IDENTIFICATION OF POTENTIAL BIOMARKERS OF ORAL SQUAMOUS CELL CARCINOMA USING INTEGRATED PROTEOMICS AND GLYCOPROTEOMICS ANALYSES

WONG YIN LING

FACULTY OF DENTISTRY
UNIVERSITY OF MALAYA
KUALA LUMPUR

2020

Universiti Malaya

IDENTIFICATION OF POTENTIAL BIOMARKERS OF ORAL SQUAMOUS CELL CARCINOMA USING INTEGRATED PROTEOMICS AND GLYCOPROTEOMICS ANALYSES

WONG YIN LING

THESIS SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

FACULTY OF DENTISTRY
UNIVERSITY OF MALAYA
KUALA LUMPUR

2020

Universiti Malaya

UNIVERSITY OF MALAYA
ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: Wong Yin Ling

Matric No: DHA140006

Name of Degree: Doctor of Philosophy

Title of Thesis: Identification of Potential Biomarkers of Oral Squamous Cell Carcinoma Using Integrated Proteomics and Glycoproteomics Analyses

Field of Study: Oral Oncology

I do solemnly and sincerely declare that:

(1) I am the sole author/writer of this Work;
(2) This Work is original;
(3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
(4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
(5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya ("UM"), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
(6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate's Signature Date:

Subscribed and solemnly declared before,

Witness's Signature Date:

Name:

Designation:

Universiti Malaya

IDENTIFICATION OF POTENTIAL BIOMARKERS OF ORAL SQUAMOUS CELL CARCINOMA USING INTEGRATED PROTEOMICS AND GLYCOPROTEOMICS ANALYSES

ABSTRACT

Oral cancer is one of the major health concerns worldwide. The 5-year survival rate of oral squamous cell carcinoma (OSCC) has remained at approximately 50% over the past decades. Most OSCCs are diagnosed at an advanced stage due to the delay in diagnosis, suggesting an imperative need in identifying reliable biomarkers to improve the diagnosis, prognosis, and treatment of OSCC. This study aimed to identify potential biomarkers that are involved in the development and progression of OSCC using integrated proteomics and glycoproteomics analyses. Proteomics analysis was performed on serum samples of 10 patients with oral potentially malignant disorder (OPMD), 40 patients with OSCC, and 10 healthy volunteers as control using two-dimensional gel electrophoresis (2-DE) followed by silver staining. Based on the analysis, 5 proteins (AAT, APOA1, IGKC, SAMP, and VDBP) were up-regulated and 5 (AMBP, CLU, HP, PRDX2, and RBP4) down-regulated in OPMD when compared with control ($p < 0.05$). In OSCC, 4 proteins (IGHA2, IGHG2, IGKC, and TF) were up-regulated and 5 (ALB, AMBP, CLU, HP, and LRG1) down-regulated in the early stage of OSCC, whereas 5 proteins (AAT, APOA1, C3, IGHG2, and VDBP) were up-regulated and one (PRDX2) down-regulated in the advanced stage of OSCC when compared with control ($p < 0.05$). As for glycoproteomics analysis, the serum samples were subjected to 2-DE coupled with Concanavalin A and Jacalin lectin for the detection of *N*- and *O*-glycosylated proteins, respectively. A total of 5 glycoproteins (AAT, AHSG, APOA1, CLU, and HP) that exhibited tumour-specific glycosylation changes in OPMD and OSCC were identified. Most of these identified proteins and glycoproteins were acute-phase proteins, which indicated the presence of chronic inflammation in OSCC. Furthermore, bioinformatics

Universiti Malaya

analysis revealed the involvement of the identified proteins in platelet degranulation, activation of classical complement pathway, LXR/RXR activation, and acute phase response signalling pathways during development and progression of OSCC. Based on these findings, AAT, AHSG, APOA1, CLU, and HP were selected for further validation using enzyme-linked immunosorbent assay (ELISA) and immunohistochemistry (IHC). In the ELISA analysis, AAT, AHSG, APOA1, CLU, and HP showed consistent findings with the proteomics and glycoproteomics analyses. However, based on the IHC results, only CLU and HP were found to be corroborated with the findings. These identified potential biomarkers may play important roles to improve the detection of OSCC. Nevertheless, further investigation is warranted to determine their roles in OSCC.

Keywords: Oral Squamous Cell Carcinoma, Biomarker, Proteomics, Glycoproteomics, Lectin

Universiti Malaya

PENGENALPASTIAN BIOPENANDA YANG BERPOTENSI BAGI KANSER MULUT JENIS KARSINOMA SEL SKUAMUS DENGAN MENGGUNAKAN ANALISIS PROTEOMIK DAN GLIKOPROTEOMIK BERSEPADU

ABSTRAK

Kanser mulut adalah salah satu masalah kesihatan yang dibimbangkan di seluruh dunia. Kadar kemadirian 5 tahun bagi kanser mulut jenis karsinoma sel skuamus (OSCC) masih kekal pada lebih kurang 50% sejak beberapa dekad yang lalu. Kebanyakan OSCC didiagnosis pada peringkat lanjut disebabkan kelewatan dalam diagnosis, ini mencadangkan pengenalpastian biopenanda yang andal adalah keperluan penting untuk meningkatkan diagnosis, prognosis, dan rawatan OSCC. Kajian ini bertujuan untuk mengenal pasti biopenanda berpotensi yang terlibat dalam pembangunan dan perkembangan OSCC dengan menggunakan analisis proteomik dan glikoproteomik bersepadu. Analisis proteomik dilakukan pada sampel serum 10 pesakit gangguan mulut berpotensi malignan (OPMD), 40 pesakit OSCC, dan 10 sukarelawan yang sihat sebagai kawalan dengan menggunakan gel elektroforesis 2-dimensi (2-DE) diikuti oleh pewarnaan perak. Berdasarkan analisis, 5 protein (AAT, APOA1, IGKC, SAMP, dan VDBP) beregulasi menaik dan 5 (AMBP, CLU, HP, PRDX2, dan RBP4) beregulasi menurun dalam OPMD apabila dibandingkan dengan kawalan ($p < 0.05$). Pada OSCC, 4 protein (IGHA2, IGHG2, IGKC, dan TF) beregulasi menaik dan 5 (ALB, AMBP, CLU, HP, dan LRG1) beregulasi menurun dalam peringkat awal OSCC, manakala 5 protein (AAT, APOA1, C3, IGHG2, dan VDBP) beregulasi menaik dan satu (PRDX2) beregulasi menurun dalam peringkat lanjutan OSCC apabila dibandingkan dengan kawalan ($p < 0.05$). Bagi analisis glikoproteomik, sampel serum tertakluk kepada 2-DE diikuti dengan lektin Concanavalin A dan Jacalin untuk mengesan protein berglikosilat-*N* dan -*O*. Sejumlah 5 glikoprotein (AAT, AHSG, APOA1, CLU, dan HP) telah menunjukkan perubahan glikosilasi secara tumor spesifik dalam OPMD dan OSCC. Kebanyakan

Universiti Malaya

protein dan glikoprotein yang dikenal pasti didapati adalah protein fasa-akut. Tambahan lagi, analisis bioinformatik menunjukkan protein-protein yang dikenal pasti ini terlibat dalam degranulasi platelet, pengaktifan laluan komplemen klasik, pengaktifan LXR/RXR, dan laluan pengisyaratan tindak balas fasa akut semasa pembangunan dan perkembangan OSCC. Berdasarkan penemuan, AAT, AHSG, APOA1, CLU, dan HP telah dipilih untuk pengesahan lanjut dengan menggunakan asai imunoserap terangkai ensim (ELISA) dan imunohistokimia (IHC). Dalam analisis ELISA, AAT, AHSG, APAO1, CLU, dan HP menunjukkan penemuan yang konsisten dengan analisis proteomik dan glikoproteomik. Manakala, keputusan IHC menunjukkan hanya CLU dan HP menyokong penemuan. Biopenanda-biopenanda yang dikenal pasti ini mungkin memainkan peranan penting untuk meningkatkan pengesanan OSCC. Walau bagaimanapun, penyelidikan lanjut diperlukan untuk menentukan peranan meraka dalam OSCC.

Kata kunci: Kanser Mulut Jenis Karsinoma Sel Skuamus, Biopenanda, Proteomik, Glikoproteomik, Lektin

Universiti Malaya

ACKNOWLEDGEMENTS

First and foremost, my deepest gratitude goes to my supervisors, Assoc. Prof. Dr. Chen Yeng, Prof. Dato' Dr. Zainal Ariff Bin Abdul Rahman, and Dr. Anand Ramanathan for their supervision, guidance, and support. Thank you for your patient and understanding throughout this journey to make this study a success.

I would like to convey my gratitude to the University of Malaya Fellowship Scheme (SBUM) for the financial support throughout my postgraduate study. I also would like to take this opportunity to thank all the lecturers and staff members from Department of Oral and Craniofacial Science, Oral Cancer Research and Coordinating Centre (OCRCC), Dental Research Management Unit (DRMU), Balai Ungku Aziz Research Laboratory (BUARL), Craniofacial & Molecular Biology Research Laboratory (CMBRL), Research & Diagnostic Laboratory (RDL), Oral Pathology Diagnostic and Research Laboratory (OPDRL), and Medical Biotechnology Lab (MBL) for their kind support and assistance. I am also grateful to Dr. Mohammad Zabri Johari for his advice related to statistical analysis.

Many thanks to my lab mates and friends for delightful company and support throughout the journey. Last but not the least, I would like to give my deepest appreciation to my parents and my family members for their patience, encouragement, and understanding.

Universiti Malaya

TABLE OF CONTENTS

IDENTIFICATION OF POTENTIAL BIOMARKERS OF ORAL SQUAMOUS CELL CARCINOMA USING INTEGRATED PROTEOMICS AND GLYCOPROTEOMICS ANALYSES Abstract iii

PENGENALPASTIAN BIOPENANDA YANG BERPOTENSI BAGI KANSER MULUT JENIS KARSINOMA SEL SKUAMUS DENGAN MENGGUNAKAN ANALISIS PROTEOMIK DAN GLIKOPROTEOMIK BERSEPADU Abstrak v

Acknowledgements vii

Table of Contents viii

List of Figures xii

List of Tables xiv

List of Symbols and Abbreviations xvi

List of Appendices xxiii

CHAPTER 1: INTRODUCTION 1

1.1 Introduction 1

1.2 Research Objectives 6

CHAPTER 2: LITERATURE REVIEW 7

2.1 Oral Cancer 7

2.1.1 Epidemiology 8

2.1.2 Aetiology 10

2.1.2.1 Chemical factors 10

2.1.2.2 Biological factors 12

2.1.3 Oral Potentially Malignant Disorder (OPMD) and Oral Cancer 15

2.1.4 Oral Cancer Staging and Grading 17

Universiti Malaya

2.1.5 Diagnosis, Prognosis, and Treatment 19
2.1.6 Oral Carcinogenesis 23
2.2 Cancer Biomarker Discovery 27
2.2.1 Proteomics in Biomarker Discovery 28
2.2.2 Glycoproteomics in Biomarker Discovery 30
2.3 Proteomics and Glycoproteomics Approaches to Identify Potential Biomarkers 34
2.3.1 Two-dimensional Gel Electrophoresis (2-DE) and Mass Spectrometry 34
2.3.2 Lectin-based Glycoproteomics Approaches 36

CHAPTER 3: METHODOLOGY 40
3.1 Sample Collection 40
3.2 Study Design 40
3.3 Two-dimensional Gel Electrophoresis (2-DE) 42
3.3.1 First-dimensional Separation 42
3.3.2 Second-dimensional Separation 43
3.3.3 Silver Staining 45
3.3.4 Image Analysis 46
3.4 Glycoproteomics Study 47
3.4.1 *N*-glycoprotein Study 47
3.4.2 *O*-glycoprotein Study 48
3.5 Mass Spectrometry (MS) Analysis 49
3.5.1 In-gel Trypsin Digestion 49
3.5.2 Sample Preparation for Mass Spectrometry 50
3.5.3 Protein Identification 51
3.6 Bioinformatics Analysis 52
3.7 Enzyme-linked Immunosorbent Assay (ELISA) 53
3.8 Immunohistochemistry (IHC) 55

Universiti Malaya

CHAPTER 4: RESULTS 58

4.1 Study Subjects 58

4.2 2-DE Gel Analysis 60

4.2.1 Image Analysis of 2-DE Gel 60

4.2.2 2-DE Protein Profile Analysis 63

4.3 Glycoproteomics Analysis 75

4.3.1 *N*-glycoprotein Analysis 75

4.3.1.1 Con A affinity chromatography 75

4.3.1.2 Image analysis of *N*-glycoprotein profile 76

4.3.1.3 *N*-glycoprotein profile analysis 78

4.3.2 *O*-glycoprotein Analysis 79

4.3.2.1 Image analysis of *O*-glycoprotein profile 79

4.3.2.2 *O*-glycoprotein profile analysis 81

4.4 Bioinformatics Analysis 82

4.5 Enzyme-linked Immunosorbent Assay (ELISA) Analysis 93

4.6 Immunohistochemistry (IHC) Analysis 102

CHAPTER 5: DISCUSSION 116

5.1 Analysis of Serum Protein and Glycoprotein Profiles Using Proteomics and Glycoproteomics Approaches 117

5.1.1 Analysis of Serum Protein Profiles Using Proteomics Approach 117

5.1.2 Analysis of Serum Glycoprotein Profiles Using Glycoproteomics Approach 122

5.2 Functional Annotation and Protein Interaction Analysis 131

5.3 Validation of Potential OSCC Biomarkers 138

CHAPTER 6: CONCLUSION 142

Universiti Malaya

References 144

List of Publications and Papers Presented 176

Appendix A: Ethical Approval 177

Appendix B: Mass Spectra of Tryptic Peptides 178

Appendix C: Prediction of Glycosylation Sites 189

Appendix D: Association of Socio-demographic and Clinicopathological Characteristics with AAT, AHSG, APOA1, CLU, and HP Expression in OSCC 199

