in vivo mouse model that the increased expression of ANXA2 is associated with an improved ability of the cells to metastasize [70]. Monge et al. analyzed the role of the transcription factor ERM/ETV5 in myometrial invasion by 2D-DIGE to evaluate the differential expression of proteins in EC cell lines overexpressing ERM/ETV5 [71]. The authors have demonstrated that ERM/ETV5 acts by involving matrix metalloproteinase-2 to provide the migratory and invasive capabilities associated with the switch to myometrial infiltration in Hec-1A EC cell line [71]. They further characterized a role for ETV5 in a modulated response to oxidative stress associated with the promotion of invasion in EC [71]. Ihata et al. analyzed the amino acid concentrations in plasma samples by LC-MS and reported that the amino acid profile index is a potential new modality that could eventually play a significant role in the preoperative evaluation of EC and its screening [72].

The MALDI technique has also been widely used for the discovery of gynecological cancer biomarkers but in a more limited context for EC. Qiu et al. have reported that MALDI-TOF-MS is a high-throughput, sensitive, highly predictive, and rapid method for the early detection, diagnosis, and prognosis of EC and will be widely used in future clinical screening work [73]. In this technique, molecules including proteins, peptides, lipids and metabolites are first co-crystallized with large excess of suitable matrix (usually a weak organic acid) and then spotted on to a MALDI target plate. MALDI-TOF-MS is then performed, a process in which the laser transfers the energy to the aromatic matrix molecules and the matrix transfer a proton to the analyte, resulting in soft energy transfer without fragmentation. According to the m/z ratios, the ionized analytes are then separated in the TOF mass analyzer and the detector transforms the incoming ions into an electric current that is proportional to their abundance [74]. In the case of protein analysis, the generated protein/peptide mass list can then be compared to protein databases that allow for the matching of peptide masses and identification of proteins, a process known as PMF. In addition using tandem MS, the parent ion can be further fragmented into daughter ions by the collision with an inert gas and peptide sequence is inferred by identifying the daughter ions. Casado-Vela et al. used an integrated approach combining three different but complementary strategies (in-solution digest followed by reverse phase HPLC-MS/MS, protein separation by SDS-PAGE followed by HPLC-MS/MS analysis, 2D-PAGE followed by MALDI-TOF/TOF analysis) to describe comprehensive proteomic analysis of EC fluid aspirate that led to the successful identification of 803 different proteins in the International Protein Index human database (v3.48) [75]. Therefore, EC biomarkers can be used for screening, estimating risk of disease, distinguishing benign from malignant, determining prognosis, and monitoring status of the disease, either to detect recurrence or response to therapy. Importantly, some biomarkers are only used in a specific setting, whereas others can serve more than one purpose [76].

6 Emerging techniques for the proteomic analysis of endometrial cancer

Traditional proteomic approaches as described above, continue to be the most widely applied technologies for protein identification; nonetheless, a major drawback of these proteomics techniques is that not every peptide from a complex biological sample can be detected. As a result, the proteomes will be under sampled. New techniques are emerging that offer innovative solutions to the analysis of low abundant proteins from complex biological samples [77]. Those proteomics technologies have the potential for direct clinical applications including the identification of novel biomarkers for the monitoring of therapy response, and disease progression, delivering on the great promise of personalized cancer medicine [78].

6.1 Data-independent acquisition

Protein identification by LC-MS is a straightforward technique, as the detection of only two unique peptides from a protein is considered sufficient. However, relative quantification by LC-MS is a more complex process and the failure to identify or detect a peptide does not necessarily confirm absence, as the peptides may simply be below the threshold of detection [79]. Traditional LC-MS or shotgun approaches are based on data-dependent acquisition (DDA), where peptides are selected for isolation and fragmentation based on their relative abundance within the sample. This makes the technique stochastic and can result in under sampling of the injected

© 2015 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

analyte [80]. Abundance measurements can be inaccurate as the quantification of a peptide is highly influenced by coeluting peptides and their signal to noise ratios, meaning peptides of a higher abundance are quantified more accurately and more often than peptides of low abundance. We do however; know that many cancer-relevant proteins including signaling and regulatory proteins are typically expressed in low concentrations. As a result, the DDA approach tends to miss out on acquiring the most-valuable information [81]. Data-independent acquisition (DIA) aims to solve the problem of DDA by measuring all precursor ions within a defined mass range of ~25 Da, fragmenting and measuring all of the ions within that mass range [82]. Typically a 25 Da mass window is stepped across the entire mass range covered in traditional LC-MS/MS (~150–2200 Da) in cycle times compatible with the coupled chromatographic separation [82]. Vowinckel et al. shown the number of precisely quantifiable peptides can be increased up to five-fold when using DIA compared to DDA, with a CV less than 5% [83].

6.2 MALDI imaging

Another innovative technology that can be used for both biomarker discovery and validation is MALDI imaging (MALDI-MSI), which facilitates the direct analysis of protein distribution and abundance in thin tissue sections [84]. This technique has the potential to detect and characterize tumor margins and their environment in a spatial context [56]. MALDI-MSI is a valuable method for the identification of biomarkers and can complement histology, immunohistochemistry [85], and molecular pathology in various fields of histopathological diagnostics, especially with regard to identification and grading of tumors [86]. Whereas, MALDI-MSI is optimally suited to detect a large number of molecular features in a given tissue sample, the tissue microarray method enables expansion of this analysis to hundreds of tissue samples in a single experiment [87,88]. In MALDI-MSI, proteins or in situ digested peptides are directly ionized from the surface of the tissue sample and a characteristic histological pattern is derived for hundreds or thousands of individual proteins/peptides simultaneously. The identification of proteins may be further investigate by PMF, a process of correlating the acquired mass spectra with the theoretical peptides generated from a given protein or proteins to find the closest match or by in situ MS/MS [89]. During in situ MS/MS, a desired peptide is isolated based on its m/z ratio and fragmented (i.e. breaking peptide bonds within the peptides), producing a series of fragment ions that are detected as a MS/MS spectrum. The fragmentation pattern is compared to the theoretical fragmentation pattern for every peptide in the proteome to find the closest match. In this way the sequence of the peptide ion is inferred from its fragmentation pattern [90]. Realistically in the case of MALDI-MSI, this form of MS/MS is limited to highly abundant ionizable peptides whereas the identifications of lower abundance peptides can be made by matching back to fragmentation spectra acquired from LC-MS/MS using a tissue extract.

MALDI-MSI has been successfully used to classify HER2 receptor expression in breast cancer [91]. Meding et al. have shown the power of MALDI-MSI in classification of metastasis for different tumor entities [92]. In addition, their results indicated that even closely related entities such as primary tumor of the colon and its liver metastasis could be classified efficiently [92]. Quaas et al. detected several molecular signals associated with phenotypic features of oesophageal cancer cells, highlighting the potential of MALDI-MSI to identify new molecular markers with relevance to oesophageal cancer [87]. Another strength of MALDI-MSI is that it can be used in the discovery phase for the identification of tumor-derived proteins and then the verification of the presence of those particular proteins can be carried out in body fluid (blood, urine, or saliva), aiming to develop noninvasive diagnostic tests. Addona et al. developed a pipeline integrating proteomic technologies from the discovery to the verification stages of plasma biomarker identification and applied it to identify early biomarkers of cardiac injury [93]. Thus, combining the histopathology reports with MALDI-MSI data will help surgeons to make informed medical decisions.

In the similar approach, we performed MALDI imaging on EC formalin-fixed paraffin embedded (FFPE) tissue samples collected from the Royal Adelaide Hospital (RAH), South Australia. Ethics approval was granted by RAH Human ethics committee.

In this study, 6 μm FFPE tissue sections were sliced and mounted onto indium tin oxide-coated conductive glass slides (Bruker Daltonics, Bremen, Germany) and samples were prepared as described by Gustafsson et al. [12]. Briefly upon Citric acid antigen retrieval (CAAR), in situ tryptic digestion was performed using an Image prep station (Bruker Daltonics, Bremen, Germany) at 37°C for 2 h. CHCA matrix solution was prepared at 7 mg/mL in 50% ACN/0.2% TFA and was overlayed onto the tissue section using an Image Prep station. Following matrix deposition, analysis was performed on an UltrafleXtreme MALDI-TOF/TOF (Bruker Daltonics, Bremen, Germany) in positive reflectron mode with FlexControl V3.0.1 and FlexImaging V4.0.1 (Bruker Daltonics) software package. The data was acquired at 100 μm lateral resolution with a laser frequency of 2000 Hz over m/z 800–4500. Following MS acquisition, an overlapped MALDI spectral profile of tumor and healthy regions was generated by FlexAnalysis software (V4.0.1, Bruker Daltonics, Bremen, Germany) and as expected, a number of m/z values have been identified that are differentially expressed between healthy and tumor regions (Fig. 2).

In order to determine data-dependent visualization of tissue morphological regions, the raw data was loaded into SCiLS lab software (V2014b, Bruker Daltonics, Bremen, Germany). Data were processed by baseline removal and TIC normalization. Healthy and tumor regions were then grouped based on annotations done by a pathologist and discriminatory m/z values were calculated by ROC (receiver operating

© 2015 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

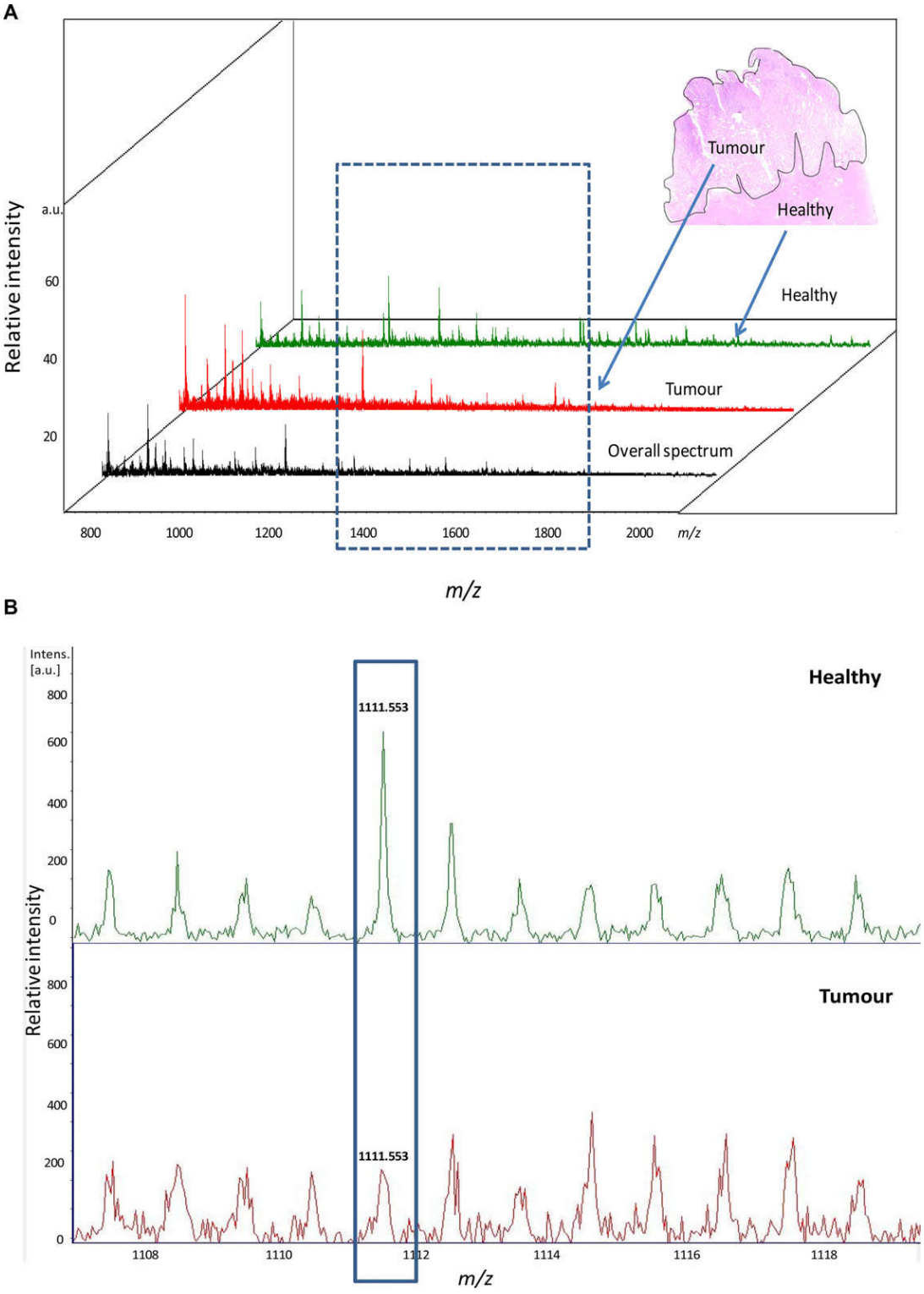


Figure 2. In this analysis in situ tryptic digestion was performed on 6 μm thick FFPE tissue. The matrix used was α-cyano-4-hydroxycinnamic acid and data were collected in positive reflectron mode using an UltrafleXtreme MALDI-TOF/TOF (Bruker Daltonics, Bremen, Germany). An overlapped MALDI spectral profile of tumor and healthy regions was generated by FlexAnalysis software (V4.0.1, Bruker Daltonics, Bremen, Germany). (A) The overlapped spectral profile show a multitude of differentially expressed *m/z* species between healthy and tumor regions simultaneously that allows a straightforward correlation of the expression pattern within the tissue morphology. (B) Comparative representative spectra of *m/z* 1111.553 Da ± 0.250 Da between healthy and tumor regions.