Appendix E: Association of Socio-demographic and Clinicopathological Characteristics with AAT, AHSG, APOA1, CLU, and HP Expression in OSCC 202

Universiti Malaya

LIST OF FIGURES

Figure 2.1: Histological progression of oral carcinogenesis 23

Figure 3.1: Overview workflow of the study 41

Figure 4.1: Typical representative of serum protein profile using 2-DE and silver staining 61

Figure 4.2: Typical representative 2-DE serum protein profiles of (a) control, (b) OPMD, (c) early OSCC, and (d) advanced OSCC 64

Figure 4.3: Principal component analysis (PCA) plot of (a) OPMD, (b) early OSCC, and (c) advanced OSCC compared with control 71

Figure 4.4: Principal component analysis (PCA) plot of OPMD and OSCC 72

Figure 4.5: The fold change (relative protein abundance) level of identified proteins in OPMD, early OSCC, and advanced OSCC compared with control 74

Figure 4.6: Representative elution profiles of (a) control, (b) OPMD, (c) early OSCC, and (d) advanced OSCC 76

Figure 4.7: Typical representative serum *N*-glycoprotein profiles of (a) control, (b) OPMD, (c) early OSCC, and (d) advanced OSCC 77

Figure 4.8: Typical representative serum *O*-glycoprotein profiles of (a) control, (b) OPMD, (c) early OSCC, and (d) advanced OSCC 80

Figure 4.9: Gene ontology analysis of identified proteins 83

Figure 4.10: Canonical pathway analysis of identified proteins 87

Figure 4.11: Network analysis of identified proteins with molecule activity prediction 91

Figure 4.12: Mean expression level of (a) AAT, (b) AHSG, (c) APOA1, (d) CLU, and (e) HP in control, OPMD, early OSCC, and advanced OSCC 96

Figure 4.13: Photomicrograph of AAT expression in (a) positive control, (b) negative control, (c) normal oral mucosa, (d) OPMD, (e) early OSCC, and (f) advanced OSCC 105

Figure 4.14: Photomicrograph of AHSG expression in (a) positive control, (b) negative control, (c) normal oral mucosa, (d) OPMD, (e) early OSCC, and (f) advanced OSCC 106

Universiti Malaya

Figure 4.15: Photomicrograph of APOA1 expression in (a) positive control, (b) negative control, (c) normal oral mucosa, (d) OPMD, (e) early OSCC, and (f) advanced OSCC 107

Figure 4.16: Photomicrograph of CLU expression in (a) positive control, (b) negative control, (c) normal oral mucosa, (d) OPMD, (e) early OSCC, and (f) advanced OSCC 108

Figure 4.17: Photomicrograph of HP expression in (a) positive control, (b) negative control, (c) normal oral mucosa, (d) OPMD, (e) early OSCC, and (f) advanced OSCC 109

Figure 4.18: Kaplan-Meier overall survival curve for the expression of (a) AAT, (b) AHSG, (c) APOA1, (d) CLU, and (e) HP in OSCC 115

Universiti Malaya

LIST OF TABLES

Table 2.1: TNM staging of oral cancer 18

Table 3.1: Running condition for IEF on Multiphor II Electrophoresis System 43

Table 3.2: Preparation of 12.5% SDS-PAGE gel solution 44

Table 3.3: Running condition for SDS-PAGE 45

Table 3.4: Specific condition of ELISA kit 54

Table 3.5: Specific condition of primary antibodies 57

Table 4.1: Demographic characteristics of study samples 59

Table 4.2: MS identification of serum proteins 62

Table 4.3: Differentially expressed protein spots in OPMD compared with control 65

Table 4.4: Differentially expressed protein spots in early OSCC compared with control 66

Table 4.5: Differentially expressed protein spots in advanced OSCC compared with control 67

Table 4.6: Differentially expressed protein spots in OPMD, early OSCC, and advanced OSCC compared with control 67

Table 4.7: List of detected *N*-glycoproteins in OPMD, early OSCC, and advanced OSCC when compared with control 78

Table 4.8: List of detected *O*-glycoproteins in OPMD, early OSCC, and advanced OSCC when compared with control 81

Table 4.9: Functional annotation clustering analysis of identified proteins using DAVID v6.8 84

Table 4.10: List of the proteins in the activation/inhibition canonical pathways 88

Table 4.11: Top five significant molecular and cellular functions of identified proteins 88

Table 4.12: Associated network functions analysis 89

Table 4.13: Predicted glycosylation sites of the identified glycoproteins 92

Universiti Malaya

Table 4.14: Demographic characteristics of study samples in ELISA analysis..............94

Table 4.15: The expression level of AAT, AHSG, APOA1, CLU, and HP in control, OPMD, early OSCC, and advanced OSCC 95

Table 4.16: Optimal cut-off value and area under the curve (AUC) in ROC analysis ...97

Table 4.17: Association analysis between the expression level of AAT, AHSG, and APOA1 with the study groups 99

Table 4.18: Association analysis between the expression level of CLU and HP with the study groups 100

Table 4.19: Demographic characteristics of study samples in IHC analysis 103

Table 4.20: Optimal cut-off value and area under the curve (AUC) in ROC analysis .110

Table 4.21: Association analysis between the expression level of AAT, AHSG, and APOA1 with the study groups 112

Table 4.22: Association analysis between the expression level of CLU and HP with the study groups 113

Universiti Malaya

LIST OF SYMBOLS AND ABBREVIATIONS

2-DE	:	Two-dimensional gel electrophoresis
4-PL	:	Four-parameter logistic
AAT	:	α1-antitrypsin
AFP	:	alpha-fetoprotein
AHSG	:	α2-HS-glycoprotein
AINX	:	Alpha-internexin
AJCC	:	American Joint Committee on Cancer Staging
ALB	:	Albumin
AMBP	:	Alpha-1-microglobulin/bikunin precursor
ANOVA	:	Analysis of variance
APOA1	:	Apolipoprotein A1
APS	:	Ammonium persulfate
ARF	:	Tumour suppressor alternative reading frame
Asn	:	Asparagine
ASR	:	Age standardize rate
ATM	:	Ataxia telangiectasia mutated
AUC	:	Area under the curve
B7-H3	:	B7 homolog H3
BCL2	:	B cell lymphoma 2
C1GalT1	:	Core 1 β1,3-galactosyltransferase C1GalT1
C3	:	Complement C3
$CaCl_2$	:	Calcium chloride
CA125	:	Carcinoma antigen 125
CA15-3	:	Carcinoma antigen 15-3

Universiti Malaya

CA19-9	:	Carcinoma antigen 19-9
CCND1	:	Cyclin D1
CCR2	:	Chemokine receptor 2
CD44	:	Cluster of differentiation 44
CD147	:	Cluster of differentiation 147
CD163	:	Cluster of differentiation 163
CDC25	:	Cell division cycle 25
CDC25B	:	Cell division cycle 25B
CDDP	:	Cis-dichloro-diamine-platinum or cisplatin
CDKN2A	:	Cyclin-dependent kinase inhibitor 2A
CEA	:	Carcinoembryonic antigen
CHCA	:	α-cyano-4-hydroxy-cinnamic acid
CHEK1	:	Checkpoint kinase 1
CLU	:	Clusterin
CMV	:	Cytomegalovirus
Con A	:	Concanavalin A
CT	:	Computed tomography
CXCL12	:	Chemokine (C-X-C motif) ligand 2
CXCR4	:	C-X-C chemokine receptor type 4
DAB	:	3,3'-diaminobenzidine
DAVID	:	Database for Annotation, Visualization, and Integrated Discovery
DNA	:	Deoxyribonucleic acid
DPAGT1	:	Dolichyl-phosphate *N*-acetylglucosamine phosphotransferase 1
DTT	:	Dithiothreitol
EBV	:	Epstein-Bar virus
EGF	:	Epidermal growth factor

Universiti Malaya

EGFR	:	Epidermal growth factor receptor
ELISA	:	Enzyme-linked immunosorbent assay
ER	:	Endoplasmic reticulum
ERK1/2	:	Extracellular signal-regulated kinase 1 and 2
ESI	:	Electrospray ionization
FC	:	Fold change
FDA	:	Food and Drug Administration
FFPE	:	Formalin-fixed paraffin-embedded
FTICR	:	Fourier transform ion cyclotron resonance
FXR/RXR	:	Famesoid X receptor/retinoid X receptor
Gy	:	Gray
GalNAc	:	*N*-acetylgalactosamine
Glc	:	Glucose
GlcNAc	:	*N*-acetylglucosamine
GALNT2	:	Polypeptide *N*-acetylgalactosaminyltransferase 2
GLOBOCAN	:	Global Burden of Cancer
GO	:	Gene ontology
HDL	:	High-density lipoprotein
HP	:	Haptoglobin
HPV	:	Human papillomavirus
HNSCC	:	Head and neck squamous cell carcinoma
HRP	:	Horseradish peroxidase
HSV-1	:	Herpes simplex virus type 1
JAK	:	Janus kinase
IgG	:	Immunoglobulin G
IgG1	:	Immunoglobulin G1

Universiti Malaya

IgG4	:	Immunoglobulin G4
IAA	:	Iodoacetamide
IARC	:	International Agency for Research on Cancer
ICD-10	:	International Statistical Classification of Diseases 10th Revision
IEF	:	Isoelectric focusing
IGF-1		Insulin-like growth factor-1
IGFBP-2		Insulin-like growth factor binding protein-2
IGHA2	:	Immunoglobulin alpha-2 chain region
IGHG2	:	Immunoglobulin gamma-2 chain C region
IGKC	:	Immunoglobulin kappa chain C region
IHC	:	Immunohistochemistry
IL-6	:	Interleukin 6
IPA	:	Ingenuity Pathway Analysis
LRG1	:	Leucine-rich alpha-2-glycoprotein
LXR/RXR	:	Liver X receptor/retinoid X receptor
mTOR	:	Mechanistic target of rapamycin
m/z	:	Mass-to-charge
MALDI	:	Matrix-assisted laser desorption ionization
Man	:	Mannose
MAPK	:	Mitogen-activated protein kinase
$MnCl_2$	:	Manganese (II) chloride
MOCDTBS	:	Malaysian Oral Cancer Database and Tissue Bank System
MRI	:	Magnetic resonance imaging
MRM	:	Multiple reaction monitoring
MS	:	Mass spectrometry
MUCA5C	:	Mucin 5AC

Universiti Malaya

NaCl	:	Sodium chloride
NF-kB	:	Kappa-light-chain-enhancer of activated B cells
NGS	:	Next-generation sequencing
NIH	:	National Institute of Health
NO	:	Nitric oxide
OCRCC	:	Oral Cancer Research & Coordinating Centre
OPMD	:	Oral potentially malignant disorder
OSCC	:	Oral squamous cell carcinoma
OR	:	Odds ratio
pI	:	Isoelectric point
PCA	:	Principal component analysis
PD-1	:	Programmed cell death protein-1
PET	:	Positron emission tomography
PI3K	:	Phosphoinositide 3-kinases
PRDX2	:	Peroxiredoxin-2
PSA	:	Prostate-specific antigen
PTM	:	Post-translational modifications
PVDF	:	Polyvinylidene fluoride
PXDN	:	Peroxidasin
ras	:	Rat sarcoma viral
Rb	:	Retinoblastoma
RAD17	:	Cell cycle checkpoint protein RAD17
RAD50	:	DNA repair protein RAD50
RBP4	:	Retinol-binding protein 4
RNA	:	Ribonucleic acid
ROC	:	Receiver operator characteristic

Universiti Malaya

ROS	:	Reactive oxygen species
RPPA	:	Reverse-phase protein array
RT	:	Room temperature
sLe^{x}	:	sialyl Lewis X
S/N	:	Signal-to-noise
SAMP	:	Serum amyloid P-component
SCC	:	Squamous cell carcinoma
SDS-PAGE	:	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SEM	:	Standard error mean
Ser	:	Serine
SERPINA1	:	Serine proteinase inhibitor A1
SNA	:	*Sambucus nigra* agglutinin
SPSS	:	Statistical Package for the Social Sciences
SRM	:	Selected reaction monitoring
STAT	:	Signal transducers and activators of transcription
STAT3	:	Signal transducers and activators of transcription 3
Thr	:	Threonine
TBST	:	Tris-buffered saline with Tween 20
TEMED	:	N,N,N',N'-tetramethylrthylenediamine
TF	:	Serotransferrin
TGFα	:	Transforming growth factor-α
TMB	:	3,3′,5,5′-tetramethylbenzidine
TNF	:	Tumour necrosis factor
TNF-α	:	Tumour necrosis factor-α
TNM	:	Tumour-node-metastasis
TOF	:	Time-of-flight

Universiti Malaya

TP53	:	Tumour protein p53
Tris-HCl	:	Tris (hydroxymethyl) aminomethane (THAM) hydrochloride
UICC	:	Union of International Cancer Control
UniProtKB	:	Universal Protein Resource Knowledgebase
UM	:	University of Malaya
US	:	United States
VCL	:	Vinculin
VDBP	:	Vitamin D-binding protein
VEGF	:	Vascular endodermal growth factor
WGA	:	Wheat germ agglutinin
WHO	:	World Health Organization

Universiti Malaya

LIST OF APPENDICES

Appendix A: Ethical Approval... 177

Appendix B: Mass Spectra of Tryptic Peptides 178

Appendix C: Prediction of Glycosylation Sites 189

Appendix D: Association of Socio-demographic and Clinicopathological Characteristics with AAT, AHSG, APOA1, CLU, and HP Expression in OSCC 199

Appendix E: Association of Socio-demographic and Clinicopathological Characteristics with AAT, AHSG, APOA1, CLU, and HP Expression in OSCC……………………...202

Universiti Malaya

CHAPTER 1: INTRODUCTION

1.1 Introduction

Oral cancer is the largest subset of head and neck cancer. The anatomical sub-sites of oral cancer include the lip, tongue, floor of the mouth, buccal mucosa, gum, retromolar trigone, and palate. It is the 18th most common cancer worldwide with an estimated 354,864 new cases of oral cancer (2.1%) and the 16th most common cause of cancer-related mortality with an estimated 177,384 deaths (2.0%) (Ferlay et al., 2018). The incidence rate of oral cancer was reported high in Asia region, especially in South Asia (such as India and Sri Lanka) and Papua New Guinea (Bray et al., 2018; Miranda-Filho & Bray, 2020).

According to the Global Burden of Cancer (GLOBOCAN), oral cancer is the 19th most common cancer in Malaysia with an estimated 667 new cases (1.6%) and 327 deaths (1.3%) (Ferlay et al., 2018). The incidence of oral cancer differs by gender and ethnic groups, in which it is predominant among the Indian community, while tongue and buccal mucosa were the most common cancer sites among both males and females (Ghani, Razak, et al., 2019; Omar & Tamin, 2011). Moreover, the highest incidence of oral cancer was observed among Indian females with an age-standardized rate (ASR) of 4.9 per 100,000 female population based on the Malaysian National Cancer Registry Report (Azizah et al., 2019).

Oral squamous cell carcinoma (OSCC) constitutes more than 90.0% of oral cancers, while the remaining 10.0% accounts for the rare malignancies (unusual forms of SCC, minor salivary gland tumour, lymphoma, sarcoma, and melanoma) (Genden et al., 2010; Montero & Patel, 2015). Although the oral cavity is accessible for clinical examination, most OSCCs are diagnosed with signs and symptoms at the advanced stage of the disease (Warnakulasuriya, 2009). The practice of tobacco smoking, alcohol drinking, and betel

Universiti Malaya

quid chewing are important risk factors in the development of OSCC (Montero & Patel, 2015; Omar, 2013; Warnakulasuriya, 2009). Furthermore, oral potentially malignant disorder (OPMD) that includes leukoplakia, erythroplakia, oral lichen planus, and oral submucous fibrosis, has also been associated with a risk of malignant transformation to OSCC (Warnakulasuriya, Johnson, & van der Waal, 2007). Interestingly, OPMD and OSCC have similar risk factors (Johnson, Jayasekara, & Amarasinghe, 2011).

The incidence of OSCC increases with age and generally occurs in the age group of more than 45 years old. However, the incidence of OSCC in younger patients has also increased in recent years (Hussein et al., 2017; Warnakulasuriya, 2009). Despite rapid advancements in the diagnosis and treatment strategies (including surgery and chemotherapy), the prognosis of OSCC has not changed over the past decades. The 5-year survival rates for OSCC patients remain at approximately 50.0% (Montero & Patel, 2015; Omar, 2013; Warnakulasuriya, 2009). Although the 5-year survival rate of early stage OSCC is about 80.0%, it reduces to around 20.0% in the advanced stage (Elimairi, Sami, & Yousef, 2017; Omar, 2013). This could be due to the delay in diagnosis and the presence of recurrence or regional lymph node metastasis. The key challenge to decrease the mortality and morbidity of OSCC is to develop strategies for early detection and screening, which allow effective intervention and therapy. Consequently, the identification of biomarkers may potentially facilitate in the early diagnosis, prognosis, or monitoring status of OSCC.

Blood (serum and plasma) and tumour tissue samples are commonly used biological samples in the discovery of cancer biomarkers. Blood samples are an attractive source for the discovery of biomarkers as they are easily accessible for sampling with minimally invasive procedure (Pavlou, Diamandis, & Blasutig, 2013). Whereas the tumour tissue samples has the capability to link with the clinical outcomes, including progression,

Universiti Malaya

recurrence, and survival in the pathological condition (Peng et al., 2018). Moreover, proteins secreted, shed, or released by tissues could circulate into the bloodstream (Anderson & Anderson, 2002; Zhang et al., 2007). Thus, blood and tumour tissue samples can be used as an indicator to determine the physiological and pathological status of the disease.

Analysing comprehensive protein profiles using the proteomics approach have been applied in the effort to discover cancer biomarkers (Kaskas et al., 2014; Lim et al., 2016; Lin, Xu, et al., 2017; Nedjadi et al., 2020; Peng et al., 2019). Proteomics provides an excellent platform for biomarker discovery in various cancers, including breast, colorectal, and lung cancers (Da Costa et al., 2015; Gan, Chen, & Li, 2014; Wang, Lu, Zhang, Li, & Li, 2016). The two-dimensional gel electrophoresis (2-DE) and mass spectrometry analyses, which involve protein separation and identification are still widely applied in proteomics studies for cancer biomarker discovery. In the past few years, several proteomics studies have been conducted to explore the molecular events in the development of OSCC and the identification of potential biomarkers (Chen et al., 2014; Tung et al., 2013; Turhani, Krapfenbauer, Thurnher, Langen, & Fountoulakis, 2006; Yu, Chang, et al., 2011). These studies have tremendously enhanced the molecular understanding of OSCC.

Most of the proteins secreted from the cells or the extracellular surface of cells are glycoproteins. Aberrant glycosylation of glycoproteins is often associated with malignant transformation (Drake et al., 2010; Kailemia, Park, & Lebrilla, 2017; Munkley & Elliott, 2016; Pan, Brentnall, & Chen, 2016). Glycosylation is one of the most common post-translational modification processes in the proteins, of which more than 50.0% of these proteins are glycosylated (Apweiler, Hermjakob, & Sharon, 1999; Yang, Franc, & Heck, 2017). *N*- and *O*-linked glycosylation are the two main types of glycosylation (Roth,

Universiti Malaya

Yehezkel, & Khalaila, 2012; Yang et al., 2017). Aberrant expression of *N*-and *O*-linked glycoproteins could impact the regulation of numerous cellular regulation processes, including cell proliferation, invasion, and cellular interaction with their surrounding environment during malignant transformation and cancer progression (Pan et al., 2016; Stowell, Ju, & Cummings, 2015). Moreover, most known cancer biomarkers are glycoproteins, including carcinoma antigen 125 (CA 125), CA 15-3, CA 19-9, carcinoembryonic antigen (CEA), and prostate-specific antigen (PSA) (Kailemia et al., 2017). Thus, glycoproteins have emerged as one of the promising targets in the detection of cancer biomarkers.

As a subfield of proteomics, glycoproteomics particularly focuses on analysing the entire complement of glycoproteins in various biological processes. Lectins are often used in the glycoproteomics study as lectins are carbohydrate-binding proteins, which can specifically bind to carbohydrate epitopes of glycoproteins (Sharon & Lis, 2004). Recently, several glycoproteomics studies have focused on identifying potential biomarkers in different types of cancer, including ovarian, liver, and breast cancers by using lectins to target the glycoproteins (Abbott et al., 2010; Liu et al., 2017; Semaan, Wang, Marshall, & Sang, 2012; Tan et al., 2015). These studies have identified several glycoproteins that displayed glycosylation changes in the progression and development of cancer.

The glycosylation changes that cause an increase in glycoprotein levels have been observed in OSCC as well (Manoharan, Padmanabhan, Kolanjiappan, Ramachandran, & Suresh, 2004). Several previous glycoproteomics studies have shown that alteration of glycoproteins is related to the glycosylation changes in OSCC (Chang, Lin, et al., 2019; Chen et al., 2015; Chen, Chong, et al., 2013; Lin, Huang, Liu, Yang, & Huang, 2014). These findings revealed the association of glycoprotein alteration with cellular invasion

Universiti Malaya

and proliferation in OSCC progression. Thus, glycoproteomics study is essential in the discovery of potential biomarkers in OSCC.

To date, there is still a need for improvement in the diagnosis, prognosis, and treatment of OSCC to increase the survival rates of the OSCC patients. Additionally, the molecular complexity in the development and progression of OSCC is still not fully understood. Therefore, integrated proteomics and lectin-based glycoproteomics approaches were applied in this study to investigate the protein and glycoprotein profiles associated with OSCC. The protein and glycoprotein profiles of OSCC were compared with the profiles of OPMD and control to identify differentially expressed proteins and glycoproteins in the development and progression of OSCC. This integrated strategy may contribute important insights into the molecular understanding of the oral carcinogenesis. The identified proteins could be used as potential biomarkers of OSCC in screening, detection, and monitoring disease progression. This may influence the treatment strategies and management of OSCC patients.

Universiti Malaya

1.2 Research Objectives

The aim of this study is to identify potential biomarkers that are involved in the development and progression of OSCC using integrated proteomics and glycoproteomics analyses. This study is expected to enhance the current understanding of OSCC.

The specific objectives of the study are:

(a) To develop serum protein profiles of OPMD, OSCC, and normal controls using two-dimensional gel electrophoresis (2-DE).

(b) To obtain serum *N*- and *O*-glycoprotein profiles of OPMD, OSCC, and normal controls using 2-DE coupled with lectin-based approaches.

(c) To identify differentially expressed proteins and glycoproteins in OPMD and OSCC as compared with normal controls.

(d) To determine the biological functions and pathway annotations of differentially expressed proteins and glycoproteins that may be involved in the development and progression of OSCC.

(e) To validate the differentially expressed proteins and glycoproteins using enzyme linked immunosorbent assay (ELISA) and immunohistochemistry (IHC).

Universiti Malaya

CHAPTER 2: LITERATURE REVIEW

2.1 Oral Cancer

Cancer consists of a large group of diseases characterized by an uncontrollable growth of abnormal cells in any part of the body with potential to invade and spread to other parts of the body. It is one of the leading causes of death worldwide, accounting for an estimated 9,555,027 deaths and 18,078,957 new cases (Ferlay et al., 2018). Head and neck cancer represent the sixth most common cancer worldwide with approximately 888,000 new cases and 45,300 deaths annually (Ferlay et al., 2019; Ferlay et al., 2018). The incidence of head and neck cancer will increase by 30.0%, with 1,080,000 new cases by 2030 (Ferlay et al., 2019; Ferlay et al., 2018).

Oral cancer is a subset of head and neck cancers, and it is also known as mouth cancer. It is delineated as the uncontrollable growth of cells that invades and causes damage to the surrounding tissue in the oral cavity. Moreover, it has a potential to spread to the lymph nodes and distant organs. The World Health Organization (WHO) International Statistical Classification of Diseases 10th Revision (ICD-10) version for 2010 (https://icd.who.int/browse10/2010/en#/C00-C97) has defined oral cancer as the malignancy emerging from the anatomical subsites that correspond to the rubrics C00-C06. These subsites are comprised of lip (C00), base of tongue (C01), other and unspecified parts of tongue (C02), gum (C03), floor of the mouth (C04), palate (C05), and other and unspecified parts of the mouth (C06).

More than 90.0% of oral cancers are diagnosed as oral squamous cell carcinoma (OSCC), which originates from abnormal cells on the surface layer of the lip or mouth lining (Genden et al., 2010; Montero & Patel, 2015). The remaining subtypes of oral cancer represent rare or uncommon forms of squamous cell carcinoma. About 2.0 to 12.0% of all diagnosed oral cancers are verrucous carcinoma, a slow-growing type of

Universiti Malaya

squamous cell carcinoma (Rekha & Angadi, 2010). Minor salivary gland cancer is also one of the uncommon malignant oral cancers and it only accounts for less than 5.0% of all oral cancers (Montero & Patel, 2015).

2.1.1 Epidemiology

Oral cancer is the 18th most common cancer worldwide with an estimated 354,864 new cases (2.1%) and the 16th most common cause of cancer-related mortality with an estimated 177,384 deaths (2.0%) (Ferlay et al., 2018). A total of 274,000 oral cancer cases were reported globally in 2002 (Parkin, Bray, Ferlay, & Pisani, 2005). Noticeably, the oral cancer incidence rate has increased in recent decades. The incidence rates of oral cancer are varied geographically. The incidence rate of oral cancer is high in Pacific Islands (Melanesia and Papua New Guinea), South Asia (such as India and Sri Lanka), and Australia/New Zealand, whereas the incidence rate is low in Western Africa, Central America, and Northern Africa (Bray et al., 2018; Miranda-Filho & Bray, 2020).

More than half of oral cancers are reported in the Asia region. It ranked the 11th most common cancer with an estimated 227,906 new cases (2.6%) and an estimated 129,939 deaths (2.4%) (Ferlay et al., 2018). Of these, Southeast Asia region accounts for an estimated 16,818 new cases (7.4%) and an estimated 8,542 deaths (6.6%) (Ferlay et al., 2018). In Malaysia, oral cancer is the 19th most common cancer with an estimated 667 new cases (1.6%) and 327 deaths (1.3%) (Ferlay et al., 2018). Furthermore, the oral cancer cases were increased from 353 cases in 2007 to 1,975 cases in 2016 based on the Malaysian National Cancer Registry Report (Azizah et al., 2019; Omar & Tamin, 2011).