© 2015 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

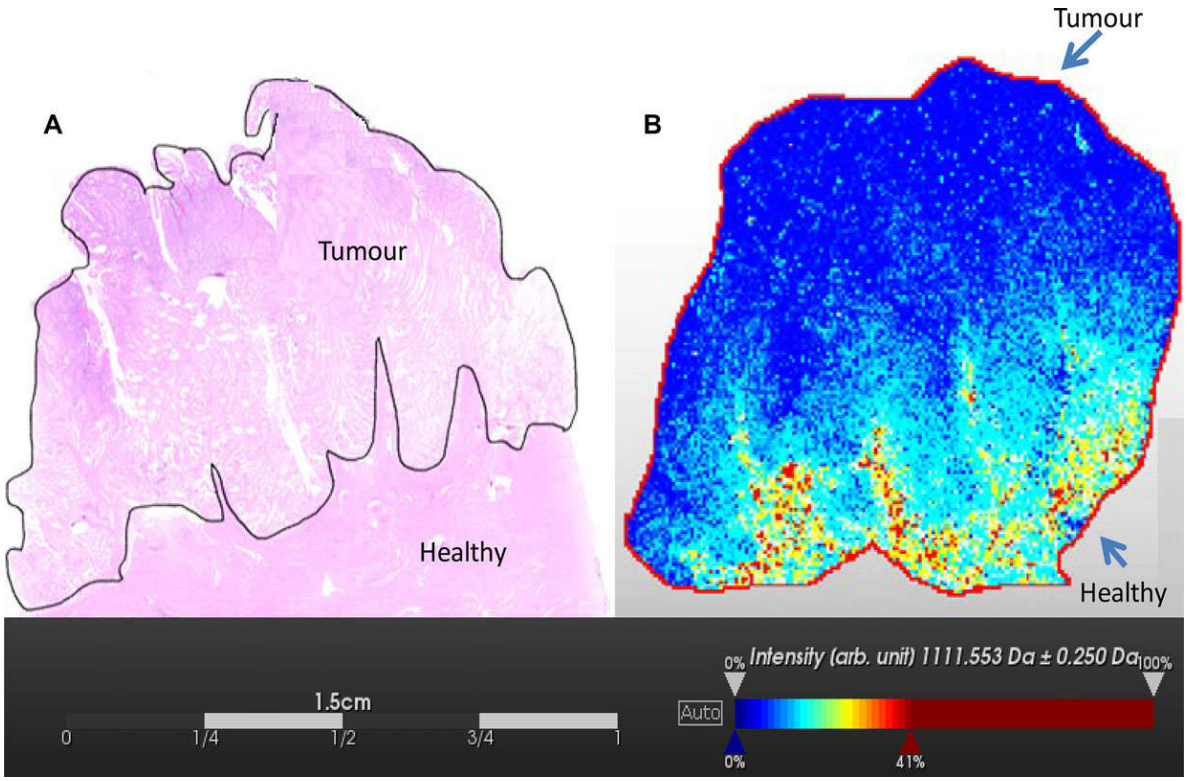


Figure 3. MALDI-MSI of an EC FFPE sample. The expression pattern of *m/z* 1111.553 Da ± 0.250 Da is visualized between healthy and tumor regions (B), when comparison is made with corresponding H&E stained annotated image (A). Visualization of difference in the intensities was performed in SCiLS lab software (V2014b, Bruker Daltonics, Bremen, Germany) with edge preserving image denoising and automatic hotspot removal applied. This *m/z* species is seen to be specifically downregulated in tumor region as compared to healthy region.

characteristic) curve. ROC curve compares sensitivity verses specificity across a range of *m/z* values by plotting a curve called an AUC. The AUC assumes values between 0 and 1 and expresses the discrimination power of the *m/z* signal. A perfect discrimination would yield an AUC value equal to 1 (abundant in group 1) or 0 (abundant in group 2) [94]. The AUC closer to 0.5 indicates that there is no discrimination between two sample groups. An example of a discriminatory *m/z* value (*m/z* 1111.553 Da ± 0.250 Da) is shown in Fig. 3. Here the spatial intensity of *m/z* 1111.553 Da ± 0.250 Da is clearly higher in the healthy tissue regions as compared to the tumor tissue region, shown also in the ROC curve (Fig. 4). The results from this study will provide crucial new tools to assist in the diagnosis and prognosis of EC, with the ultimate aim to prevent overtreatment of patients whose primary tumors do not have metastatic potential.

6.3 Targeted approach

Sensitive detection of low abundant proteins in complex samples has been typically achieved by ELISA that require antibodies specific to the protein(s) of interest. However, this poses a problem as development of antibodies is associated with high cost, long development lead times, and high failure rates [95]. To overcome this problem, for the past few decades intensive research has been carried out aiming to develop more targeted strategies that have been designed specifically to analyze preselected peptides/proteins of interest within complex samples.

Unlike untargeted approaches, which aim to detect all proteins within a complex sample in an unbiased manner, multiple reaction monitoring (MRM) is a targeted approach, in which the protein(s) of interest are preselected for quantification and analyzed using highly sensitive triple quadrupole MS [96]. MRM-MS provides high sensitivity and accuracy needed in the discovery phase and offers high reproducibility necessary for clinical validation [97]. Due to the high specificity, sensitivity, and throughput of MRM, it can be used for the validation of a single protein of interest or a subset of proteins [98]. To ensure good quantification by MRM, three of the most intense ion transitions (precursor ion fragmentation to product ions) per protein are monitored. DeSouza et al. demonstrated that the MRM of iTRAQ labeled peptides enables absolute quantification of pyruvate kinase in cancerous and normal endometrial tissues [99]. Elizabeth et al. pioneered a new technique that couples the ability to perform MRM on a triple quadrupole mass spectrometer with a MALDI source. The authors have shown that this approach has the feasibility for the precise and accurate quantitation

© 2015 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

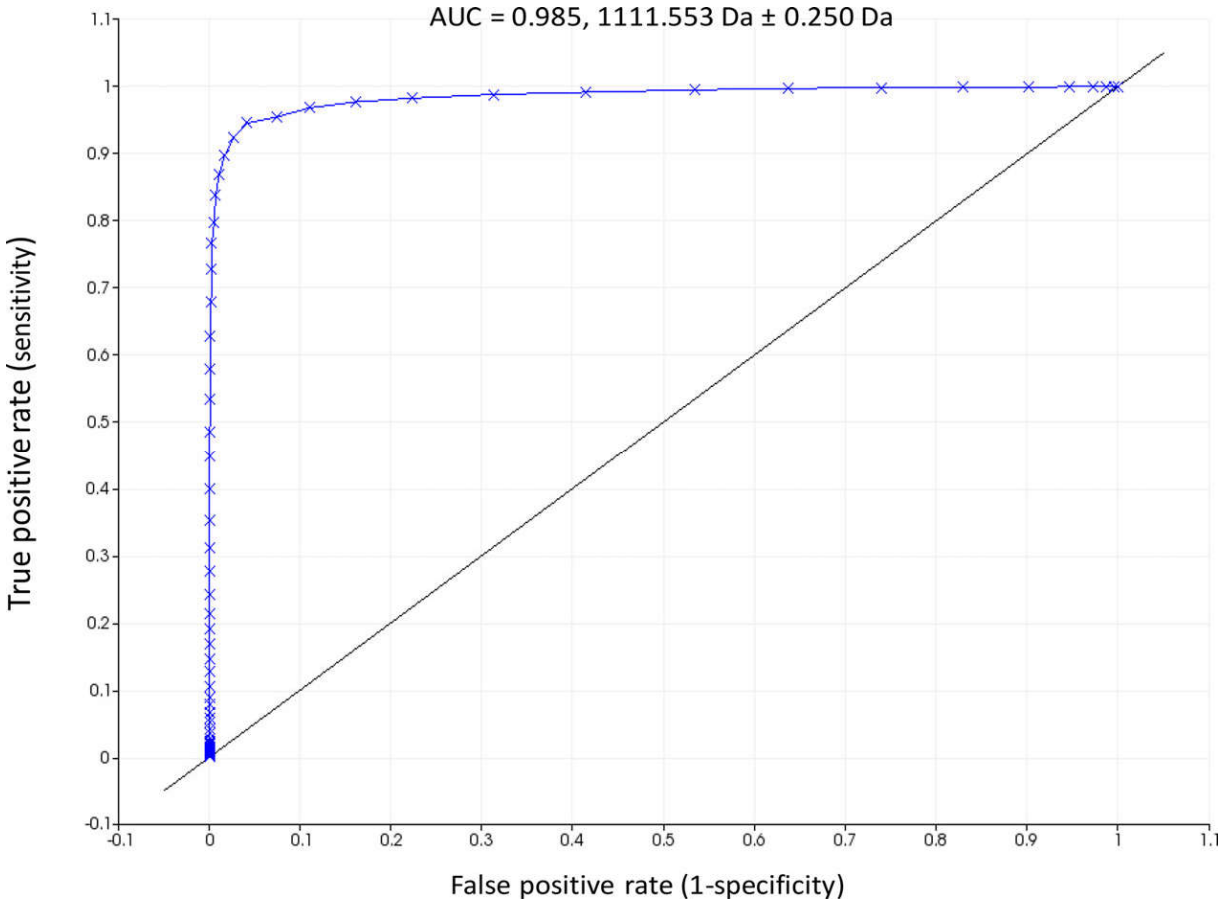


Figure 4. The ROC curve of *m/z* 1111.553 Da ± 0.250 Da (AUC = 0.985) for healthy versus tumor region.

of tissue protein concentrations over two orders of magnitude, while maintaining the spatial localization information for the proteins [100]. The greatest advantage of the MALDI MRM-based imaging technique is improved sensitivity and selectivity of the analysis, enabling measurement of the accurate protein concentration with the addition of stable isotope labeled internal standards. This technique combines the spatial information gained by traditional MALDI MSI with the accurate quantification achievable by MRM and may developed as an optimal strategy for the analysis of abundant peptide biomarkers.

7 Conclusion

Above-mentioned studies have discovered a number of candidate EC biomarkers, for example CA-125, CA 19-9, CA 15-3, HE4 for the diagnosis; and L1CAM, COX-2, Survivin, c-erb B2 for the prognosis of EC. Unfortunately due to their lack of sensitivity and specificity, none of those protein biomarkers are currently used in clinical practice. Recently, the major focus of biomarker research has shifted from single biomarker discovery to the multiparametric analysis of proteins, as the proteomic pattern analysis can obtain exceptional results. As part of this review, we have introduced two emerging techniques with exceptional sensitivity and dynamic range. These techniques are refining our capability to identify and quantify relative changes in protein expression and will ultimately result in the discovery of new EC biomarkers panels.

Moreover, the integration of the data obtained from various "omics" approaches such as genomics, transcriptomics, and proteomics will help to reduce the false positives and false negatives obtained from single "omic" approaches and provide us with a number of potential protein targets. Taken together, this is an exciting time for EC biomarker research and we are hopeful that novel EC biomarker panels will perform well and will be used in the clinic in the future.

The authors have declared no conflict of interest.

8 References

[1] American Cancer Society, *Cancer Facts and Figures 2014*. American Cancer Society, Atlanta 2014.

[2] Bokhman, J. V., Two pathogenetic types of endometrial carcinoma. *Gynecol. Oncol.* 1983, *15*, 10–17.

© 2015 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

[3] Toiyama, Y., Tanaka, K., Kitajima, T., Shimura, T. et al., Elevated serum angiopoietin-like protein 2 correlates with the metastatic properties of colorectal cancer: a serum biomarker for early diagnosis and recurrence. *Clin. Cancer Res.* 2014, *20*, 6175–6186.

[4] Bie, Y., Zhang, Z., Diagnostic value of serum HE4 in endometrial cancer: a meta-analysis. *World J. Surg. Oncol.* 2014, *12*, 169.

[5] Qendro, V., Lundgren, D. H., Rezaul, K., Mahony, F. et al., Large-scale proteomic characterization of melanoma expressed proteins reveals nestin and vimentin as biomarkers that can potentially distinguish melanoma subtypes. *J. Proteome Res.* 2014, *13*, 5031–5040.

[6] Le Faouder, J., Laouirem, S., Alexandrov, T., Ben-Harzallah, S. et al., Tumoral heterogeneity of hepatic cholangiocarcinomas revealed by MALDI imaging mass spectrometry. *Proteomics* 2014, *14*, 965–972.

[7] Kralj, E., Žakelj, S., Trontelj, J., Pajič, T. et al., Monitoring of imatinib targeted delivery in human leukocytes. *Eur. J. Pharm. Sci.* 2013, *50*, 123–129.

[8] Ubhi, B. K., Cheng, K. K., Dong, J., Janowitz, T. et al., Targeted metabolomics identifies perturbations in amino acid metabolism that sub-classify patients with COPD. *Mol. Biosyst.* 2012, *8*, 3125–3133.

[9] Meding, S., Martin, K., Gustafsson, O. J., Eddes, J. S. et al., Tryptic peptide reference data sets for MALDI imaging mass spectrometry on formalin-fixed ovarian cancer tissues. *J. Proteome Res.* 2013, *12*, 308–315.

[10] Weiland, F., Zammit, C. M., Reith, F., Hoffmann, P., High resolution two-dimensional electrophoresis of native proteins. *Electrophoresis* 2014, *35*, 1893–1902.

[11] Danihlik, J., Sebela, M., Petrivalsky, M., Lenobel, R., A sensitive quantification of the peptide apidaecin 1 isoforms in single bee tissues using a weak cation exchange preseparation and nanocapillary liquid chromatography coupled with mass spectrometry. *J. Chromatogr. A* 2014, *1374*, 134–144.

[12] 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 Commun. Mass Spectrom.* 2013, *27*, 655–670.

[13] Li, J., Zhou, L., Wang, H., Yan, H. et al., A new sample preparation method for the absolute quantitation of a target proteome using O labeling combined with multiple reaction monitoring mass spectrometry. *Analyst* 2015, *140*, 1281–1290.

[14] Ueda, Y., Enomoto, T., Kimura, T., Miyatake, T. et al., Serum biomarkers for early detection of gynecologic cancers. *Cancers* 2010, *2*, 1312–1327.

[15] Engelsen, I. B., Akslen, L. A., Salvesen, H. B., Biologic markers in endometrial cancer treatment. *APMIS* 2009, *117*, 693–707.

[16] Dong, P., Kaneuchi, M., Konno, Y., Watari, H. et al., Emerging therapeutic biomarkers in endometrial cancer. *BioMed. Res. Intern.* 2013, *2013*, 11.

[17] Galvao, E. R., Martins, L. M., Ibiapina, J. O., Andrade, H. M., Monte, S. J., Breast cancer proteomics: a review for clinicians. *J. Cancer Res. Clin. Oncol.* 2011, *137*, 915–925.

[18] Ge, H., Walhout, A. J., Vidal, M., Integrating 'omic' information: a bridge between genomics and systems biology. *Trends Gen.* 2003, *19*, 551–560.

[19] Zhang, Y., Wang, J., Controversies in the management of endometrial carcinoma. *Obst. Gynecol. Int.* 2010, *2010*, 862908.

[20] Haltia, U. M., Butzow, R., Leminen, A., Loukovaara, M., FIGO 1988 versus 2009 staging for endometrial carcinoma: a comparative study on prediction of survival and stage distribution according to histologic subtype. *J. Gynecol. Oncol.* 2014, *25*, 30–35.

[21] Jacques, S. M., Qureshi, F., Munkarah, A., Lawrence, W. D., Interinstitutional surgical pathology review in gynecologic oncology: I. Cancer in endometrial curettings and biopsies. *Int. J. Gynecolog. Pathol.* 1998, *17*, 36–41.

[22] Ueda, Y., Enomoto, T., Kimura, T., Miyatake, T. et al., Serum biomarkers for early detection of gynecologic cancers. *Cancers* 2010, *2*, 1312–1327.

[23] Powell, J. L., Hill, K. A., Shiro, B. C., Diehl, S. J., Gajewski, W. H., Preoperative serum CA-125 levels in treating endometrial cancer. *J. Reprod. Med.* 2005, *50*, 585–590.

[24] Jhang, H., Chuang, L., Visintainer, P., Ramaswamy, G., CA 125 levels in the preoperative assessment of advanced-stage uterine cancer. *Am. J. Obst. Gynecol.* 2003, *188*, 1195–1197.

[25] Scambia, G., Gadducci, A., Panici, P. B., Foti, E. et al., Combined use of CA 125 and CA 15-3 in patients with endometrial carcinoma. *Gynecol. Oncol.* 1994, *54*, 292–297.

[26] Brennan, D. J., Hackethal, A., Metcalf, A. M., Coward, J. et al., Serum HE4 as a prognostic marker in endometrial cancer—a population based study. *Gynecol. Oncol.* 2014, *132*, 159–165.

[27] Brennan, D. J., Hackethal, A., Metcalf, A. M., Coward, J. et al., Serum HE4 as a prognostic marker in endometrial cancer—a population based study. *Gynecol. Oncol.* 2014, *132*, 159–165.

[28] Moore, R. G., Brown, A. K., Miller, M. C., Badgwell, D. et al., Utility of a novel serum tumor biomarker HE4 in patients with endometrioid adenocarcinoma of the uterus. *Gynecol. Oncol.* 2008, *110*, 196–201.

[29] Hareyama, H., Sakuragi, N., Makinoda, S., Fujimoto, S., Serum and tissue measurements of CA72-4 in patients with endometrial carcinoma. *J. Clin. Pathol.* 1996, *49*, 967–970.

[30] Konno, R., Takano, T., Sato, S., Yajima, A., Serum soluble Fas level as a prognostic factor in patients with gynecological malignancies. *Clin. Cancer Res.* 2000, *6*, 3576–3580.

[31] Sawada, M., Okudaira, Y., Matsui, Y., Shimizu, Y., Immunosuppressive acidic protein in patients with gynecologic cancer. *Cancer* 1984, *54*, 652–656.

[32] Hakala, A., Kacinski, B. M., Stanley, E. R., Kohorn, E. I. et al., Macrophage colony-stimulating factor 1, a clinically useful tumor marker in endometrial adenocarcinoma: comparison

© 2015 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

with CA 125 and the aminoterminal propeptide of type III procollagen. *Am. J. Obst. Gynecol.* 1995, *173*, 112–119.

[33] Suzuki, M., Ohwada, M., Sato, I., Nagatomo, M., Serum level of macrophage colony—stimulating factor as a marker for gynecologic malignancies. *Oncology* 1995, *52*, 128–133.

[34] Mu, A. K., Lim, B. K., Hashim, O. H., Shuib, A. S., Detection of differential levels of proteins in the urine of patients with endometrial cancer: analysis using two-dimensional gel electrophoresis and o-glycan binding lectin. *Int. J. Mol. Sci.* 2012, *13*, 9489–9501.

[35] Frederick, P. J., Straughn, J. M., Jr., The role of comprehensive surgical staging in patients with endometrial cancer. *Cancer Control Center* 2009, *16*, 23–29.

[36] Larson, D. M., Connor, G. P., Broste, S. K., Krawisz, B. R., Johnson, K. K., Prognostic significance of gross myometrial invasion with endometrial cancer. *Obst. Gynecol.* 1996, *88*, 394–398.

[37] 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. *Obst. Gynecol.* 1984, *63*, 825–832.

[38] Cocco, E., Bellone, S., El-Sahwi, K., Cargnelutti, M. et al., Serum amyloid A: a novel biomarker for endometrial cancer. *Cancer* 2010, *116*, 843–851.

[39] Lo, S. S., Cheng, D. K., Ng, T. Y., Wong, L. C., Ngan, H. Y., Prognostic significance of tumour markers in endometrial cancer. *Tumor Biol.* 1997, *18*, 241–249.

[40] Lambropoulou, M., Papadopoulos, N., Tripsianis, G., Alexiadis, G. et al., Co-expression of survivin, c-erbB2, and cyclooxygenase-2 (COX-2): prognostic value and survival of endometrial cancer patients. *J. Cancer Res. Clin. Oncol.* 2010, *136*, 427–435.

[41] Yilmaz, E., Koyuncuoglu, M., Gorken, I. B., Okyay, E. et al., Expression of matrix metalloproteinase-2 and survivin in endometrioid and nonendometrioid endometrial cancers and clinicopathologic significance. *J. Gynecol. Oncol.* 2011, *22*, 89–96.

[42] Zeimet, A. G., Reimer, D., Huszar, M., Winterhoff, B. et al., L1CAM in early-stage type I endometrial cancer: results of a large multicenter evaluation. *J. Natl. Cancer Inst.* 2013, *105*, 1142–1150.

[43] Bansal, N., Yendluri, V., Wenham, R. M., The molecular biology of endometrial cancers and the implications for pathogenesis, classification, and targeted therapies. *Cancer Control* 2009, *16*, 8–13.

[44] Hecht, J. L., Mutter, G. L., Molecular and pathologic aspects of endometrial carcinogenesis. *J. Clin. Oncol.* 2006, *24*, 4783–4791.

[45] Arango, H. A., Hoffman, M. S., Roberts, W. S., DeCesare, S. L. et al., Accuracy of lymph node palpation to determine need for lymphadenectomy in gynecologic malignancies. *Obst. Gynecol.* 2000, *95*, 553–556.

[46] Bilbao-Sieyro, C., Ramirez, R., Rodriguez-Gonzalez, G., Falcon, O. et al., Microsatellite instability and ploidy status define three categories with distinctive prognostic impact in endometrioid endometrial cancer. *Oncotarget* 2014, *5*, 6206–6217.

[47] Bilbao, C., Rodriguez, G., Ramirez, R., Falcon, O. et al., The relationship between microsatellite instability and PTEN gene mutations in endometrial cancer. *Int. J. Cancer* 2006, *119*, 563–570.

[48] Lax, S. F., Kendall, B., Tashiro, H., Slebos, R. J., Hedrick, L., The frequency of p53, K-ras mutations, and microsatellite instability differs in uterine endometrioid and serous carcinoma: evidence of distinct molecular genetic pathways. *Cancer* 2000, *88*, 814–824.

[49] Moreno-Bueno1, G., Hardisson, D., Sánchez, C., Sarrió, D. et al., Abnormalities of the APC/b-catenin pathway in endometrial cancer. *Oncogene* 2002, *21*, 7981–7990.

[50] Saegusa, M., Hashimura, M., Yoshida, T., Okayasu, I., β-Catenin mutations and aberrant nuclear expression during endometrial tumorigenesis. *Br. J. Cancer* 2001, *84*, 209–217.

[51] Lluis Catasus, A. G., Miriam cuatrecasas and Jaime Prat. Concomitant PI3K–AKT and p53 alterations in endometrial carcinomas are associated with poor prognosis. *Modern Pathol.* (2009). 2009, *22*, 522–529.

[52] Doll, A., Abal, M., Rigau, M., Monge, M. et al., Novel molecular profiles of endometrial cancer-new light through old windows. *J. Steroid Biochem. Mol. Biol.* 2008, *108*, 221–229.

[53] Takes, R. P., Baatenburg de Jong, R. J., Wijffels, K., Schuuring, E. et al., Expression of genetic markers in lymph node metastases compared with their primary tumours in head and neck cancer. *J. Pathol.* 2001, *194*, 298–302.

[54] René Bernards, Weinberg, R. A., A progression puzzle. 2002, *418*, 823.

[55] Fidler, I. J., Kripke, M. L., Metastasis results from preexisting variant cells within a malignant tumor. *Science* 1977, *197*, 893–895.

[56] Ramaswamy, S., Ross, K. N., Lander, E. S., Golub, T. R., A molecular signature of metastasis in primary solid tumors. *Nat. Genet.* 2003, *33*, 49–54.

[57] Yi, Z., Jingting, C., Yu, Z., Proteomics reveals protein profile changes in cyclooxygenase-2 inhibitor-treated endometrial cancer cells. *Int. J. Gynecol. Cancer* 2009, *19*, 326–333.

[58] Maxwell, G. L., Hood, B. L., Day, R., Chandran, U. et al., Proteomic analysis of stage I endometrial cancer tissue: identification of proteins associated with oxidative processes and inflammation. *Gynecol. Oncol.* 2011, *121*, 586–594.

[59] Aebersold, R., Mann, M., Mass spectrometry-based proteomics. *Nature* 2003, *422*, 198–207.

[60] Hochstrasser, D. F., Proteome in perspective. *Clin. Chem. Lab. Med.* 1998, *36*, 825–836.

[61] Pertea, M., Salzberg, S. L., Between a chicken and a grape: estimating the number of human genes. *Genome Biol.* 2010, *11*, 206.

[62] Ewing, B., Green, P., Analysis of expressed sequence tags indicates 35,000 human genes. *Nat. Genet.* 2000, *25*, 232–234.

[63] Ong, S.-E., Whole proteomes as internal standards in quantitative proteomics. *Genome Med.* 2010, *2*, 49.