Globally, oral cancer is more common in males than in females. The incidence and mortality rate with age standardize rates (ASR) in males are 5.8 and 2.8 per 100,000 population, respectively, while the incidence and mortality rates with ASR in females are 2.3 and 1.2 per 100,000 population, respectively (Bray et al., 2018). Nevertheless, the

Universiti Malaya

trend appears to be shifting in some populations, where the incidence rates of oral cancer have been decreasing in males and increasing among females (Miranda-Filho & Bray, 2020). In Malaysia, oral cancer is more common in females. According to the Malaysian Cancer Registry Report, the incidence of oral cancer is highest in Indian females with ASR of 4.9 per 100,000 female population (Azizah et al., 2019). The trend is similar to Brunei, Thailand, and Laos that also showed a higher incidence rate in females than in males, in which the incidence rates with ASR are 9.0, 6.2, and 4.2 per 100,000 female population, respectively (Cheong et al., 2017; Rao, Mejia, Roberts-Thomson, & Logan, 2013).

The 5-year survival rate for oral cancer is still at approximately 50.0% over the past decades (Montero & Patel, 2015; Omar, 2013; Warnakulasuriya, 2009). However, the 5-year survival rate of oral cancer is much lower in India and Thailand, in which the range is 38.0 to 42.0% and 18.2 to 43.1%, respectively (Rao et al., 2013). In Malaysia, approximately 63.1% of oral cancer patients are diagnosed at an advanced stage (Azizah et al., 2019), in which the 5-year survival rate is 40.9% (Ghani, Ramanathan, et al., 2019). Additionally, the 5-year survival rate for the early stage of oral cancer is almost 80.0%, and it drops to 20.0% at the advanced stage (Elimairi et al., 2017; Omar, 2013).

The risk of developing oral cancer increases with age. It generally occurs in older age group, where most of the cases occur in people aged 50 years old or over (Warnakulasuriya, 2009). Recently, an increasing trend of incidence in younger patients has been reported worldwide (Hussein et al., 2017; Warnakulasuriya, 2009). The incidence of oral cancer in younger patients is about 0.4 to 3.6%, and it is more common in males than in females (Schmitd, Tjioe, Assao, & Oliveira, 2015). Among most oral cancer cases, younger patients have a better overall survival rate than older patients (Lee et al., 2020; Schmitd et al., 2015; Warnakulasuriya, 2009). Nevertheless, recent studies

Universiti Malaya

have not found significant differences between younger and older patients in the biological characteristics and molecular basis of oral cancer (Costa et al., 2018; Kapila, Natarajan, & Boaz, 2017).

The most common site of oral cancer among European and US populations is tongue, which account for 40.0 to 50.0% of oral cancers (Warnakulasuriya, 2009). Moreover, it is the most common site for oral cancer in Asian patients (Rao et al., 2013). Tongue is also the most common site for oral cancer in younger patients with 39.0 to 77.0% of the cases reported (Schmitd et al., 2015). The Asian population also tends to have oral cancer from buccal mucosa and gingiva site due to betel quid and/or tobacco chewing habits (Rao et al., 2013; Warnakulasuriya, 2009). Similarly, tongue and buccal mucosa are the most common sites for oral cancer in Malaysia, of which the Indian community were commonly diagnosed oral cancer at buccal mucosa site due to the high prevalence of betel quid chewing in this ethnic group (Ghani, Razak, et al., 2019).

2.1.2 Aetiology

The aetiology of oral cancer is multifactorial, in which the aetiological factors can act separately or synergistically. Chemical factors such as tobacco, alcohol, and betel quid as well as biological factors including human papillomavirus (HPV), herpes simplex virus type 1 (HSV-1), candidiasis, orodental factors, family history, and dietary pattern have been shown to associate with the risk of developing oral cancer.

2.1.2.1 Chemical factors

It is known that the practice of tobacco smoking, alcohol drinking, and betel quid chewing habits are important risk factors of oral cancer (Montero & Patel, 2015; Omar, 2013; Warnakulasuriya, 2009). Tobacco contains diverse carcinogenic materials such as polycyclic hydrocarbons and nitrosamines, which can cause malignant transformation. Moreover, tobacco (smoking and smokeless) and second-hand smoking have been

Universiti Malaya

classified as group I carcinogens for oral cancer (IARC, 2004a). It has been reported that tobacco can stimulate epigenetic alteration in oral epithelium, impair immune functions, and cause DNA (deoxyribonucleic acid) damage to induce oral cancer (Jiang, Wu, Wang, & Huang, 2019). The practice of tobacco smoking and the use of smokeless tobacco are associated with about 75.0% of oral cancer cases (Elimairi et al., 2017). In fact, the risk of oral cancer for smokers is six times higher than for non-smokers (Markopoulos, 2012).

Alcohol may contain carcinogens including nitrosamines and urethane contaminants. Excessive alcohol drinking can change the mucosal morphology with a reduction in epithelial thickness by increasing the mucosa permeability to toxins. It could also promote oral cancer by affecting the systemic and redox metabolisms that cause disturbance on retinoid, zinc, iron, and methyl groups as well as oxidative damage to DNA, proteins, and lipids (Liu, Chen, Sun, & Chen, 2015). The risk of oral cancer is six times higher in alcohol drinkers than in non-drinkers (Markopoulos, 2012). Furthermore, the practice of tobacco smoking and alcohol drinking has a potent synergistic effect in the aetiology of oral cancer. It can also induce many epigenetic modifications such as promote methylation level, alter methylation pattern, and increase histone modification during carcinogenesis (Ghantous, Schussel, & Brait, 2018). Therefore, the risk of oral cancer for the combination of smokers and alcohol drinkers is 30 times higher than in non-smokers and non-drinkers (Elimairi et al., 2017).

Betel quid (*pan* or *paan*) usually contains betel vine leaf, betel (areca nut), slaked lime, and tobacco with chemical substances such as nitrosamines, alkaloids, and polyphenols. It has been recognized as a group I carcinogen in humans by the International Agency for Research on Cancer (IARC) (IARC, 2004b). Betel quid chewing can cause cytogenetic changes in oral epithelium to initiate the process of carcinogenesis (Omar, 2013). Recently, betel quid derived nitrosamines were found to be directly absorbed into the oral

Universiti Malaya

tissues (Elimairi et al., 2017). Additionally, the practice of betel quid chewing is widely prevalent in the Asian population, especially in Indian and Taiwanese (IARC, 2004b; Markopoulos, 2012; Omar, 2013). In Malaysia, this habit is more commonly practiced in the Indian community and indigenous people of Sabah and Sarawak, especially in females (Ghani et al., 2011). Moreover, the risk of oral cancer is seven times higher for chewers of betel quid with tobacco and three times higher for chewers of betel quid without tobacco than in non-chewers (Gupta & Johnson, 2014). While the synergistic effect of smoking, alcohol drinking, and betel quid chewing increase the risk of developing oral cancer by more than 40 times (Lin, Jiang, Wu, Chen, & Liu, 2011; Petti, Masood, & Scully, 2013).

2.1.2.2 Biological factors

As a small non-enveloped DNA virus, human papillomavirus (HPV) infects cells in the mucosal and squamous epithelium. The HPV infection is one of the causes for oropharynx SCC (tonsils, oropharynx, and base of tongue) with a prevalence rate of 18.0 to 72.0% (Elimairi et al., 2017). The E6 and E7 viral oncoproteins encoded by HPV have been shown to inactivate tumour protein p53 (TP53) and retinoblastoma (Rb) tumour suppressor genes, respectively, thereby interfering with cell cycle regulation and keratinocyte differentiation as well as causing DNA damage and apoptosis (Elimairi et al., 2017). The HPV strains 16 and 18 are the most common high-risk types that present in head and neck cancers (Elimairi et al., 2017; Schmitd et al., 2015). Recently, the epidemiological trend suggested that HPV-related oral cancers are more expected to occur in younger patients due to the changes in socio-economic behaviour (Hussein et al., 2017; Schmitd et al., 2015). However, the prevalence of HPV infection in oral cancer is varied from 0 to 100.0% (Schmitd et al., 2015).

Universiti Malaya

On the other hand, the presence of herpes simplex virus type 1 (HSV-1), Epstein-Bar virus (EBV), and cytomegalovirus (CMV) may lead to oral cancer (Inchingolo et al., 2020). HSV-1, EBV, and CMV are enveloped, double-stranded DNA viruses from the herpesviridae family that usually affect the skin and mucosal sites. HSV-1 usually causes sores or blisters in the oral cavity and has been associated with oral cancer as it could promote IgG antibody response (Jain, 2016). Studies also showed that the synergistic effect of tobacco smoking and HSV-1 in the development of oral cancer and the prevalence rate of HSV-1 was 15.0% (Jain, 2016; Sand & Jalouli, 2014).

EBV is a well-known oncogenic virus in Burkitt's lymphoma and nasopharyngeal carcinoma. However, it has also been associated with oral cancer, in which the prevalence of EBV infection ranges from 0 to 100% (Guidry, Birdwell, & Scott, 2018). Dysplastic changes in oral epithelial cells may make the cells more susceptible to EBV infection and that infection may alter the phenotype and behaviour of the infected cell (Elimairi et al., 2017; Sand & Jalouli, 2014). Recent studies have indicated that the co-infections of EBV and CMV may contribute to the development of oral cancer (Jassim, Mahmood, & Musa, 2020; Naqvi et al., 2020). CMV is a common opportunistic pathogen that infects majority of adults worldwide. However, the prevalence of CMV in oral cancer is low and the virus probably acts as an infection cofactor or has a minor role in oral cancer (Naqvi et al., 2020; Saravani et al., 2015).

The presence of candidiasis (Candida infection) has been suggested as a potential risk factor for oral cancer. It is known that most oral fungal infections are caused by *Candida albicans*. These *Candida* species were able to produce an inflammatory environment and dysregulate the innate immune system that leads to tumour cell proliferation (Elimairi et al., 2017). Recent studies have revealed that chronic infection by *Candida albicans* that

Universiti Malaya

exists in host tissue is associated with the malignant transformation of oral leukoplakia and oral carcinogenesis (Satiman et al., 2020; Shukla et al., 2019).

There is a relationship between oral hygiene and the risk of developing oral cancer. The presence of poor oral hygiene and periodontitis may induce inflammatory bacteriome in oral cancer tissues (Perera et al., 2018). It has been shown that poor oral hygiene is strongly related to oral cancer, and it can promote the carcinogenic potential of tobacco and alcohol (Mathur, Singhavi, Malik, Nair, & Chaturvedi, 2019). Additionally, poor oral hygiene may induce inflammatory pathways and lead to poor survival of head and neck cancer patients (Chang, Lee, et al., 2019). Therefore, good oral hygiene is important to reduce the risk of developing oral cancer.

Furthermore, there is a higher risk of developing oral cancer in people with a family history of head and neck cancers (Omar, 2013). Patients with immune defects (suppression or deficiency) and potentially malignant disorders could also have a higher risk of developing oral cancer (Elimairi et al., 2017; Warnakulasuriya, 2009). On the other hand, the dietary pattern is suggested to have a role in oral cancer development. High consumption of combination (dairy, fermented/salted food, meat/by-products) and traditional food (starches and beverages) has been shown to increase the risk of oral cancer (Helen-Ng et al., 2012). The consumption of very hot drinks and food could increase the risk of oral cancer as well (Patel, Pathak, Patel, & Sutariya, 2018). Nevertheless, the high consumption of fruits and vegetables, which are rich in antioxidant properties could reduce the risk of oral cancer (Elimairi et al., 2017; Helen-Ng et al., 2012; Omar, 2013; Patel et al., 2018).

Universiti Malaya

2.1.3 Oral Potentially Malignant Disorder (OPMD) and Oral Cancer

Precancerous lesions and conditions have been widely used to describe the clinical and morphological presentations that may potentially transform into cancer. In 2005, WHO recommended the term “potentially malignant disorders” to include the lesions and conditions that indicates the risk of malignancy transformation (Warnakulasuriya et al., 2007). A range of precancerous lesions (leukoplakia, erythroplakia, and palatal lesions in reverse smokers), precancerous conditions (lichen planus, actinic keratosis, oral submucous fibrosis, and discoid lupus erythematosus), and hereditary disorders (dyskeratosis congenita and epidermolysis bullosa) are defined as oral potentially malignant disorders (OPMDs) (Warnakulasuriya et al., 2007).

OPMDs could manifest morphological alterations to increase the risk for malignant transformation, which precedes to the development of OSCC. Histopathological assessment for the presence and grade of oral epithelial dysplasia is a common method to predict the malignant transformation risk of OPMDs (Warnakulasuriya, Reibel, Bouquot, & Dabelsteen, 2008). The oral epithelial dysplasia is traditionally divided into grades of mild, moderate, and severe based on the cellular changes. No or mild dysplasia is considered as low risk of malignant transformation, and moderate or severe dysplasia is indicated as a high risk of malignant transformation (Warnakulasuriya et al., 2008). The presence of dysplasia is observed in most of the erythroplakia cases as well as in 1.0 to 30.0% of leukoplakia cases, whereas oral submucous fibrosis lesions do not have any dysplastic features (Dionne, Warnakulasuriya, Zain, & Cheong, 2015).

Leukoplakia represents the most common OPMD, while erythroplakia, oral lichen planus, and oral submucous fibrosis are also frequently seen in OPMDs (Speight, Khurram, & Kujan, 2018). Leukoplakia is defined as “a white patch or plaque of questionable risk having excluded other known diseases or disorders that carry no

Universiti Malaya

increased risk for cancer" (Warnakulasuriya et al., 2007). It is a lesion that has a dynamic spectrum from hyperkeratosis to different grades of oral epithelial dysplasia. It is not inevitable as non-dysplastic lesions could also transform into cancer (Speight et al., 2018). The worldwide prevalence rate of leukoplakia is reported at approximately 2.6% with a malignant transformation rate of 1.4% per year (Petti, 2003).