[64] Kulasingam, V., Diamandis, E. P., Tissue culture-based breast cancer biomarker discovery platform. *Int. J. Cancer* 2008, *123*, 2007–2012.

© 2015 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

[65] Grossmann, J., Roschitzki, B., Panse, C., Fortes, C. et al., Implementation and evaluation of relative and absolute quantification in shotgun proteomics with label-free methods. *J. Proteomics* 2010, *73*, 1740–1746.

[66] Arentz, G., Weiland, F., Oehler, M. K., Hoffmann, P., State of the art of 2D DIGE. *Proteomics* 2015, *9*(3–4), 277–288.

[67] Tiss, A., Timms, J., Menon, U., Gammerman, A., Cramer, R., Proteomics approaches towards early detection and diagnosis of ovarian cancer. *J. Immunother. Cancer* 2014, *2*(Suppl 1), O5.

[68] Tran, J. C., Zamdborg, L., Ahlf, D. R., Lee, J. E. et al., Mapping intact protein isoforms in discovery mode using top-down proteomics. *Nature* 2011, *480*, 254–258.

[69] Liu, H., Sadygov, R. G., Yates, J. R., A model for random sampling and estimation of relative protein abundance in shotgun proteomics. *Anal. Chem.* 2004, *76*, 4193–4201.

[70] Alonso-Alconada, L., Santacana, M., Garcia-Sanz, P., Muinelo-Romay, L. et al., Annexin-A2 as predictor biomarker of recurrent disease in endometrial cancer. *Int. J. Cancer* 2014, *136*, 1863–1873.

[71] Monge, M., Colas, E., Doll, A., Gil-Moreno, A. et al., Proteomic approach to ETV5 during endometrial carcinoma invasion reveals a link to oxidative stress. *Carcinogenesis* 2009, *30*, 1288–1297.

[72] Ihata, Y., Miyagi, E., Numazaki, R., Muramatsu, T. et al., Amino acid profile index for early detection of endometrial cancer: verification as a novel diagnostic marker. *Int. J. Clin. Oncol.* 2014, *19*, 364–372.

[73] Qiu, F., Gao, Y. H., Jiang, C. G., Tian, Y. P., Zhang, X. J., Serum proteomic profile analysis for endometrial carcinoma detection with MALDI-TOF MS. *Arch. Med. Sci.* 2010, *6*, 245–252.

[74] Hillenkamp, F., Karas, M., Beavis, R. C., Chait, B. T., Matrix-assisted laser desorption/ionization mass spectrometry of biopolymers. *Anal. Chem.* 1991, *63*, 1193A–1203A.

[75] Casado-Vela, J., Rodriguez-Suarez, E., Iloro, I., Ametzazurra, A. et al., Comprehensive proteomic analysis of human endometrial fluid aspirate. *J. Proteome Res.* 2009, *8*, 4622–4632.

[76] Henry, N. L., Hayes, D. F., Cancer biomarkers. *Mol. Oncol.* 2012, *6*, 140–146.

[77] Meehan, K. L., Rainczuk, A., Salamonsen, L. A., Stephens, A. N., Proteomics and the search for biomarkers of female reproductive diseases. *Reproduction* 2010, *140*, 505–519.

[78] Ong, S.-E., Mann, M., Mass spectrometry-based proteomics turns quantitative. *Nat. Chem. Biol.* 2005, *1*, 252–262.

[79] Arentz, G., Weiland, F., Oehler, M. K., Hoffmann, P., State of the art of 2D DIGE. *Proteomics Clin. Appl.* 2014, *9*, 277–288.

[80] Wang, N., Li, L., Exploring the precursor ion exclusion feature of liquid chromatography–electrospray ionization quadrupole time-of-flight mass spectrometry for improving protein identification in shotgun proteome analysis. *Anal. Chem.* 2008, *80*, 4696–4710.

[81] Tate, S., Larsen, B., Bonner, R., Gingras, A.-C., Label-free quantitative proteomics trends for protein–protein interactions. *J. Proteomics* 2013, *81*, 91–101.

[82] Vowinckel, J., Capuano, F., Campbell, K., Deery, M. J. et al., The beauty of being (label)-free: sample preparation methods for SWATH-MS and next-generation targeted proteomics. *F1000Research* 2013, *2*, 272.

[83] Cornett, D. S., Reyzer, M. L., Chaurand, P., Caprioli, R. M., MALDI imaging mass spectrometry: molecular snapshots of biochemical systems. *Nat. Meth.* 2007, *4*, 828–833.

[84] Bateson, H., Saleem, S., Loadman, P. M., Sutton, C. W., Use of matrix-assisted laser desorption/ionisation mass spectrometry in cancer research. *J. Pharmacol. Toxicol. Methods* 2011, *64*, 197–206.

[85] Kriegsmann, J., Kriegsmann, M., Casadonte, R., MALDI TOF imaging mass spectrometry in clinical pathology: a valuable tool for cancer diagnostics (review). *Int. J. Oncol.* 2014, *46*, 893–906.

[86] Quaas, A., Bahar, A. S., von Loga, K., Seddiqi, A. S. et al., MALDI imaging on large-scale tissue microarrays identifies molecular features associated with tumour phenotype in oesophageal cancer. *Histopathology* 2013, *63*, 455–462.

[87] Steurer, S., Borkowski, C., Odinga, S., Buchholz, M. et al., MALDI mass spectrometric imaging based identification of clinically relevant signals in prostate cancer using large-scale tissue microarrays. *Int. J. Cancer* 2013, *133*, 920–928.

[88] Breuer, E. K., Murph, M. M., The role of proteomics in the diagnosis and treatment of women's cancers: current trends in technology and future opportunities. *Int. J. Proteomics* 2011, *2011*.

[89] Wang, P., Whiteaker, J. R., Paulovich, A. G., The evolving role of mass spectrometry in cancer biomarker discovery. *Cancer Biol. Ther.* 2009, *8*, 1083–1094.

[90] Rauser, S., Marquardt, C., Balluff, B., Deininger, S.-O. et al., Classification of HER2 receptor status in breast cancer tissues by MALDI imaging mass spectrometry. *J. Proteome Res.* 2010, *9*, 1854–1863.

[91] Meding, S., Nitsche, U., Balluff, B., Elsner, M. et al., Tumor classification of six common cancer types based on proteomic profiling by MALDI imaging. *J. Proteome Res.* 2012, *11*, 1996–2003.

[92] Addona, T. A., Shi, X., Keshishian, H., Mani, D. R. et al., A pipeline that integrates the discovery and verification of plasma protein biomarkers reveals candidate markers for cardiovascular disease. *Nat. Biotechnol.* 2011, *29*, 635–643.

[93] Klein1, O., Nebrich, G., Trede, D., Becker, M. et al., SCiLS Lab 2D: Comparative Analysis for Uncovering Discriminative M/z-markers, 2015, SCiLS application note#5.

[94] Shi, T., Fillmore, T. L., Sun, X., Zhao, R. et al., Antibody-free, targeted mass-spectrometric approach for quantification of proteins at low picogram per milliliter levels in human plasma/serum. *Proc. Natl. Acad. Sci. USA* 2012, *109*, 15395–15400.

[95] Huttenhain, R., Malmstrom, J., Picotti, P., Aebersold, R., Perspectives of targeted mass spectrometry for protein biomarker verification. *Curr. Opin. Chem. Biol.* 2009, *13*, 518–525.

[96] Cohen Freue, G. V., Borchers, C. H., Multiple reaction monitoring (MRM): principles and application to coronary artery disease. *Circ. Cardiovasc. Genet.* 2012, *5*, 378.

© 2015 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

[97] Kitteringham, N. R., Jenkins, R. E., Lane, C. S., Elliott, V. L., Park, B. K., Multiple reaction monitoring for quantitative biomarker analysis in proteomics and metabolomics. *J. Chromatogr. B, Analy. Technol. Biomed. Life Sci.* 2009, *877*, 1229–1239.

[98] DeSouza, L. V., Taylor, A. M., Li, W., Minkoff, M. S. et al., Multiple reaction monitoring of mTRAQ-labeled peptides enables absolute quantification of endogenous levels of a potential cancer marker in cancerous and normal endometrial tissues. *J. Proteome Res.* 2008, *7*, 3525–3534.

[99] Clemis, E. J., Smith, D. S., Camenzind, A. G., Danell, R. M. et al., Quantitation of spatially-localized proteins in tissue samples using MALDI-MRM imaging. *Anal. Chem.* 2012, *84*, 3514–3522.

[100] Creasman, W., Revised FIGO staging for carcinoma of the endometrium. *Int. J. Gynaecol. Obst.* 2009, *105*, 109.

© 2015 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

Appendix IV

Lymph Node Metastasis of Primary Endometrial Cancers: Associated Proteins Revealed by MALDI Imaging

This page intentionally left blank

RESEARCH ARTICLE

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] and Peter Hoffmann[1,2**]*

[1] Adelaide Proteomics Centre, School of Biological Sciences, University of Adelaide, Adelaide, South Australia, Australia
[2] Institute for Photonics and Advanced Sensing (IPAS), University of Adelaide, Adelaide, South Australia, Australia
[3] Department of Tissue Pathology and Diagnostic Oncology, Royal Prince Alfred Hospital, Camperdown, New South Wales, Australia
[4] Faculty of Health and Medicine, University of New South Wales, Callaghan, New South Wales, Australia
[5] School of Women's and Infants' Health, University of Western Australia, Crawley, Western Australia, Australia
[6] Department of Gynaecological Oncology, Chris O'Brien Lifehouse, Camperdown, New South Wales, Australia
[7] Institute for Research in Molecular Medicine, Universiti Sains Malaysia, 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, Australia
[9] Department of Gynaecological Oncology, Royal Adelaide Hospital, Adelaide, South Australia, Australia

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.

Received: November 17, 2015
Revised: February 12, 2016
Accepted: April 5, 2016

Keywords:
Biomarker / Biomedicine / Endometrial cancer / Lymph node metastasis / Proteomics

Additional supporting information may be found in the online version of this article at the publisher's web-site

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

Correspondence: Professor Peter Hoffmann, Adelaide Proteomics Centre, University of Adelaide, Gate 8 Victoria Drive, Adelaide, South Australia, 5005
E-mail: peter.hoffmann@adelaide.edu.au
Fax: +61 (08) 0 8313 4362

Abbreviations: **α-Actin-2**, alpha-actin-2; **CCA**, canonical correlation analysis; **EC**, endometrial cancer; **FFPE**, formalin fixed paraffin embedded; **LC-MS/MS**, liquid chromatography mass spectrometry; **LDA**, linear discriminant analysis; **LNM**, lymph node metastasis; **LOO**, leave one out; **MALDI-MSI**, matrix assisted laser desorption/ionisation mass spectrometry imaging

*These authors contributed equally to the work.
**To whom all correspondence and requests for reprints should be addressed.

© 2016 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

Significance of the study

Endometrial cancer is the most common gynaecological malignancy in Australia with 2256 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.

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. categorised EC patients as low risk for LNM if the tumour size was $\leq$ 2cm, well or moderately differentiated, and depth of myometrial invasion was $\leq$ 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 ionisation tandem mass spectrometry (nanoLC-ESI-MS/MS). The

© 2016 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

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 Materials 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 Supporting Information 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 h. 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 visualisation of tissue morphological regions, raw data was uploaded into the SCiLS lab software (v2015a, GmbH, Bremen, Germany). Here the data was preprocessed including top hat baseline removal and total ion count (TIC) normalisation, 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

© 2016 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

matching back to peptide sequences obtained by data dependent acquisition, nanoLC-ESI-MS/MS, primary tumour 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 min. 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 min, followed by a 20 min wash with 90% B, and a 20 min 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, deconvolution, 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 two 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 four patients with, and four 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 one 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 min 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 min at 100°C (Sixth Sense, Whirlpool, VIC, Australia) in 10 mM citric acid buffer pH = 6. TMA sections were incubated overnight at 4°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 photomicrographic 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 Supporting Information 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,

© 2016 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

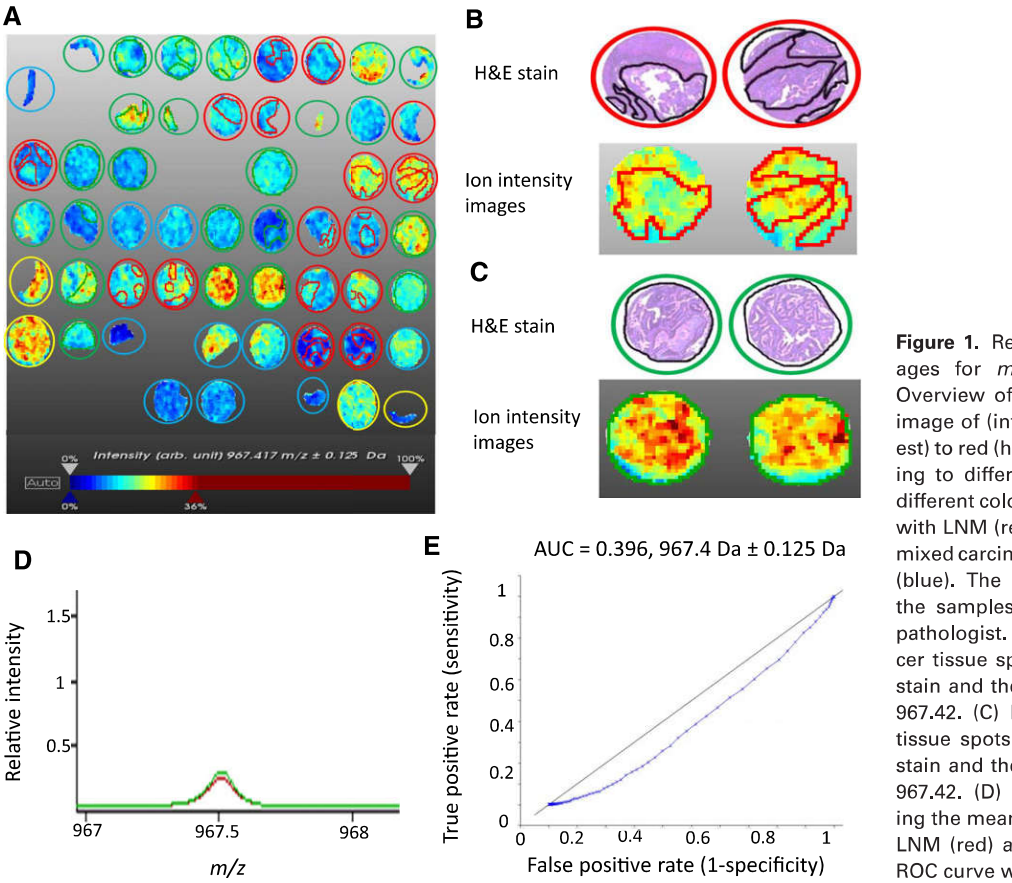


Figure 1. Representative MALDI-MSI images for *m/z* 967.42 ± 0.125 Da. (A) Overview of one TMA slide, MALDI-MSI image of (intensity range from blue (lowest) to red (highest)). The samples belonging to different groups are indicated by different coloured circles: control (yellow), with LNM (red), without 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.

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 (Supporting Information 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* 1198.67 AVFPSIVGRPR (α-Actin-2), and *m/z* 976.42 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).

Figures 1 and 2 show 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 (Figs. 1B, 2B) and without LNM (Figs. 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 (Fig. 1). However, the SCiLS analysis of *m/z* 976.42 revealed a difference between the ion intensity map between tumours with and without LNM (Fig. 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

© 2016 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

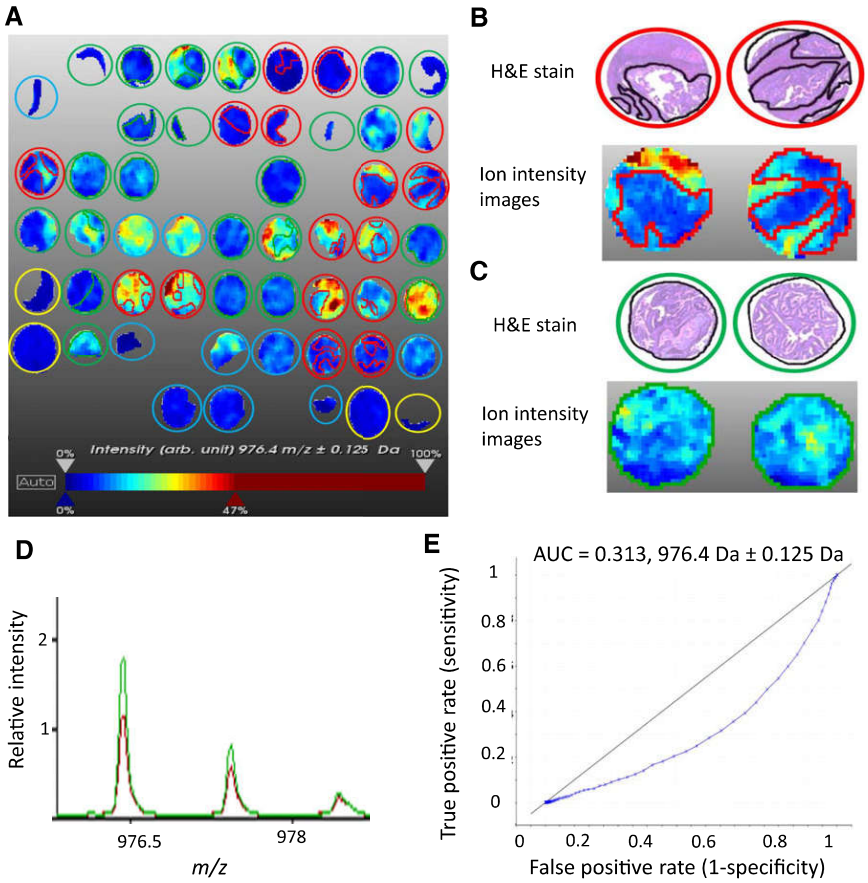


Figure 2. Representative MALDI-MSI images for *m/z* 976.42 ± 0.125 Da. (A) Overview of one TMA slide, MALDI-MSI image of (intensity range from blue (lowest) to red (highest)). The samples belonging to different groups are indicated by different coloured circles: control (yellow), with LNM (red), without 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.

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 (Fig. 3).

3.4 Validation by immunohistochemistry

The spatial expression profile of plectin was verified across the patient cohort by immunohistochemistry (Fig. 4A).

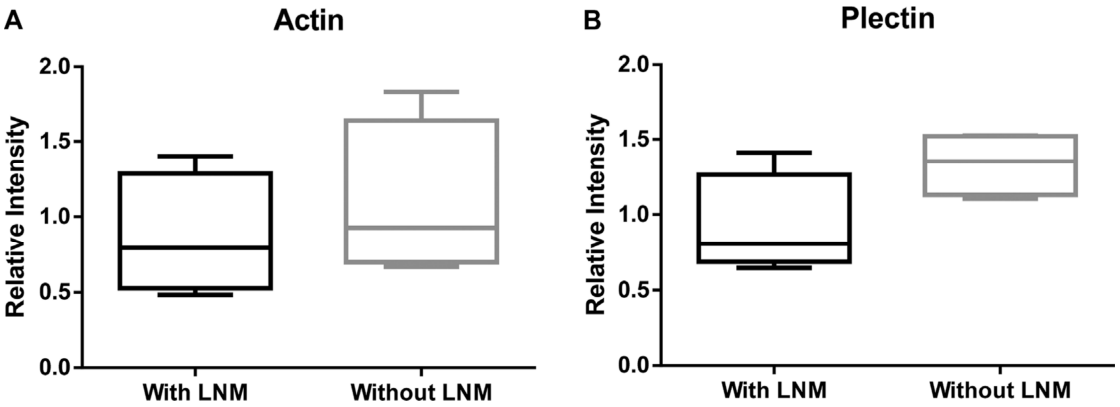


Figure 3. DIA analysis of tumour sections with ($n = 4$) and without ($n = 4$) LNM. (A) The relative abundance of plectin was analysed in comparison to normal tissue set to 1.0. (B) The relative abundance of α-Actin-2 was analysed in comparison to normal tissue set to 1.0. The error bar indicates the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

© 2016 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

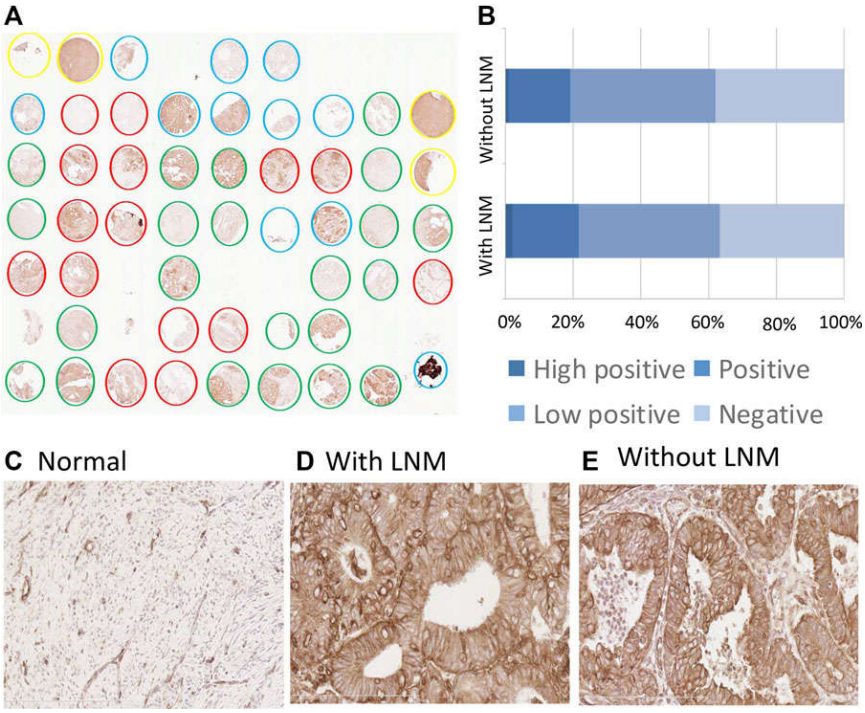


Figure 4. Immunohistochemical staining of plectin (A) 6 μm serial section of TMA 1 was used for immunohistochemistry (IHC). The different tissue types are encircled control (yellow), with LNM (red), without LNM (green) and mixed carcinoma/not included in the study (blue). (B) Quantitative analysis was 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).

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 (Fig. 4B). Plectin scarcely stained normal tissue including stroma (Fig. 4C); the staining of tumours cells was strong (Figs. 4D–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 organisation [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 Fig. 1 and has been confirmed by DIA (Fig. 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 (Fig. 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 (Fig. 5B). α-Actin-2 stained normal tissue (Fig. 5C),while the staining of tumours with LNM was reduced (Fig. 5D) when compared to tumour without LNM (Fig. 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

© 2016 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

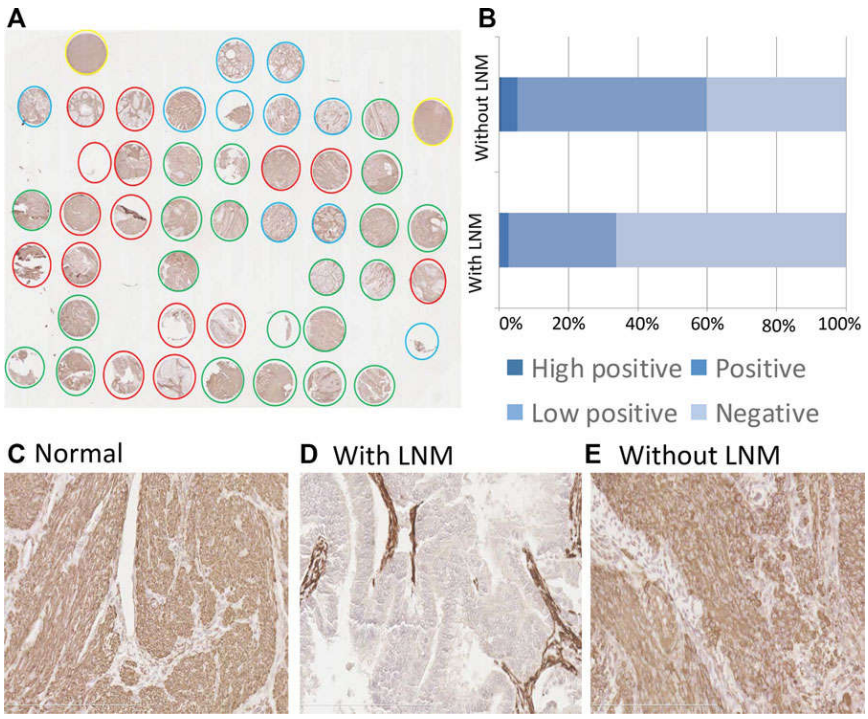


Figure 5. Immunohistochemical staining of α-Actin-2 (A) 6 μm serial section of TMA 1 was used for immunohistochemistry (IHC). The different tissue types are encircled control (yellow), with LNM (red), without LNM (green) and mixed carcinoma/not included in the study (blue). (B) Quantitative analysis was performed using IHC profiler-Image J. For each tissue section, three representative photomicrographic 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).

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 (Supporting Information Fig. 1) and label free quantification (Supporting Information Table 2). Furthermore, DIA (Supporting Information Figs. 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.

5 References

[1] Bokhman, J. V., Two pathogenetic types of endometrial carcinoma. *Gynecol. Oncol.* 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. *Obstet. Gynecol.* 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. *Obstet. Gynecol.* 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. *Obstet. Gynecol.* 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.*Int. J. Gynecol. Pathol.* 1998, *17*, 36–41.