Erythroplakia is defined as "a fiery red patch that cannot be characterized clinically or pathologically as any other definable disease" (Warnakulasuriya et al., 2007). It is a well-defined clinical lesion with bright red granular surface. Erythroplakia is a high malignant potential lesion because most erythroplakia have high-risk dysplastic features and often present as carcinoma in situ or invasive carcinoma (Speight et al., 2018). Although the worldwide prevalence rate of erythroplakia is uncertain, it has low prevalence in the range from 0.01 to 0.21% (Villa, Villa, & Abati, 2011). In Malaysia and India, the prevalence rate of erythroplakia is estimated at 0.02% (Dionne et al., 2015). Nevertheless, erythroplakia showed a high malignant transformation rate of 44.9% (Villa et al., 2011).

Oral lichen planus is a chronic inflammatory disorder that affects the oral mucosa. It is mainly mediated by T-cell and increases the rate of differentiation of the stratified squamous epithelium, causing hyperkeratosis and erythema (Warnakulasuriya et al., 2007). The lesion may appear as a white lacy patch or red swollen patch. The malignant potential of lichen planus remains controversial, in which the malignant transformation rate is ranged from 0.2 to 1.0% (Dionne et al., 2015). A systemic review has indicated that oral lichen planus lesion may have a greater risk of malignant transformation predominantly in females and in older patients (Fitzpatrick, Hirsch, & Gordon, 2014).

Oral submucous fibrosis is a chronic disorder characterized by fibrosis of the submucosa on the oral cavity, resulting in difficulty to open the mouth. Different populations may have different affected sites within the mouth (Warnakulasuriya et al.,

Universiti Malaya

2007). Oral submucous fibrosis is predominant in the Asian population with the habit of betel quid and areca chewing (Elimairi et al., 2017). The malignant transformation rate of oral submucous fibrosis is ranged from 1.9 to 9.1%, and the risk is higher with the presence of epithelial dysplasia or concomitant lesions of leukoplakia in oral submucous fibrosis (Speight et al., 2018).

Although the global incidence and prevalence of OPMD are unavailable, the prevalence of OPMD was ranged from 0.2 to 11.3% (Dionne et al., 2015). The prevalence of OPMD is reported higher in Southeast Asia, Sri Lanka, Taiwan, and Papua New Guinea, and it is predominant in males (Johnson et al., 2011). OPMD lesions are frequently found in the younger age groups, and most likely to transform into cancer in older individuals (Speight et al., 2018). The risk factors of OPMDs are alike the risk factors of oral cancers, which includes the practice of tobacco smoking, alcohol drinking, betel quid chewing, and HPV infection (Johnson et al., 2011). However, the development of OPMDs in non-smokers could transform into OSCC due to an underlying genetic predisposition that may be involved in the disease (Speight et al., 2018).

2.1.4 Oral Cancer Staging and Grading

The staging of oral cancer is referred to the tumour-node-metastasis (TNM) system, which describes the primary tumour size status, regional lymph nodes status, and distant metastasis status. The TNM staging system is established by the Union of International Cancer Control (UICC) and the American Joint Committee on Cancer Staging (AJCC). The staging system always classified as stage 0 (carcinoma in situ), early stage (stages I and II), and advanced stage (stages III and IV) (Table 2.1). The TNM staging system is the most widely used prognostic system as it is a simple design and allows the TNM stage grouping for the tumour (Edge et al., 2010; Montero & Patel, 2015). The tumour size and

Universiti Malaya

lymph nodes metastasis are reliable indicator of patient prognosis and treatment planning (Warnakulasuriya, 2014).

Table 2.1: TNM staging of oral cancer

Stage	Tumour size status (T)	Lymph nodes status (N)	Metastasis status (M)
0	T_{is}	N0	M0
I	T1	N0	M0
II	T2	N0	M0
III	T3	N0	M0
	T1, T2 or T3	N1	M0
IV	T4	N0 or N1	M0
	Any T	Any N	M1

(adapted from Edge et al., 2010)

T_{is} = carcinoma in situ
T1 = tumour ≤ 2 cm in greatest dimension
T2 = tumour > 2 and ≤ 4 cm in greatest dimension
T3 = tumour > 4 cm in greatest dimension
T4 = moderately advanced or very advanced disease
N0 = no regional lymph nodes
N1 = enlarged ipsilateral lymph nodes present < 3 cm
N2 = ipsilateral single > 3-6 cm, ipsilateral multiple < 6 cm or bilateral/contra lateral < 6 cm
N3 = any palpable node > 6 cm
M0 = no distant metastasis
M1 = distant metastasis

The histopathological grading system was initially introduced by Broders in 1920. It is based on the assessment of the degree of keratinization, cellular and nuclear polymorphism, and mitotic activity (Pindborg, Reichart, Smith, & Van der Waal, 1997). This grading system was further modified and assessed on the differentiation status of the cells within the tumour cell population. Till now, this modification of Broders' grading system is widely used in the histopathological assessment of oral cancer. It is an essential tool to predict the clinical and biological behaviour of oral cancer. The morphological assessment tumours were classified into well, moderately, and poorly differentiated.

Universiti Malaya

The histological and cytological features in well-differentiated OSCC closely resemble those of the squamous epithelial lining of the oral mucosa. It contains 75 to 100% differentiated cells and 0 to 25% undifferentiated cells. Moderately differentiated OSCC displays neoplasm features intermediate between well-differentiated and poorly differentiated. It contains 50 to 70% differentiated cells and 25 to 50% undifferentiated cells. The histological and cytological features in poorly differentiated OSCC are only a minor similarity to the normal stratified squamous epithelium of the oral mucosa. More than 50% of the cells are undifferentiated.

At the same time, Broders' histological grading has limited value in the prognosis and treatment response of OSCC (Pindborg et al., 1997). This could be due to OSCC usually consists of heterogeneous cell populations with various degrees of differentiation in invasiveness and metastasis behaviour (Anneroth, Batsakis, & Luna, 1987). Nevertheless, this histological grading could be taken as a valuable diagnostic and predictive factor of lymph node metastasis (Akhter, Hossain, Rahman, & Molla, 2011).

2.1.5 Diagnosis, Prognosis, and Treatment

Most of the oral cancers are frequently diagnosed with signs and symptoms at the advanced stages (Markopoulos, 2012; Warnakulasuriya, 2009). The sign and symptoms at the early stage of OSCC are often vague and painless, but it may have burning sensation or pain at the advanced stage of OSCC (Markopoulos, 2012). The related signs and symptoms in the development of OSCC usually include non-healing ulceration, non-healing extraction sockets, swelling of the mouth and neck, firm and fixed lesions, and lymph node enlargement (Elimairi et al., 2017; Markopoulos, 2012). If there is any oral symptoms persist for more than two weeks, it is imperative to consider OSCC or some underlying serious conditions (Markopoulos, 2012).

Universiti Malaya

A comprehensive head and neck examination that involved the physical examination of the oral cavity and clinical assessment of the lymph nodes involvement is required in patients with suspected oral cancer. This is followed by histopathological examination on suspected tissue biopsy. On the other hand, several non-invasive diagnostic aids such as toluidine blue staining, oral brush biopsy, and optical biopsy have been used for the diagnosis of oral cancer (Omar, 2013). Toluidine blue staining is a practical, inexpensive, and sensitive diagnostic tool. It can optimally identify a biopsy site with dysplastic and cancerous lesions (Elimairi et al., 2017). Whereas the oral brush biopsy is a reliable, easy, and risk-free technique that is well accepted by the patients. This biopsy method obtains oral epithelial cells to identify dysplastic and cancerous lesions, but it may cause false-negative results (Elimairi et al., 2017). Optical biopsy detects the dysplastic and cancerous lesions based on the assumption that their metabolic or structural have different absorbance and reflectance properties for light. The optical biopsy systems included chemiluminescence (Vizilite), fluorescence (Veloscope), tissue fluorescence spectroscopy, and Raman spectroscopy (Omar, 2013). However, these systems are costly and have low sensitivity or specificity (Elimairi et al., 2017).

Concurrently, imaging techniques such as computed tomography (CT) scanning, magnetic resonance imaging (MRI) scanning, chest radiography, and positron emission tomography (PET) scanning are used in the diagnosis of oral cancer to investigate the primary site of the tumour and suspected sites of lymph node or distant metastasis. CT scanning is beneficial in the evaluation of bone and neck nodes, whereas MRI scanning is capable in the characterization of soft tissue and perineural invasion (Montero & Patel, 2015). While chest radiography can help to assess the metastasis sites of OSCC such as lungs, hilar lymph nodes, ribs, and vertebrae (Omar, 2013). PET scanning is a nuclear medicine imaging to produce three-dimensional imaging using the radioactive drug, 18-fluorodeoxyglucose in the detection of unknown primary tumour sites (Omar, 2013). By

Universiti Malaya

combining these current imaging modalities would have a better efficacy and sensitivity to evaluate the locoregional lymph node and distant metastasis, but it may give false-positive results (Palasz, Adamski, Gorska-Chrzastek, Starzynska, & Studniarek, 2017).

After confirmation of diagnosis, the main treatments for oral cancer patients are mostly surgery and radiotherapy. Lymph node involvement and tumour size are important prognostic factors for oral cancer patients (Sankaranarayanan, Ramadas, Amarasinghe, Subramanian, & Johnson, 2015). A better treatment outcome is accomplished in patients diagnosed at the early stage of oral cancer. It can reduce the unfavourable effects of treatment such as surgery reconstruction and have a better quality of life (Elimairi et al., 2017).

Surgery and radiotherapy without functional morbidity are the primary treatment for early stage of oral cancer. Nevertheless, the treatment options depend on several factors, which include the tumour sites, cosmetic and functional consequences, age of the patient, presence of other illnesses, patient's preference, and the availability of expertise (Sankaranarayanan et al., 2015). Alternatively, external beam radiotherapy and/or brachytherapy that use radioactive source may be employed as a single modality for the early stage of oral cancer with a well-defined primary tumour (Huang & O'Sullivan, 2013).

A combination modality strategy of surgery, radiotherapy, and/or chemotherapy is the treatment that often used for the advanced stage of oral cancer. However, the locoregional treatment failure rates are high due to the aggressive form of the advanced oral cancer (Sankaranarayanan et al., 2015). Surgery followed by postoperative radiotherapy and/or chemotherapy is the preferred modality for patients with unfavourable pathological features. Whereas radiotherapy and chemotherapy are usually reserved for patients in the

Universiti Malaya

advanced stage of oral cancer without bone involvement or patients who are unsuited for surgery (Sankaranarayanan et al., 2015).

The radiotherapy treatment is usually recommended to begin with a typical radiation dose of approximately 60 Gy within 6 weeks after the surgery (Omura, 2014). While adjuvant chemotherapy is a standard treatment for the advanced stage of oral cancer where the most common regimen consists of CDDP (cis-dichloro-diamine-platinum or cisplatin) had shown a survival benefit (Omura, 2014). It has been suggested that surgery with concurrent radiotherapy and chemotherapy should be considered for those high-risk patients with pathologically positive lymph nodes (Huang & O'Sullivan, 2013).

Despite recent advances in the multimodality treatment of oral cancer, the survival rates of oral cancer patients have not improved much. The development of targeted therapies can be a complement to current cancer treatment to improve the efficacy of the treatment outcome. Several targeted therapies, which include monoclonal antibodies and small-molecule inhibitors that target epidermal growth factor receptor (EGFR), vascular endodermal growth factor (VEGF), mechanistic target of rapamycin (mTOR), and programmed cell death protein-1 (PD-1) have been proposed (Kozakiewicz & Grzybowska-Szatkowska, 2018). These targeted molecules are associated with cell proliferation, apoptosis, angiogenesis, and invasion of OSCC. The monoclonal EGFR antibody, cetuximab as well as the monoclonal PD-1 antibodies, pembrolizumab and nivolumab were approved by the United States Food and Drug Administration (FDA) for the treatment of advanced and recurrent or metastatic head and neck cancers (Kozakiewicz & Grzybowska-Szatkowska, 2018). Nevertheless, the therapeutic effect of these therapies in OSCC remains uncertain.

Universiti Malaya

2.1.6 Oral Carcinogenesis

Cancer can be defined by six hallmarks, which include self-sufficiency in proliferation signals, evasion of growth suppressors, resistance to cell death, limitless replicative potential, increased angiogenesis, activation of invasion, and metastasis (Hanahan & Weinberg, 2000). In the past decade, reprogramming of energy metabolism and evading immune destruction had counted in as two emerging hallmarks of cancer (Hanahan & Weinberg, 2011). All these features had revealed the conceptual framework of cancer development and progression for a better understanding on the complex biology of cancer.

Oral carcinogenesis is a multistep and multifactorial process that involved the accumulation of genetic changes such as mutation, amplification of oncogenes, and inactivation of tumour suppressor genes, where these changes are modulated by genetic predisposition and environmental carcinogen exposure (Choi & Myers, 2008). It is believed that the histopathological progression of oral carcinogenesis reflected the accumulation of genetic changes (Lingen & Kumar, 2005; Tanaka & Ishigamori, 2011). Oral carcinogenesis initially begins with the phenotypic changes, through hyperplasia to dysplasia, followed by carcinoma *in situ*, and transform into invasive squamous cell carcinoma (Figure 2.1). However, not all lesions exhibiting dysplasia or potentially malignant lesions will eventually develop into cancer (Neville & Day, 2002; Warnakulasuriya et al., 2007).

Figure 2.1: Histological progression of oral carcinogenesis
(adapted from Lingan & Kumar, 2005)

Universiti Malaya

The genetic changes in oral carcinogenesis are mainly due to the activation of oncogenes and inactivation of tumour suppressor genes, which affect the cell cycle regulation, cell proliferation, differentiation, and apoptosis (Choi & Myers, 2008; Hanahan & Weinberg, 2000; Krishna, Singh, Kumar, & Pal, 2015). Oncogenes are genes derived from the alteration of cellular proto-oncogenes, which may cause the growth of cancer cells. Several oncogenes such as EGFR, rat sarcoma viral (*ras*), *c-myc*, and cyclin D1 (CCND1) are associated with oral carcinogenesis (Choi & Myers, 2008; Krishna et al., 2015).

EGFR is a biological receptor of epidermal growth factor (EGF) and transforming growth factor-α (TGF-α). It plays a key role in cell proliferation, apoptosis, and angiogenesis. The overexpression of EGFR and TGF-α is observed in OPMD and OSCC, where the co-expression of EGFR and TGF-α stimulates cell proliferation by the autocrine and paracrine mechanism (Srinivasan & Jewell, 2001). Furthermore, the overexpression and amplification of EGFR are frequently seen in OSCC (Choi & Myers, 2008; Huang et al., 2012; Sarkis, Abdullah, Majeed, & Talabani, 2010). The overexpression of EGFR was associated with poor prognosis, and the amplification of EGFR is associated with the advanced stage of OSCC (Huang et al., 2012).

The *ras* oncogene family consists of three members (K-*ras*, H-*ras*, and N-*ras*) that involved in cell growth regulation and signal transduction. They encode protein p21 by binding with guanosine triphosphatase, which can be constitutively stimulated through mutation (Krishna et al., 2015). H-*ras* is one of the most frequently mutated oncogenes in OSCC (Murugan, Munirajan, & Tsuchida, 2012). The H-*ras* mutations are common in the Asian population associated with tobacco smoking and betel quid chewing habits (Murugan et al., 2012).

Universiti Malaya

As a transcription factor, *c-myc* has a role in the regulation of cell proliferation, differentiation, and apoptosis. It activates the cyclin-dependent kinase, facilitating transcription, and consequently regulating cell proliferation (Perez-Sayans et al., 2011). Approximately 80.0% of the overexpression of *c-myc* in OSCC is mainly caused by gene amplification (Pai et al., 2009). The overexpression of *c-myc* was found to be involved in the initial phases of malignant transformation in oral precancerous lesions and associated with poor prognosis in OSCC (Eversole & Sapp, 1993).

CCND1 is a transcription factor involved in the regulation of the cell cycle. The overexpression of CCND1 has been reported in 45.0 to 66.0% of OSCC (Miyamoto, Uzawa, Nagaoka, Hirata, & Amagasa, 2003; Saawarn, Astekar, Saawarn, Dhakar, & Gomateshwar Sagari, 2012). It is believed that the gene amplification of *CCND1*, and consequently of protein overexpression is an early event during oral carcinogenesis. Furthermore, the overexpression of CCND1 is associated with tumour size, lymph node metastasis, poor prognosis, and low survival status in the Asian population (Zhao et al., 2014).

Tumour suppressor genes are negative cell regulators that normally suppress cell division or growth. The inactivation of tumour suppressor genes is a crucial event in malignant transformation as well. Several tumour suppressor genes such as retinoblastoma (*Rb* gene), cyclin-dependent kinase inhibitor 2A (CDKN2A), and tumour protein p53 (TP53) are associated with oral carcinogenesis (Choi & Myers, 2008; Krishna et al., 2015).

Rb gene is one of the key regulators of cell cycle progression. The loss of Rb expression is an early event of oral carcinogenesis. It has been observed in 23.0 to 25.0% of leukoplakia and 56.0 to 66.0% of OSCC (Nakahara et al., 2000; Pande, Mathur, Shukla,

Universiti Malaya

& Ralhan, 1998). A significant loss of Rb expression with the overexpression of CCND1 was also found to be involved in the transition of OPMD to OSCC (Soni et al., 2005).

CDKN2A encodes p16(INK4A) and p14(ARF) proteins to take part as inhibitors of cell cycle progression in the Rb and TP53 pathways. The loss of CDKN2A expression due to promoter hypermethylation has been observed in 57.0 to 60.0% of OPMD and 60.0 to 82.0% of OSCC (Asokan, Jeelani, & Gnanasundaram, 2014; Bhatia et al., 2014). These studies indicated that it is involved in the early event of oral carcinogenesis. Additionally, the loss of CDKN2A expression is associated with poor prognosis and low survival rate of OSCC (Padhi et al., 2017).

TP53 is a tumour suppressor gene that regulates cell cycle progression, cell proliferation, DNA repair, and induces apoptosis. It is commonly mutated in nearly 70.0% of all cancers (Krishna et al., 2015). The overexpression of TP53 has been detected in 16.0 to 77.0% of OPMD and 34.0 to 77.0% of OSCC (Ghanghoria, Ghanghoria, & Shukla, 2015; Khanna, Vidhyarthi, Khanna, Singh, & Singh, 2012; Rowley, Sherrington, Helliwell, Kinsella, & Jones, 1998). Although the overexpression of TP53 was not always correlated with *TP53* mutation, *TP53* mutation is the common event in oral carcinogenesis (Khanna et al., 2012; Rowley et al., 1998). Moreover, the overexpression of TP53 with a high probability of mutation is associated with poor prognosis and survival rate of OSCC (Cutilli, Leocata, Dolo, & Altobelli, 2016).

Universiti Malaya

2.2 Cancer Biomarker Discovery

According to the National Institute of Health (NIH), a biological marker or biomarker is defined as any biological characteristic that can be used as an indicator of normal biological processes, pathogenic processes, or pharmacological responses to a therapeutic invention (Biomarkers Definitions Working, 2001). Therefore, biomarkers have many valuable utilizations for the detection and monitoring of disease. Biomarkers can be deoxyribonucleic acid (DNA), ribonucleic acid (RNA), protein, metabolite, or any other biological molecules associated with a diseased state (Kulasingam & Diamandis, 2008).

Biomarkers can be classified based on their utilities, which are diagnostic biomarkers, disease staging biomarkers, prognostic biomarkers, and biomarkers for monitoring of therapy response to intervention (Biomarkers Definitions Working, 2001). Diagnostic biomarkers can be used to identify those patients with a disease or abnormal condition, and disease staging biomarkers can be used to rule out the severity of the disease. On the other hand, prognostic biomarkers and biomarkers for monitoring of therapy response can be used to predict tumour behaviour and the efficacy of a therapy, respectively. These biomarkers can predict a clinical outcome and indicate the risk of benefit or harm to patients (Strimbu & Tavel, 2010).

Ideally, a cancer biomarker should be easily accessible, reliable, and cost-effectiveness to measure in a minimally invasive or non-invasive method with high sensitivity and specificity (Pavlou et al., 2013). Biological samples, which include tumour tissues or body fluids such as blood (serum and plasma), saliva, urine, sputum, cerebrospinal fluid, and bone marrow aspirate are generally used for the discovery of cancer biomarkers. Among these biological samples, blood is the most easily accessible and it contains proteins secreted, shed, or otherwise released by the tissues (Pavlou et al., 2013; Zhang et al., 2007).

Universiti Malaya

In recent years, the emergence of "omics" technologies such as microarray, next-generation sequencing (NGS), and mass spectrometry has contributed to the advancements in cancer biomarker research (Kulasingam & Diamandis, 2008; Mabert et al., 2014). Microarray and NGS technologies have been developed for genomics and transcriptomics analyses to assess gene mutation, single nucleotide polymorphism, and changes in gene expression. Using microarray and NGS technologies, numerous studies were conducted to identify the genetic variants and gene expression associated with cancer risks (Fabris et al., 2016; Idris, Ahmad, Scott, Vassiliou, & Hadfield, 2013; Li, Fu, & Xiao, 2015; Liu, So, & Fan, 2015). Meanwhile, the majority of proteomics-based cancer biomarker discovery studies were conducted using mass spectrometry technology to detect the changes in protein expression (Chung et al., 2014; Crutchfield, Thomas, Sokoll, & Chan, 2016; Zhang et al., 2004).

2.2.1 Proteomics in Biomarker Discovery

The term "proteome" was first coined to describe all the protein complement encoded by a genome (Wilkins et al., 1996). The proteome is inherently more complex than genome and transcriptome due to the proteome is highly dynamic and changes in response to the physiological status of the organism (Hudler, Kocevar, & Komel, 2014). The proteome also varies from cell to cell, time to time, and in response to different stimuli.

Although cancer is well characterized by the accumulation of genetic changes, proteins are the functional molecules that most probably reflect the actual disease state as they represent the biological endpoint (Hudler et al., 2014). Proteomics is the large-scale study of the composition, structure, function, and interaction of proteins in complex biological samples. Therefore, proteomics is one of the most extensively used approaches for the discovery of cancer biomarkers (Sallam, 2015).

Universiti Malaya

Proteomics can be classified into three major categories, which are expression proteomics, structural proteomics, and functional proteomics (Graves & Haystead, 2002). Expression proteomics aims to study the expression of protein levels qualitatively and quantitatively between samples that are in different conditions, for instance, healthy individuals and those who have disease. Structural proteomics attempts to understand three-dimensional structural of protein complexes in a specific cellular organelle. Whereas functional proteomics facilitates to understand the biological functions of the protein and molecular mechanisms of protein-protein interaction networks. Hence, proteomics has become essential in biomarker discovery where different properties of proteins are studied to inspect the defective proteins' structure, function, and interaction networks within the cancer states (Sallam, 2015).

By using proteomics analysis, three biomarkers, apolipoprotein A1, transthyretin, and inter-alpha-trypsin inhibitor heavy chain H4 were identified for the early detection of ovarian cancer (Zhang et al., 2004). Several potential biomarkers associated with cancer progression were reported in breast, colorectal, lung, and bladder cancers (Da Costa et al., 2015; Gan et al., 2014; Kelemen et al., 2020; Nedjadi et al., 2020; Wang, Lu, et al., 2016). Distinct serum proteins that elicited in immune response could act as useful biomarkers in nasopharyngeal and thyroid cancers which can be identified through mass spectrometry technology (Lin, Xu, et al., 2017; Yekta, Oskouie, Tavirani, Mohajeri-Tehrani, & Soroush, 2018).

Furthermore, there are many studies have been conducted to identify potential biomarkers of OSCC using mass spectrometry technology and proteomics techniques such as enzyme-linked immunosorbent assay (ELISA), Western blot, and immunohistochemistry (Arellano-Garcia et al., 2010; Chanthammachat et al., 2013; Chen et al., 2014; Csosz et al., 2018; Jessie et al., 2013; Jou et al., 2010; Thiel et al., 2011; Tung

Universiti Malaya

et al., 2013; Turhani et al., 2006; Wu, Chu, Hsu, Chang, & Liu, 2015; Yu, Chang, et al., 2011; Yu et al., 2016). The identified biomarkers displayed aberrant protein expression and participated in various biological processes such as cell signalling, cell proliferation, invasion, and survival. It was observed that most of the identified biomarkers for the detection of OSCC were proteins with up-regulated levels, such as α1-antitrypsin, haptoglobin, complement C3, hemopexin, transthyretin, transferrin, and guanylate binding protein-1 (Jessie et al., 2013; Jou et al., 2010; Yu, Chang, et al., 2011). Nonetheless, the downregulation of tetranectin was found as a potential prognostic biomarker in metastatic OSCC (Arellano-Garcia et al., 2010).

2.2.2 Glycoproteomics in Biomarker Discovery

Glycoproteomics is an important subfield of proteomics. Post-translational modifications (PTMs) that occur during or after protein biosynthesis involve in modulating the protein functions. Thus, the analysis of PTMs could yield insight into their roles to the disease states such as inflammation and cancer. PTM generally refers as enzymatic modification on the amino acid side chains in a protein (Walsh, Garneau-Tsodikova, & Gatto, 2005). Currently, more than 200 different types of PTMs have been discovered and implicated in the regulation of cellular processes such as protein stability, signal transduction, and protein-protein interaction (Pagel, Loroch, Sickmann, & Zahedi, 2015). These PTMs include phosphorylation, glycosylation, ubiquitylation, methylation, and acetylation (Walsh et al., 2005).

Glycosylation is one of the most common PTM processes, with more than 50.0% of all proteins are predicted to be glycosylated (Apweiler et al., 1999; Yang et al., 2017). It involves the enzymatic attachment of a range of different carbohydrate molecules (monosaccharide or oligosaccharide) to the protein surface. Most proteins that synthesized in the endoplasmic reticulum (ER) and Golgi apparatus undergo

Universiti Malaya

glycosylation. Glycosylation can act as a key regulatory mechanism in numerous biological processes such as maintenance of protein conformation, protein protection from proteolytic degradation, protein mobility, and signal transduction (Roth et al., 2012). There are several types of glycosylation based on the linkage of sugar-peptide bond and the oligosaccharide as attached, which includes *N*-, *O*-, and *C*-linked glycosylation, glypiation, and phosphoglycosylation.

Mainly, *N*- and *O*-glycosylation (*N*- and *O*-linked glycosylation) are the two most frequently detected types of glycosylation (Roth et al., 2012; Yang et al., 2017). *N*-glycosylation occurs on asparagine (Asn) residue within the consensus sequence Asn-X-Ser(Thr), where X can be any amino acid except proline, whereas *O*-glycosylation occurs on the side chain hydroxyl oxygen of serine (Ser) or threonine (Thr) residue (Kornfeld & Kornfeld, 1985; Steen, Rudd, Dwek, & Opdenakker, 1998). Dissimilar to *N*-glycosylation, there is no consensus sequence that has been reported for *O*-glycosylation.