[7] Creasman, W. T., Odicino, F., Maisonneuve, P., Quinn, M. A. et al., Carcinoma of the Corpus Uteri. *Int. J. Gynecol. Obstet.* 2006, *95*, Supplement 1, S105–S143.

[8] Lodish, H., Berka, A., Zipursky, S. L., Matsudaira, P. 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. *J. Proteome Res.* 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. *J. Pathol.* 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. *Mod. Pathol.* 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 J. Gastroenterol.* 2009, *15*, 2228–2233.

© 2016 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

[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. *J. Natl. Cancer Inst.* 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. *Hum. Pathol.* 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. *J. Proteome Res.* 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 Sci.* 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. *J. Proteome Res.* 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. *Anal. Bioanal. Chem.* 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* 2015, *10*, 217–229.

[22] Winderbaum, L., Koch, I., Mittal, P., Hoffmann, P., Canonical correlation analysis based variable ranking for the classification of MALDI-MS imaging data of tissue microarrays. *Proteomics* 2016, doi: 10.1002/pmic.201500451.

[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. *J. Proteome Res.* 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 Commun. Mass Spectrom.* 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. *J. 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. *Int. J. 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. *Anal. Chem.* 2015, *87*, 11448–11454.

[31] Liotta, L. A., Stracke, M. L., Aznavoorian, S. A., Beckner, M. E., Schiffmann, E., Tumor cell motility. *Semin. Cancer Biol.* 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. *Clin. Cancer Res.* 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. *Scientific World Journal* 2012, 252034.

[39] Rønnov-Jessen, L., Petersen, O. W., A function for filamentous alpha-smooth muscle actin: retardation of motility in fibroblasts. *J. Cell Biol.* 1996, *134*, 67–80.

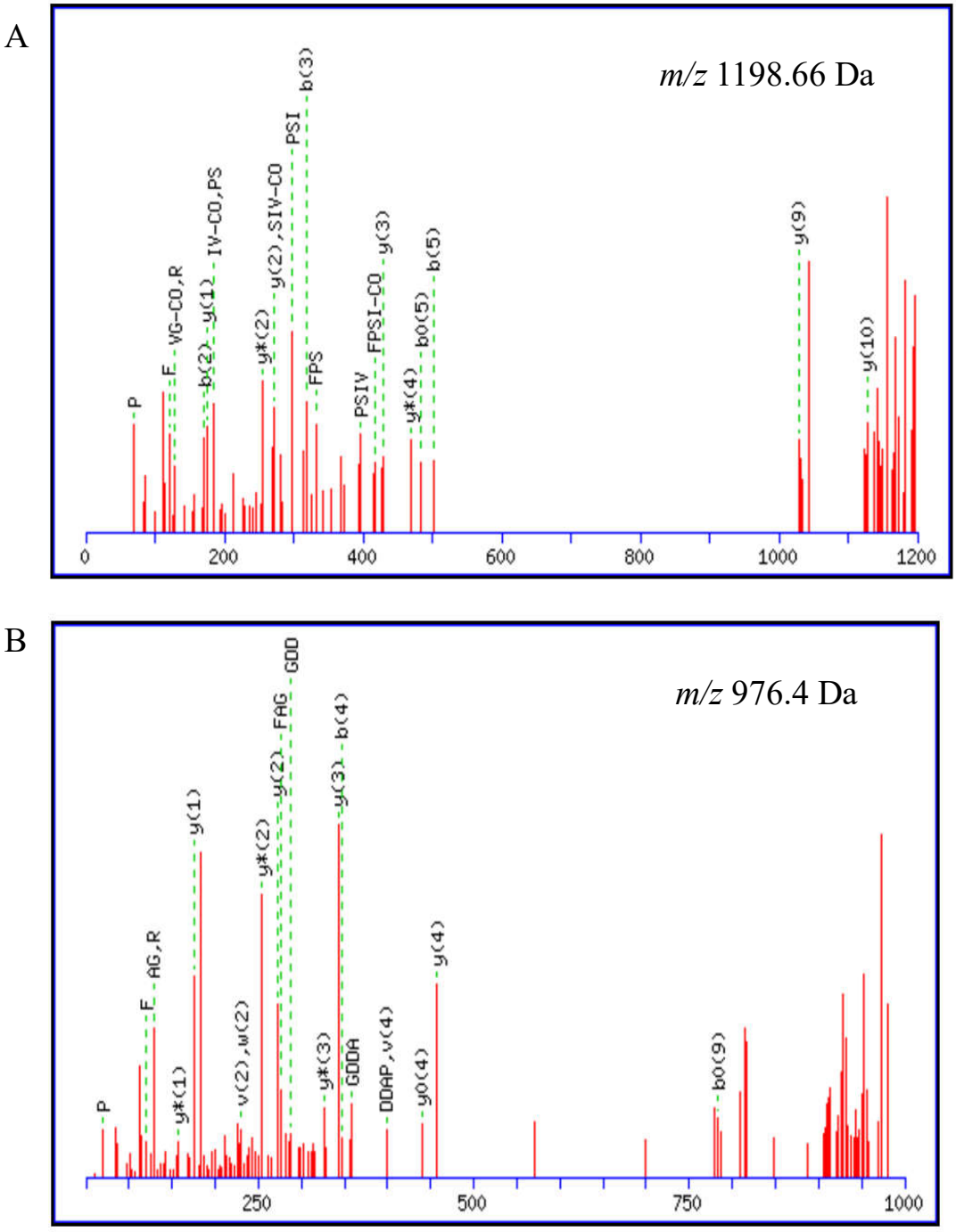
[40] Uzquiano, M. C., Prieto, V. G., Nash, J. W., Ivan, D. S. et al., Metastatic basal cell carcinoma exhibits reduced actin expression. *Mod. Pathol.* 2008, *21*, 540–543.

© 2016 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

This page intentionally left blank

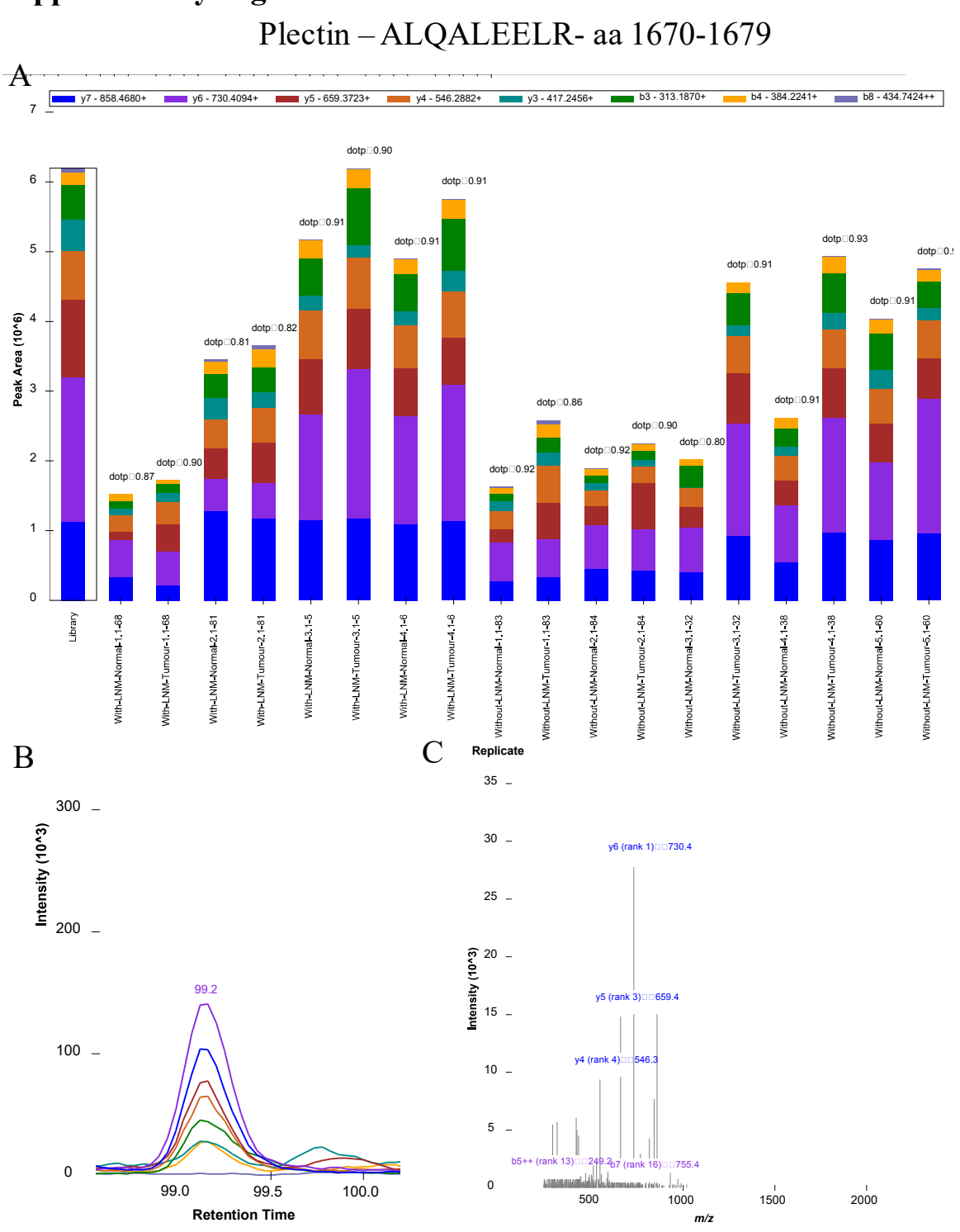
4.4 Supplementary File

Supplementary Figure 1



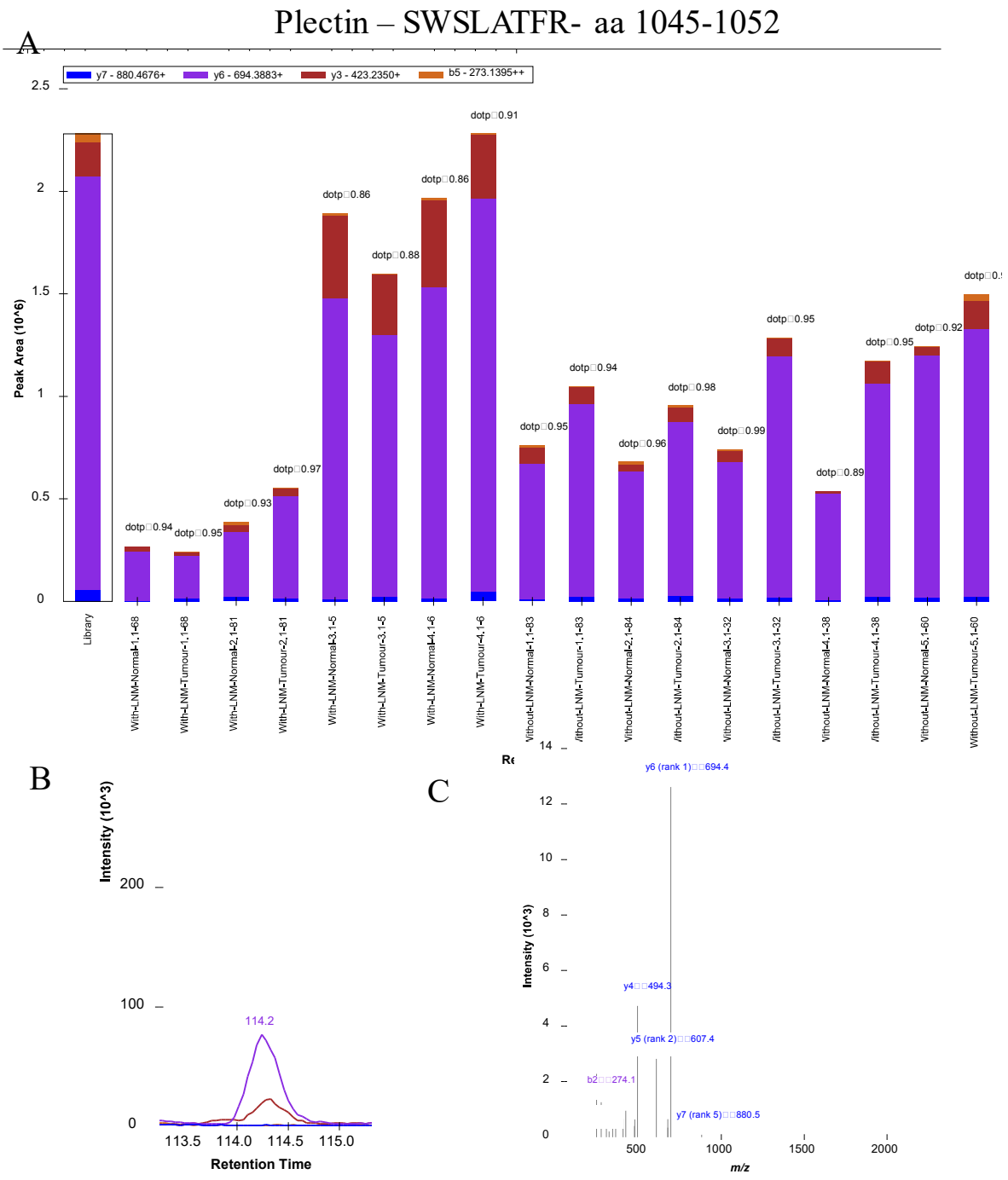
Supplementary Figure 1: Tryptic peptide identified by *in situ* MS/MS as α-Actin-2, Aortic smooth muscle (ACTA_Human). 6µm FFPE EC TMA was analysed by tryptic peptide MALDI-MSI followed by *in situ* MS/MS on an ultrafleXtreme MALDI-TOF/TOF instrument. Selected *m/z* values from Supplementary Table 2 were fragmented *in situ* to generate a list of daughter ions to identify the peptide sequence (A) *m/z* 1198.66 Da AVFPSIVGRPR (B) *m/z* 976.4 Da AGFAGDDAPR