N-glycosylation begins as a co-translational event that often happens in ER, where common oligosaccharide precursor ($Glc_3Man_9GlcNAc_2$) that containing 14 sugar oligosaccharides (2 *N*-acetylglucosamines, 9 mannoses, and 3 glucoses) is biosynthesized, and the proteins are transferred to Golgi apparatus (Colley, Varki, & Kinoshita, 2015). The proteins are further diversified in the Golgi apparatus. Thus, there are three major types of oligosaccharide structures: high mannose, complex, and hybrid type in *N*-glycoprotein (Kornfeld & Kornfeld, 1985; Roth et al., 2012).

O-glycosylation of secreted and membrane-bound proteins is a post-translational event that only takes place in the Golgi apparatus (Colley et al., 2015). The most common type of *O*-glycosylation involves the attachment of *N*-acetylgalactosamine (GalNAc), and it is also known as mucin-type *O*-glycosylation (Roth et al., 2012; Steen et al., 1998). The attachment of GalNAc in *O*-glycosylation can be further extended with several

Universiti Malaya

monosaccharides, including GalNAc, *N*-acetylglucosamine (GlcNAc), galactose, and fucose (Brockhausen & Stanley, 2015). Hence, there are four common core structures that have been identified: core 1 (GalGalNAc), core 2 (GlcNAcGalNAc), core 3 (GlcNAcGalNAc), and core 4 (GlcNAc2GalNAc) in *O*-glycoprotein (Brockhausen & Stanley, 2015).

For decades, aberrant glycosylation is often associated with proliferation, invasion, metastasis, angiogenesis, cell adhesion, and immune response during cancer development and progression (Dube & Bertozzi, 2005; Hakomori, 2002; Kailemia et al., 2017; Pinho & Reis, 2015). Aberrant expression of *N*-and *O*-glycoproteins has been detected in many cancers, which include pancreatic, lung, and breast cancers (Pan et al., 2016; Rho, Roehrl, & Wang, 2009; Semaan et al., 2012). It has also been suggested that aberrant glycosylation is often associated with all the hallmarks of cancer (Munkley & Elliott, 2016).

Numerous glycoproteomics studies were conducted to detect potential glycoprotein cancer biomarkers. Glycoproteomics analysis on ovarian cancer has discovered several glycoproteins with tumour-specific glycosylation changes such as biglycan precursor, periostin, and thrombospondin 1 (Abbott et al., 2010). Unique glycoprotein signatures and their glycoproteomics characterization have been reported as potential biomarkers of bladder cancer (Yang et al., 2011). Similarly, the glycoproteomics approach was employed to identify biomarkers for breast cancer (Fry, Sinclair, Timms, Leathem, & Dwek, 2013; Semaan et al., 2012; Tan et al., 2015). The detection of glycoproteins in pancreatic cancer indicated the important implications of glycoproteins in the development and progression of cancer (Drabik et al., 2017).

In OSCC, altered expression of glycoproteins that caused by aberrant glycosylation in blood or malignant tissues also has been associated with cellular invasion and cancer

Universiti Malaya

progression (Guu et al., 2017; Lin, Lin, Shi, Chang, & Wu, 2015; Manoharan et al., 2004; Shah, Telang, Raval, Shah, & Patel, 2008). The dysregulation of the *N*-glycosylation-regulating gene (DPAGT1), which interacts with the Wnt signalling pathway and inhibits E-cadherin adhesion, has been associated with OSCC progression (Jamal et al., 2012; Liu, Sengupta, et al., 2013). Whereas the *O*-glycosylation of EGFR that mediated by polypeptide *N*-acetylgalactosaminyltransferase 2 (GALNT2) is associated with OSCC cell migration and invasion (Lin, Huang, et al., 2014). By using glycoproteomics analysis, several glycoproteins were identified as potential OSCC diagnostics biomarkers such as apolipoprotein A-IV, leucine-rich alpha-2-glycoprotein (LRG1), B7-H3, CD147, IgG1, IgG4, haptoglobin, and transferrin (Chang, Lin, et al., 2019; Chen et al., 2015; Chen, Chong, et al., 2013; Saraswat et al., 2018).

Most of the currently used cancer biomarkers that clinically utilized for cancer diagnosis, prognosis as well as monitoring disease recurrence and treatment response are glycoproteins. These biomarkers include carcinoma antigen 125 (CA 125) for ovarian cancer, CA 15-3 for breast cancer, CA 19-9 for pancreatic cancer, carcinoembryonic antigen (CEA) for colon cancer, and prostate-specific antigen (PSA) for prostate cancer (Kailemia et al., 2017; Pinho & Reis, 2015). All these biomarkers have been shown with aberrant glycosylation in cancer (Pinho & Reis, 2015). More remarkably, AFP-L3 (a glycosylated form of alpha-fetoprotein) that reacts with lectin *Lens culinaris* agglutinin has been approved by the Food and Drug Administration (FDA) as a more specific marker for liver cancer (Drake et al., 2006). Therefore, this has been shown that the detection of biomarkers has enormous potential for further optimization and increase assay specificity.

Universiti Malaya

2.3 Proteomics and Glycoproteomics Approaches to Identify Potential Biomarkers

2.3.1 Two-dimensional Gel Electrophoresis (2-DE) and Mass Spectrometry

Current proteomics approaches are mostly mass spectrometry-based approaches, which can be divided into top-down and bottom-up strategies. The top-down strategy analysed intact proteins and their fragments, whereas the bottom-up strategy is the most widely used strategy that analysed peptides derived from digested proteins (Feist & Hummon, 2015; Yates, Ruse, & Nakorchevsky, 2009). Moreover, the gel-based mass spectrometry approach is one of the most commonly used bottom-up strategies (Feist & Hummon, 2015; Sallam, 2015).

Two-dimensional gel electrophoresis (2-DE) was developed in 1975 for the separation and fractionation of complex protein mixtures (O'Farrell, 1975). This gel-based method was further enhanced to have better technical performance in the 1980s (Bjellqvist et al., 1982; Gorg, Postel, & Gunther, 1988). In the first-dimensional separation, the proteins are separated according to their charges or isoelectric points (pI) by isoelectric focusing (IEF), and later the proteins are separated according to their masses or molecular weights by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) in the second-dimensional separation (Pomastowski & Buszewski, 2014).

After second-dimensional separation, the proteins are resolved into spot forms and visualized using staining methods such as Coomassie brilliant blue and silver staining. Nevertheless, silver staining is more sensitive than Coomassie brilliant blue staining for detection of the low molecular weight and small amounts of proteins (Shevchenko, Wilm, Vorm, & Mann, 1996). The stained gel images can be captured and analysed with 2-DE specialized software. The spots of interest are then excised and digested for further identification by mass spectrometry.

Universiti Malaya

Mass spectrometry (MS) is widely used for proteomics biomarker discovery over the years. It is a versatile and sensitive technique to detect, identify, and quantitate molecules based on their mass-to-charge ratio (m/z). MS uses mass analysis for large-scale characterization of proteomics (Yates et al., 2009). The main components of all mass spectrometry instruments are an ion source, a mass analyser, and an ion detector.

Electrospray ionization (ESI) and matrix-assisted laser desorption ionization (MALDI) are widely used methods in the soft ionization. These two methods were developed in the late 1980s to allow proteins and peptides to be analysed by MS (Karas, Ingendoh, Bahr, & Hillenkamp, 1989; Whitehouse, Dreyer, Yamashita, & Fenn, 1985). The ESI source produces ions from the liquid phase, whereas MALDI uses a laser energy absorbing matrix to generate ions (Yates et al., 2009). Both methods have high sensitivity and accuracy for the characterization of biomolecules. Nonetheless, MALDI is the most popular ionization technique in molecular MS analysis because of its robustness, high throughput, and ease of automation (Pomastowski & Buszewski, 2014).

A mass analyser separates ions according to their m/z ratio after ionization. Quadrupole, ion trap, time-of-flight (TOF), orbitrap, and fourier transform ion cyclotron resonance (FTICR) are the common types of mass analysers. Each mass analyser has specific characteristics and analytical performance. TOF analyser is typically coupled with MALDI ionization source to perform pulsed ionization, while quadrupole, ion trap, orbitrap, and FTICR are often interfaced with ESI sources (Yates et al., 2009). Once ions emerge from the mass analyser, the number of ions corresponding to an m/z is recorded by ion detectors such as electron multiplier or microchannel plate detector.

The protein identification was performed using peptide mass fingerprinting, which is based on the accurate mass measurement of peptides and complemented via protein sequence database search (Zhang, Annan, Carr, & Neubert, 2014). Hence, it is important

Universiti Malaya

to obtain high mass accuracy of tryptic peptides for more precise searches upon protein identification. Hybrid mass spectrometers that combined more than one mass analyser have been built for more accurate and sensitivity capabilities. Thus, MALDI-TOF/TOF mass spectrometer represents the advancement of MALDI ionization with tandem mass spectrometry (MS/MS) in the TOF/TOF instrument (Medzihradszky et al., 2000).

The TOF/TOF instrument can only be couple with MALDI (Zhang et al., 2014). The TOF analyser allows the ions to fly through the "flight tube" and reach the detector in a field-free vacuum. In MALDI-TOF/TOF, two TOF analysers are placed in series where the first TOF selects ions by subjecting to high-energy collision-induced dissociation, and the fragment ions were measured by the second TOF (Medzihradszky et al., 2000). The resulting fragmented peptides have a more specific sequence to provide higher accuracy for protein identification (Zhang et al., 2014).

2.3.2 Lectin-based Glycoproteomics Approaches

As a subfield of proteomics, glycoproteomics that focuses on analysing the entire complement of glycoproteins is rapidly emerging. It integrates the separation or enrichment of glycoproteins and proteomics technologies for the identification and characterization of glycoproteins. Lectins are often used in glycoproteomics study because they are carbohydrate-binding proteins that could recognize glycoproteins and specifically bind to the carbohydrate epitopes of glycoproteins (Drake et al., 2006; Kailemia et al., 2017; Sharon & Lis, 2004). Lectins were first revealed in 1888 by Hermann Stillmark, who observed ricin that was isolated from seeds of castor beans (*Ricinus communis*) agglutinated human and some animal red blood cells (Sharon & Lis, 2004). Lectins present ubiquitously in nature, including plants, animals, and microorganisms. Plant lectins have been studied most intensively as most of the lectins were isolated from plants (Dan, Liu, & Ng, 2016; Gupta, 2020; Hashim, Jayapalan, &

Universiti Malaya

Lee, 2017). Lectins were mainly found in the seeds, and they were also found in roots, tubers, leaves, flowers, rhizomes, fruits, and stems.

Based on the sequence similarities and structural homology, lectins were categorised into 12 different families, namely *Agaricus bisporus* agglutinin homologs, amaranthins, class V chitinase homologs with lectin activity, cyanovirin family, *Euonymus europaeus* agglutinin family, *Galanthus nivalis* agglutinin family, proteins with hevein domains, jacalins, legume lectin, lysine motif domain, nictaba family, and ricin-B family (Jiang, Ma, & Ramachandran, 2010). These lectins have their carbohydrate-binding specificities, such as GlcNAc, GalNAc, galactose, and mannose. For example, wheat germ agglutinin (WGA) from hevein family binds specifically to GlcNAc, *Sambucus nigra* agglutinin (SNA) from ricin-B family exhibits specificity to GalNAc, and concanavalin A (Con A) from legume family that selectively binds to mannose (Dan et al., 2016).

Lectins have become a useful tool in glycoproteomics study because of the wide availability, high affinity, and relative specificity (Gupta, 2020; Hashim et al., 2017). Lectin-based methods including lectin affinity chromatography, lectin blotting, lectin microarray, enzyme-linked lectin assay, and lectin histochemistry/cytochemistry have been applied for glycoprotein enrichment and detection (Dai, Zhou, Qiu, Liu, & Fan, 2009; Dan et al., 2016; Hashim et al., 2017). Lectin affinity chromatography is commonly used to separate glycoproteins from non-glycosylated proteins with lectin that bound to an appropriate matrix in a column, while lectin blotting is analogous to immunoblotting where the proteins are transferred onto a nitrocellulose or polyvinylidene fluoride (PVDF) membrane and incubated with a specific lectin after completion of SDS-PAGE (Dai et al., 2009; Dan et al., 2016). Similarly, lectin microarray, enzyme-linked lectin assay, and lectin histochemistry/cytochemistry are also developed from proteomics techniques that

Universiti Malaya

use lectin instead of the antibody to detect the glycoproteins (Dan et al., 2016; Hashim et al., 2017).

In the past few years, lectin-based glycoproteomics approaches with MS analysis have been employed in cancer biomarker discovery (Chantaraamporn et al., 2020; Dai et al., 2009; Drake et al., 2010; Drake et al., 2006; Hashim et al., 2017; Kailemia et al., 2017; Lastovickova, Strouhalova, & Bobalova, 2020; Pan, Chen, Aebersold, & Brentnall, 2011). Con A and Jacalin are one of the most common lectins used in the detection of *N*-glycoprotein and *O*-glycoprotein, respectively. Con A is a legume lectin found in jack bean seeds, *Canavalia ensiformis*. It binds specifically to mannose and glucose where mannose has a higher affinity (Dan et al., 2016; Lis & Sharon, 1998). The use of Con A affinity chromatography coupled with 2-DE and MS analysis has identified several differentially expressed *N*-glycoproteins such as haptoglobin, hemopexin, and clusterin in colorectal cancer (Rodriguez-Pineiro, Ayude, Rodriguez-Berrocal, & de la Cadena, 2004). Several *N*-glycoproteins such as transthyretin, α1-antitrypsin, and complement C3 were detected in lung cancer with the application of Con A affinity chromatography coupled with 2-DE and MS analysis (Tran, Nguyen, Nguyen, & Phan, 2008). Recently, this strategy had also demonstrated the discovery of several candidate biomarkers that bound with Con A lectin in breast and gastric cancers (Tan et al., 2015; Uen et al., 2013).