Supplementary Figure 2



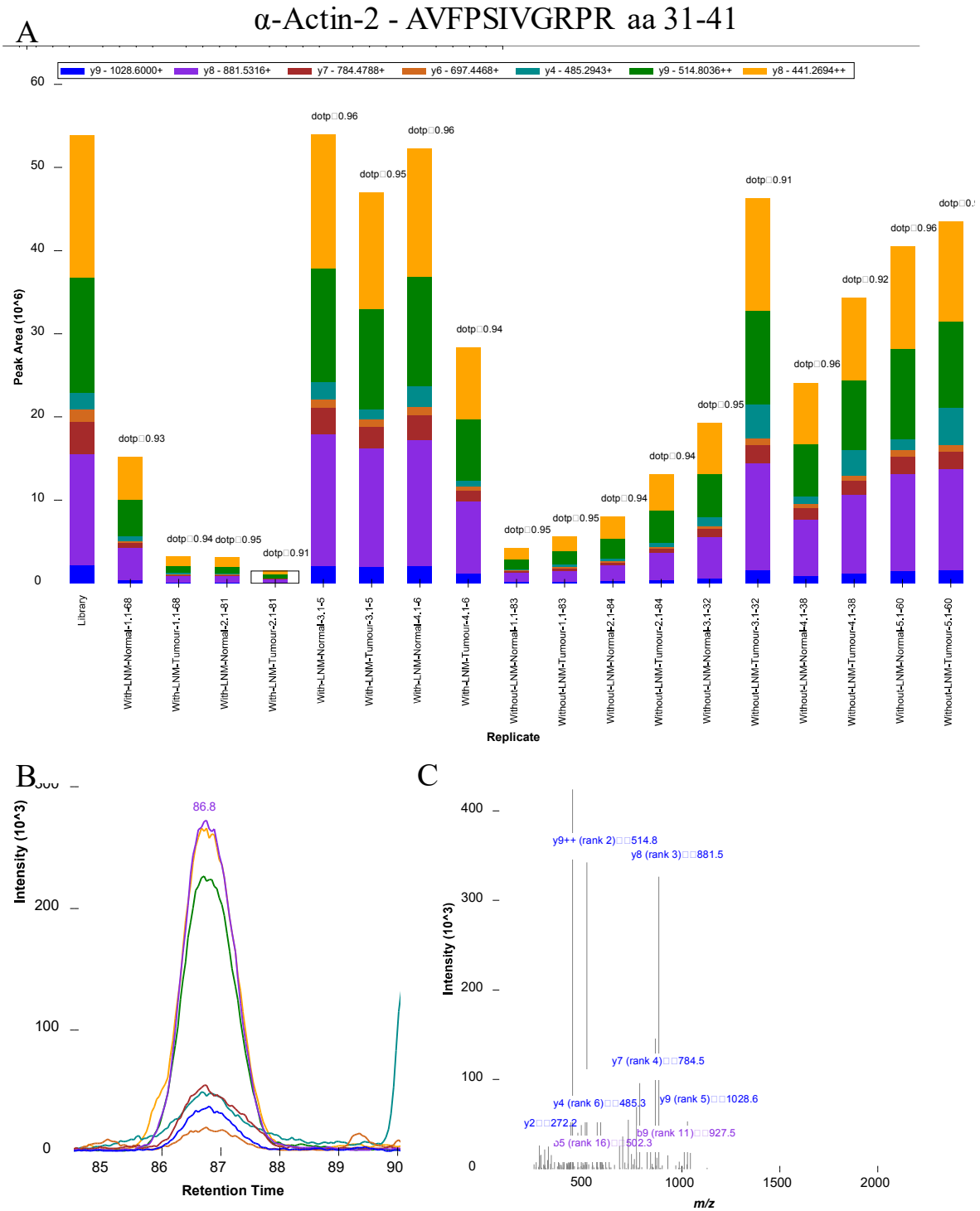
Supplementary Figure 2: DIA analysis confirmed proteins associated with LNM. The relative abundance of plectin peptide ALQALEELR was analysed by DIA of with (n=4) and without (n=4) LNM tumour sections in comparison to their paired normal tissue.(A) Peak area intensity (B) Peptide transition (C) Spectral Library

Supplementary Figure 3



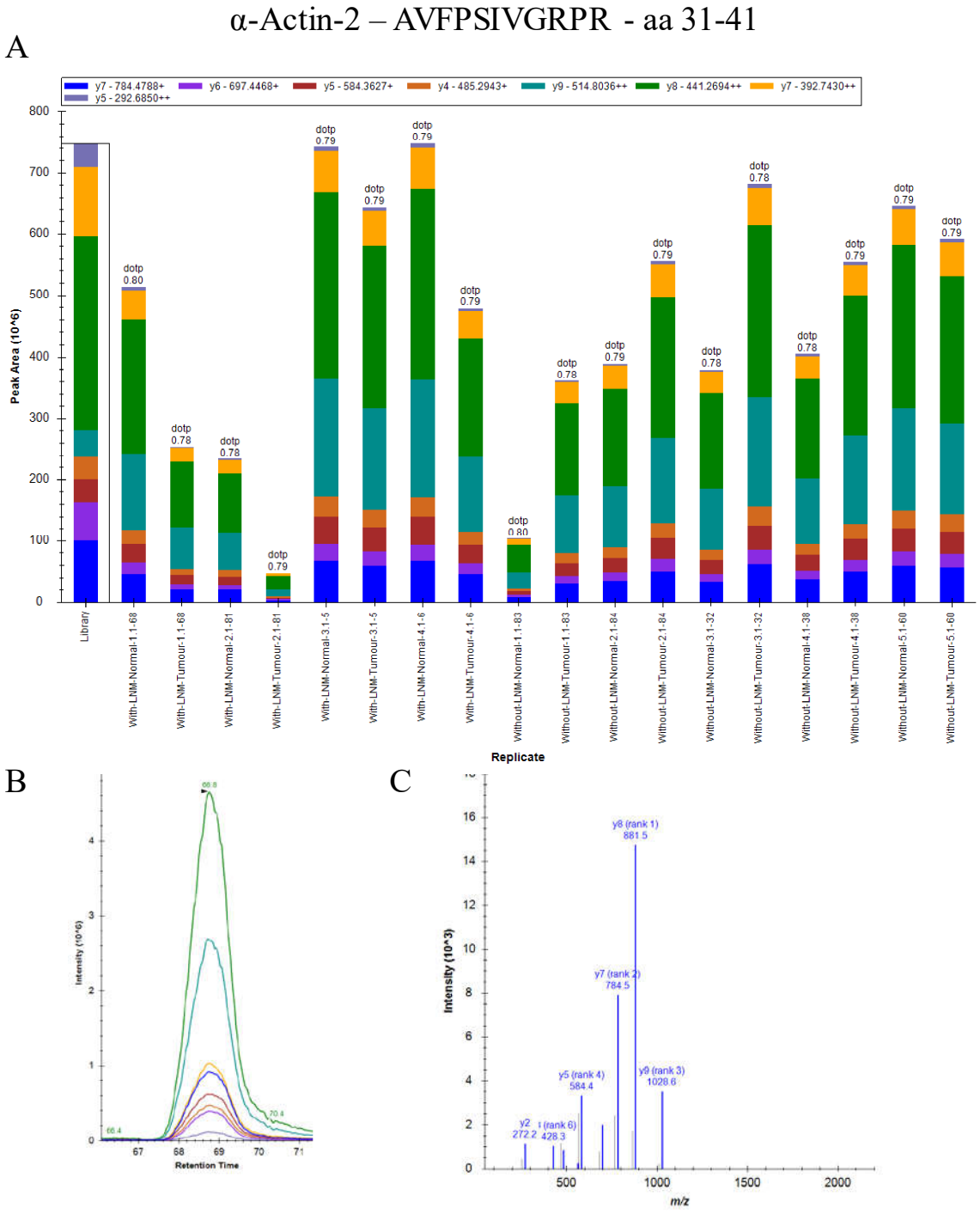
Supplementary Figure 3: DIA analysis confirmed proteins associated with LNM. The relative abundance of plectin peptide SWSLATFR was analysed by DIA of with (n=4) and without (n=4) LNM tumour sections in comparison to their paired normal tissue.(A) Peak area intensity (B) Peptide transition (C) Spectral Library

Supplementary Figure 4



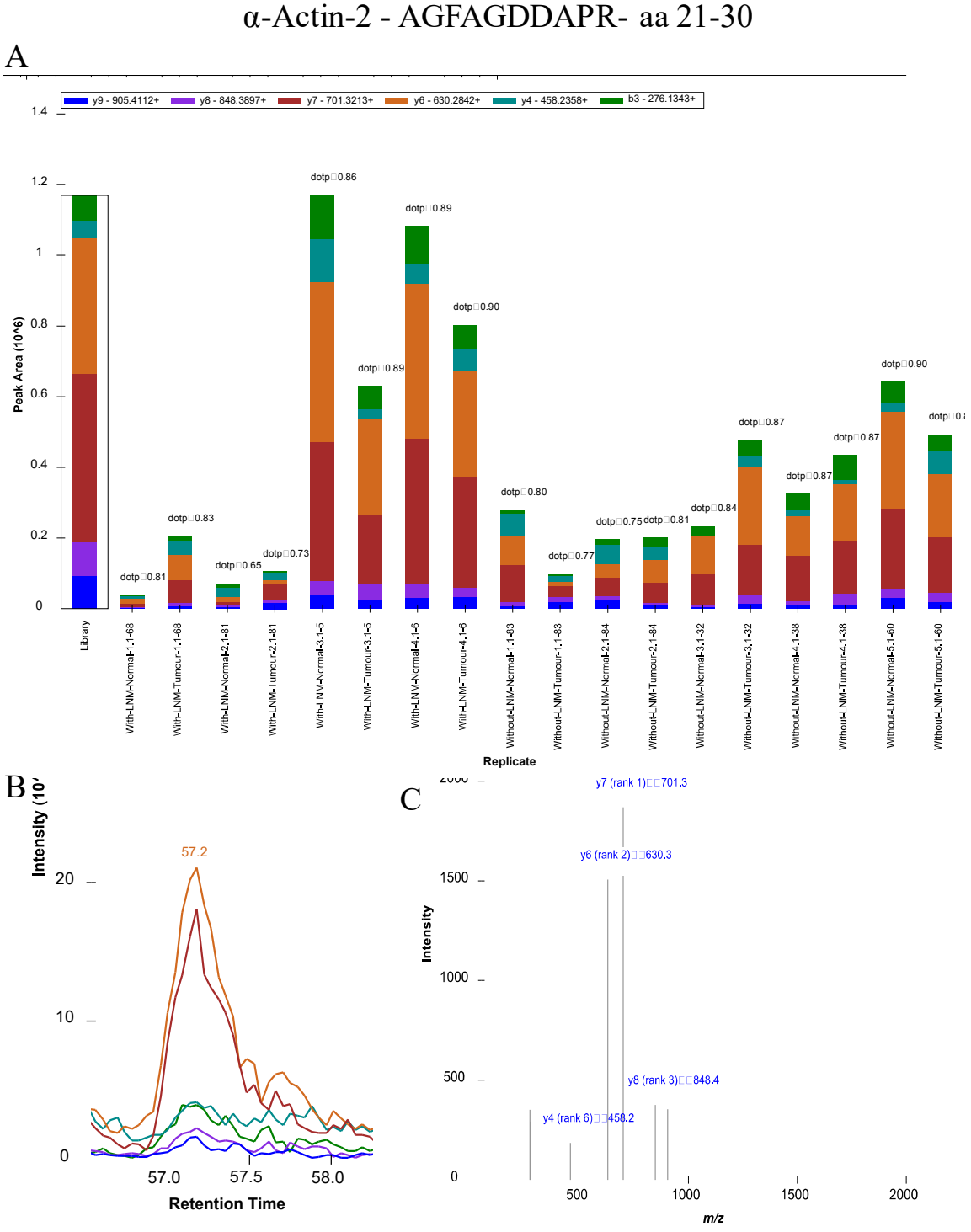
Supplementary Figure 4: DIA analysis confirmed proteins associated with LNM. The relative abundance of α-Actin-2 peptide AVFPSIVGRPR was analysed by DIA of with (n=4) and without (n=4) LNM tumour sections in comparison to their paired normal tissue.(A) Peak area intensity (B) Peptide transition (C) Spectral Library

Supplementary Figure 5



Supplementary Figure 5: DIA analysis confirmed proteins associated with LNM. The relative abundance of α-Actin-2 peptide AVFPSIVGRPR was analysed by DIA of with (n=4) and without (n=4) LNM tumour sections in comparison to their paired normal tissue.(A) Peak area intensity (B) Peptide transition (C) Spectral Library

Supplementary Figure 6



Supplementary Figure 6: DIA analysis confirmed proteins associated with LNM. The relative abundance of α-Actin-2 peptide AGFAGDDAPR was analysed by DIA of with (n=4) and without (n=4) LNM tumour sections in comparison to their paired normal tissue.(A) Peak area intensity (B) Peptide transition (C) Spectral Library

Supplementary Table 1: Clinicopathological characteristics of endometrial cancer patient samples used in the study

Histology	No. of patients
Histological stage (FIGO 2009)	
Stage I	27
Stage II	0
Stage III	16
Histological grade	
G1	20
G2	14
G3	9
Distance to Serosa	
≤ 5mm	16
> 5mm	26
Information not available	1
Myometrium invasion (MI)	
MI ≤ 5mm	16
MI > 5mm	25
Information not available	2
Lymph Vascular Space Invasion (LVSI)	
Present	21
Absent	22
Lymph node metastasis	
Present	16
Absent	27

Supplementary Table 2: All possible nano-LC-ESI-MS/MS peptide matches to the top *m/z* peak groups as ranked by the CCA

Ranked m/z values by CCA [1]	Protein Accession Name	Peptide experimental m/z	Peptide experimental charge	Peptide experimental MR	Peptide calculated MR	Peptide score, Mascot	Peptide expected score, Mascot	Peptide sequence
967.417	PLEC_HUMAN	484.256	2	966.4974	966.4923	29.76	0.05	SWSLATFR
	NXN_HUMAN	484.3001	2	966.5856	966.5862	33.38	0.0067	EVIAGPLLR
1157.667	RS3_HUMAN	579.309	2	1156.6034	1156.5951	51.77	0.00014	IMLPWDPTGK
	SPTN1_HUMAN	386.5497	3	1156.6272	1156.6029	33	0.011	LAQFVEHWK
	CPNE1_HUMAN	579.2889	2	1156.5632	1156.5553	30.4	0.014	SDPFLEFFR
	PABP1_HUMAN	579.3409	2	1156.6673	1156.6604	29.37	0.017	FSPAGPILSIR
1198.667	ACTB_HUMAN	599.8602	2	1197.7058	1197.6982	77.62	1.70E-07	AVFPSIVGRPR
	ACTC_HUMAN	599.8602	2	1197.7058	1197.6982	77.62	1.70E-07	AVFPSIVGRPR
	ACTA_HUMAN	599.8602	2	1197.7058	1197.6982	77.62	1.70E-07	AVFPSIVGRPR
	RO60_HUMAN	599.8472	2	1197.6799	1197.6717	75.09	4.60E-07	LGLENAEALIR
	H14_HUMAN	599.8398	2	1197.665	1197.6605	62.3	9.50E-06	ASGPPVSELITK
802.417	no match	no match	no match	no match	no match	no match	no match	no match
975.417	EF1A1_HUMAN	488.2835	2	974.5524	974.5437	47.72	0.00028	LPLQDVYK
	ERF3A_HUMAN	488.2838	2	974.5531	974.58	46.91	0.00034	LPIVDKYK
	ATPB_HUMAN	488.2882	2	974.5618	974.5549	36.87	0.0029	IGLFGGAGVGK
	RL11_HUMAN	488.283	2	974.5515	974.5437	26.46	0.038	YDGIILPGK
1242.667	HS90B_HUMAN	621.8602	2	1241.7059	1241.6979	81.32	1.20E-07	ADLINNLGTIAK
	HS90A_HUMAN	621.8602	2	1241.7059	1241.6979	81.32	1.20E-07	ADLINNLGTIAK
	H90B2_HUMAN	621.8602	2	1241.7059	1241.6979	81.32	1.20E-07	ADLINNLGTIAK
	RL22_HUMAN	621.846	2	1241.6774	1241.6728	73.35	9.70E-07	AGNLGGGVVTIER
	GRP75_HUMAN	621.8458	2	1241.677	1241.6728	65.65	5.70E-06	DAGQISGLNVLR
	K2C7_HUMAN	621.8281	2	1241.6416	1241.6364	57.4	3.90E-05	GQLEALQVDGGR
	TERA_HUMAN	621.8246	2	1241.6346	1241.6252	51.76	0.00013	EVDIGIPDATGR
	2AAA_HUMAN	621.8785	2	1241.7424	1241.7343	50.48	6.40E-05	LSTIALALGVER
	SRP14_HUMAN	621.8211	2	1241.6277	1241.6227	50.45	0.00018	FQMAYSNLLR
	TLN1_HUMAN	414.8787	3	1241.6142	1241.604	28.44	0.027	FLPSELRDEH
1161.667	ACTB_HUMAN	581.3166	2	1160.6187	1160.6111	70.28	2.40E-06	EITALAPSTMK
	ACTC_HUMAN	581.3166	2	1160.6187	1160.6111	70.28	2.40E-06	EITALAPSTMK
	ACTA_HUMAN	581.3166	2	1160.6187	1160.6111	70.28	2.40E-06	EITALAPSTMK
	CUL5_HUMAN	581.3374	2	1160.6602	1160.6594	40.75	0.0026	TLWSLVAFPK
	RET4_HUMAN	581.28	2	1160.5455	1160.5325	35.46	0.0051	FSGTWYAMAK
1167.667	BCLF1_HUMAN	389.9103	3	1166.7092	1166.7023	66.54	8.00E-07	LLASTLVHSVK
	ASNS_HUMAN	584.3244	2	1166.6342	1166.6335	46.38	0.00078	ELYLFDVLR
	RS4X_HUMAN	389.8951	3	1166.6636	1166.656	44.25	0.00041	GNKPWISLPR
	HNRPM_HUMAN	584.299	2	1166.5834	1166.576	34.1	0.0086	NLPFDFTWK
	SBP1_HUMAN	584.3419	2	1166.6691	1166.6659	33.34	0.01	HEIVQTLSLK
1612.917	LV302_HUMAN	806.9034	2	1611.7923	1611.7853	127.92	3.40E-12	FSGSNSGNTATLTISR
	APOA1_HUMAN	806.9034	2	1611.7922	1611.7781	110.2	2.00E-10	LLDNWDSVTSTFSK
	NMT1_HUMAN	806.8977	2	1611.7808	1611.7603	74.14	7.70E-07	GFDVFNALDLMENK
	CLH1_HUMAN	806.9307	2	1611.8468	1611.8396	53.3	8.80E-05	ESYVETELIFALAK
1032.667	RS9_HUMAN	516.7995	2	1031.5845	1031.5764	59.12	2.50E-05	LFEGNALLR
	SYFB_HUMAN	516.8008	2	1031.5871	1031.5764	30.18	0.018	DLLFQALGR
	H31T_HUMAN	344.8695	3	1031.5868	1031.5876	26.64	0.04	YRPGTVALR
1027.667	HXK1_HUMAN	514.3199	2	1026.6253	1026.6186	44.56	0.00061	GAALITAVGVR
	PGM1_HUMAN	514.2806	2	1026.5466	1026.5386	39.81	0.0018	SIFDFSALK
	KAIN_HUMAN	514.2785	2	1026.5424	1026.5386	34	0.017	LGFTDLFSK
941.417	UBP2L_HUMAN	471.3083	2	940.6021	940.5957	59.13	1.30E-05	IDLAVLLGK
	CH60_HUMAN	314.5459	3	940.6159	940.6069	36.71	0.00051	IGIEIIKR
	TXD17_HUMAN	471.3096	2	940.6046	940.5957	35.48	0.0012	VTAVPTLLK
	ACOC_HUMAN	471.2506	2	940.4867	940.4807	32.82	0.019	LFFWNSK
	GSHB_HUMAN	471.2964	2	940.5782	940.5705	31.12	0.0055	ALAEGVLLR
	SCMC1_HUMAN	471.2448	2	940.475	940.4589	30.66	0.031	MNIFGGFR
	CPNE3_HUMAN	471.2799	2	940.5453	940.5342	25.92	0.03	SPLGEVAIR
976.417	In situ identified	In situ identified	In situ identified	In situ identified	In situ identified	In situ identified	In situ identified	AGFAGDDAPR
1406.667	HXK1_HUMAN	703.8682	2	1405.7218	1405.7123	100.28	2.20E-09	LSDETLIDIMTR
	ENOA_HUMAN	703.8667	2	1405.7188	1405.7089	80.57	2.00E-07	GNPTVEVDLFTSK
	SRPR_HUMAN	469.5788	3	1405.7145	1405.7023	71.03	1.90E-06	DAAGIAMEAIAFAR
	ENOG_HUMAN	703.8667	2	1405.7188	1405.7089	62.59	1.30E-05	GNPTVEVDLYTAK
	SAMP_HUMAN	703.8407	2	1405.6669	1405.6626	53.79	7.90E-05	AYSLFSYNTQGR
1138.667	CO6A3_HUMAN	569.8042	2	1137.5938	1137.5852	62.12	1.40E-05	SSIMAFAIGNK
	G6PD_HUMAN	380.2176	3	1137.6309	1137.6182	61.92	3.50E-05	LKLEDFFAR
	EF2_HUMAN	569.7671	2	1137.5197	1137.5091	32.49	0.0077	YEWDVAEAR

1115.417	AL1A1_HUMAN	372.564	3	1114.6701	1114.6597	63.25	3.10E-06	ILDLIESGKK
	PGS2_HUMAN	558.3289	2	1114.6432	1114.6387	62.77	7.50E-06	VSPGAFTPLVK
	TMED4_HUMAN	372.5551	3	1114.6435	1114.6346	52.94	7.20E-05	DKLTELQLR
	TOP2B_HUMAN	558.3203	2	1114.626	1114.6169	50.68	0.00038	DLIQMLVQR
	AK1A1_HUMAN	558.3176	2	1114.6207	1114.6209	50.34	0.0002	MPLIGLGTWK
	ALDR_HUMAN	558.3176	2	1114.6207	1114.6209	50.34	0.0002	MPILGLGTWK
	EIF3E_HUMAN	558.2876	2	1114.5606	1114.5369	43.46	0.0022	MLFDYLADK
	TENX_HUMAN	558.3062	2	1114.5979	1114.5983	30.26	0.041	SVTVTPLDAGR
944.417	H2A1B_HUMAN	472.7713	2	943.5279	943.524	60.32	2.50E-05	AGLQFPVGR
	H2A1D_HUMAN	472.7713	2	943.5279	943.524	60.32	2.50E-05	AGLQFPVGR
	H2A1A_HUMAN	472.7713	2	943.5279	943.524	60.32	2.50E-05	AGLQFPVGR
	RBMX_HUMAN	472.8001	2	943.5856	943.5776	50.79	4.50E-05	IVEVLLMK
	H2AV_HUMAN	472.7736	2	943.5326	943.524	47.17	0.00053	AGLQFPVGR
	WFS1_HUMAN	472.7905	2	943.5665	943.5702	26.76	0.022	TLTDLLLR
857.417	no match	no match	no match	no match	no match	no match	no match	no match
1905.917	HSPB1_HUMAN	953.5054	2	1904.9963	1904.9843	129.4	1.70E-12	LATQSNEITIPVTFESR
	IGHG2_HUMAN	635.9761	3	1904.9064	1904.8866	79.15	2.00E-07	TTPPMLDSDGSFFLYSK
	SBP1_HUMAN	953.4614	2	1904.9082	1904.9004	69.19	2.00E-06	NTGTEAPDYLATVDVDPK
915.417	H2AY_HUMAN	458.2781	2	914.5416	914.5338	61.18	1.80E-05	AGVIFPVGR
	H2AW_HUMAN	458.2781	2	914.5416	914.5338	61.18	1.80E-05	AGVIFPVGR
	ZO1_HUMAN	458.2778	2	914.541	914.5338	44.67	0.0011	LGSWLAIR
	COCA1_HUMAN	458.279	2	914.5435	914.5437	37.8	0.0054	IGVLITDGK

[1] *m/z* value in the middle of the 0.25 Da wide bins