Jacalin is a plant-based lectin found in jackfruit seeds, *Artocarpus integrifolia*. It binds specifically to galactose or GalNAc (Dan et al., 2016; Lis & Sharon, 1998). Lectin blotting and lectin histochemistry with Jacalin were used for the analysis of *O*-glycoproteins containing core structure of Galβ1–3GalNAc in colorectal cancer (Dudas, Yunker, Sternberg, Byrd, & Bresalier, 2002; Said et al., 1999). A high level of *O*-glycosylated Mucin 5AC (MUCA5C) that were bound with Jacalin was observed in pancreatic cancer as well (Yue et al., 2009). Using lectin microarray and lectin assay with

Universiti Malaya

Jacalin, glycosylated PSA was found specifically expressed in aggressive prostate cancer (Li et al., 2011). Recently, *O*-glycosylated core 1 β1,3-galactosyltransferase (C1GalT1) that was recognized by Jacalin lectin blotting had found to have a potential role in radiation therapy of oesophageal cancer (Zhang, Deng, et al., 2018).

Till date, only a few studies have been conducted using Con A and Jacalin lectin in OSCC. The intensity of histochemical staining that bound with Con A lectin was decreased during oral carcinogenesis in hamster buccal pouch mucosa (Jin & Lin, 1989). On the other hand, little differences in Con A lectin staining was reported in human tissue samples during oral carcinogenesis (Kannan et al., 1993; Mazumdar, SenGupta, Param, & Sinha, 1993). While the tissue binding pattern of Jacalin lectin could predict the malignant potential of the oral cavity (Vijayan et al., 1987), it also could be used to predict the lymph node metastatic potential of OSCC (Remani et al., 1997). Although the use of these lectins has limited exploration in OSCC, these previous studies had demonstrated the detection of lectins could help to recognize the changes in glycoproteins during the oral carcinogenesis. Furthermore, lectin-based strategies that involved 2-DE technique were used to identify the potential glycoprotein biomarkers in many cancer studies (Hashim et al., 2017). Thus, it is believed that lectin-based glycoproteomics approaches with MS analysis could be a reliable methodology for the biomarker discovery of OSCC.

Universiti Malaya

CHAPTER 3: METHODOLOGY

3.1 Sample Collection

The serum and formalin-fixed paraffin-embedded (FFPE) tissue samples were obtained from the Malaysian Oral Cancer Database and Tissue Bank System (MOCDTBS) that coordinated by Oral Cancer Research & Coordinating Centre, University of Malaya (OCRCC-UM). The serum samples of control subjects were obtained from healthy volunteers, while the control FFPE tissue samples were obtained from subjects with the minor oral surgical procedure of impacted wisdom teeth. The serum and FFPE tissue samples are independent of each other. These control subjects did not have a previous history of OSCC or other cancers. The OPMD samples were obtained from patients who were histologically diagnosed as non-dysplasia (oral lichen planus and verrucous hyperplasia) and dysplasia (leukoplakia and erythtroleukoplakia). The OSCC samples included in this study were from patients with a histopathologically confirmed diagnosis of OSCC without any prior treatment. OSCC was staged according to the TNM classification system defined by cancer staging manual in the seventh edition of the American Joint Committee on Cancer (Edge et al., 2010). The serum and tissue samples were collected randomly by OCRCC-UM with written informed consent and ethical approval [Medical Ethics Approval No.: DFOP1504/0084(L) and DFOC1803/0026(P)]

3.2 Study Design

The serum samples from control, OPMD, and OSCC were subjected to protein and glycoprotein profiling using two-dimensional gel electrophoresis (2-DE) and 2-DE coupled with lectin-based approaches, respectively. The images of 2-DE gels and lectin blots were then captured and analysed to determine the differentially expressed protein spots. These differentially expressed protein spots were identified using mass spectrometry analysis. To explore their biological pathway and interaction, the identified proteins were further analysed using bioinformatics software. Subsequently, the proteins

Universiti Malaya

of interest were validated using enzyme-linked immunosorbent assay (ELISA) and immunohistochemistry (IHC). Figure 3.1 illustrates the overview workflow of this study.

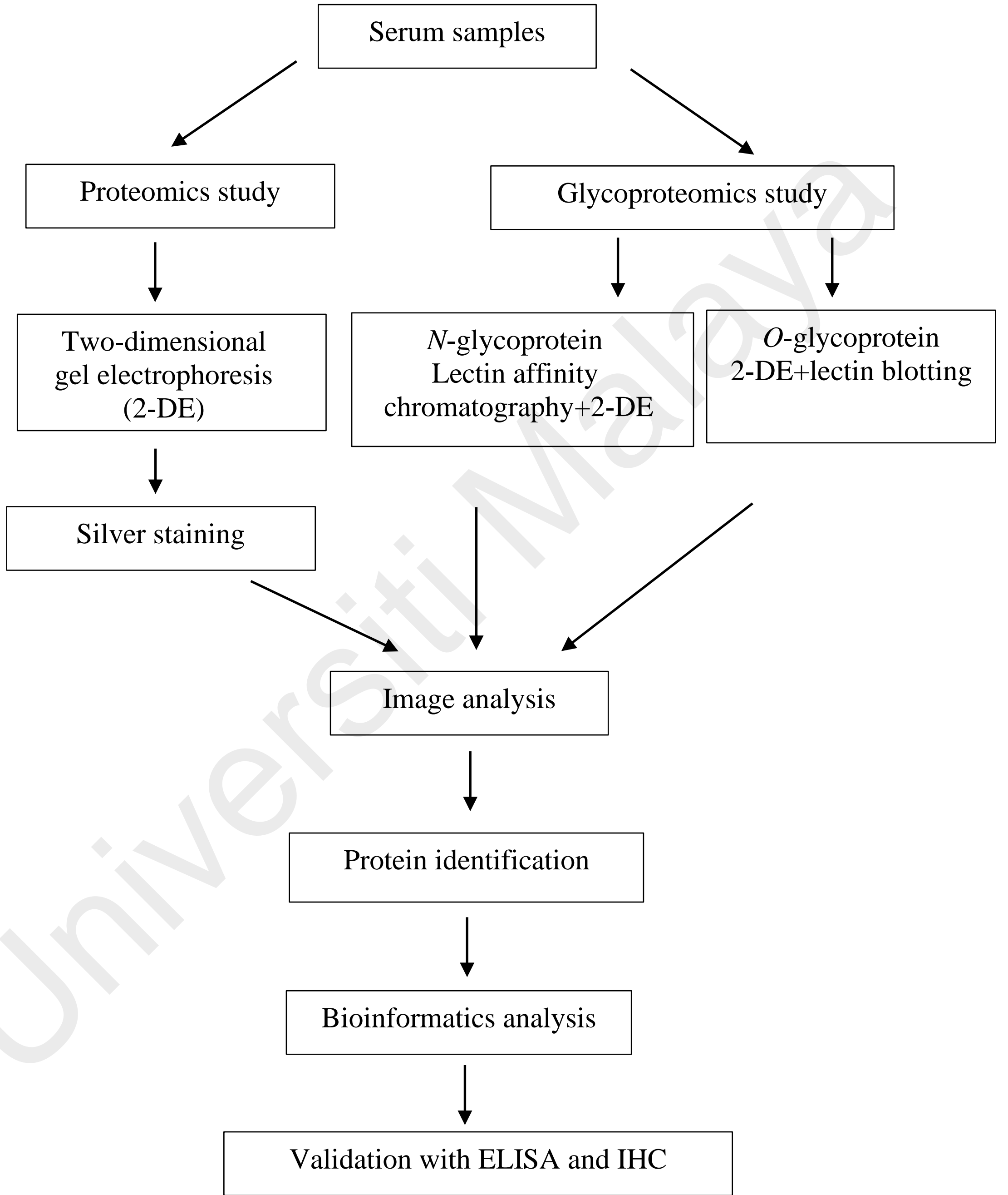


Figure 3.1: Overview workflow of the study

Universiti Malaya

3.3 Two-dimensional Gel Electrophoresis (2-DE)

Two-dimensional gel electrophoresis (2-DE) is a widely used method for the separation of complex protein mixtures. This method separates the proteins in two sequential steps. In the first step, the proteins are separated according to their charges with isoelectric focusing (IEF). While in the second step, the proteins are separated according to their masses by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE).

3.3.1 First-dimensional Separation

A total of 3 µl (approximately 150 µg protein) serum was added into rehydration buffer (8M urea, 2% NP-40, 2% IPG buffer pH4-7, 40 mM DTT, 1% bromophenol blue) with a final volume of 200 µl. The mixture was then added into the slot of the Immobiline DryStrip reswelling tray (GE Healthcare Biosciences, Uppsala, Sweden). The Immobiline DryStrip gel (11cm, pH4-7) or IPG strip (GE Healthcare Biosciences, Uppsala, Sweden) was gently put into the slot with the gel side facing down to cover the mixture. After that, the IPG strip with the mixture was overlaid with 3 ml of Plusone DryStrip cover fluid (GE Healthcare Biosciences, Uppsala, Sweden) to minimize evaporation and urea crystallization. This was followed by overnight incubation at room temperature to allow sample rehydration and uptake into the strip. The first-dimensional separation was performed using Multiphor II Electrophoresis System (GE Healthcare Biosciences, Uppsala, Sweden). The temperature was set at 20ºC throughout the experiment. Approximately 3 to 4 ml of Plusone DryStrip cover fluid was overlaid evenly on Multiphor II platform before placing the Immobiline DryStrip tray on it.

Once the Immobiline DryStrip tray was set properly, the red (anode) and the black (cathode) electrode on the tray were connected to the Multiphor II unit. The rehydrated IPG strips were then removed from the reswelling tray and placed on Immobiline Drystrip

Universiti Malaya

tray across the red electrode (anode) and the black electrode (cathode) ends. Subsequently, moistened electrode strips (GE Healthcare Biosciences, Uppsala, Sweden) were placed on both ends of the rehydrated IPG strip. Later, the electrode assembly was placed on the top of the electrode strips to give electrical contact and ensure the rehydrated IPG strips are still well aligned.

The rehydrated IPG strip was covered with Plusone DryStrip cover fluid to ensure good thermal contact as well as to minimize evaporation and urea crystallization. The running condition of isoelectric focusing (IEF) on the Multiphor II electrophoresis system was performed according to the recommended protocol by the manufacturer (Table 2.1). When IEF was completed, the focused IPG strips were stored in screw-cap tubes individually at -80°C or proceeded immediately to the second-dimensional separation.

Table 3.1: Running condition for IEF on Multiphor II Electrophoresis System

Phase	Voltage (V)	Current (mA)	Power (W)	Voltage/Hour
1	300	2	5	1
2	3500	2	5	2900
3	3500	2	5	9100

3.3.2 Second-dimensional Separation

After completion of IEF, the second-dimensional separation was performed using the SE 600 Ruby system (GE Healthcare Biosciences, Uppsala, Sweden). At first, two glass plates that mounted with side spaces and clamps to a sandwich form were assembled onto the SE 600 gel casting stand according to the manufacturer's manual. A linear gel of 12.5% sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) gel was prepared, with a total volume of approximately 25 ml for each strip (Table 3.2). The use of ammonium persulfate (APS) solution (Merck, Kenilworth, NJ, USA) and N,N,N',N'-tetramethylrthylenediamine (TEMED) (Sigma-Aldrich, St. Loius, MO, USA) is to initiate

Universiti Malaya

the polymerisation. The gel solutions were pipetted onto the glass plate and overlaid with distilled water for the gel to polymerize in a minimum of two hours.

Table 3.2: Preparation of 12.5% SDS-PAGE gel solution

Composition	Volume
30% acrylamide, 0.8% N-N'-methylenebisacrylamide	10.430 ml
4x resolving gel buffer solution (1.5M Tris base, pH8.8)	6.250 ml
10% sodium dodecyl sulphate (SDS) solution	0.250 ml
10% ammonium persulfate (APS) solution*	0.125 ml
N,N,N',N'-tetramethylrthylenediamine (TEMED)*	8.250 µl
Distilled water	Top up
Total Volume	25.000 ml

*Added immediately prior to casting the gel

The equilibration step was performed on the focused IPG strip prior to SDS-PAGE. In the first step of equilibration, the IPG strip was equilibrated with SDS equilibration buffer (75 mM Tris, pH8.8, 6 M urea, 30% v/v glycerol, 2% w/v SDS, 1% bromophenol blue) contained with dithiothreitol (DTT) (10 mg/ml) for 15 minutes. DTT was used to preserve the fully reduced state of denatured and unalkylated proteins. This was followed by the second step of equilibration. The first equilibration solution was discarded, and the IPG strip was equilibrated with SDS equilibration buffer contained with iodoacetamide (IAA) (25 mg/ml) for another 15 minutes. IAA was used to prevent point streaking and other artefacts. All the equilibration steps were performed with gentle shaking on an SSM4 See-saw rocker (Stuart, Staffordshire, UK) at room temperature.

Once the equilibration steps were completed, the IPG strip was quickly rinsed with SDS electrophoresis buffer (25 mM Tris base, 192 mM glycine, 0.1% w/v SDS) to remove excessive SDS equilibration buffer. The distilled water that overlaid on the top of the casted gel was poured away while the SDS gels are kept in the gel caster. The IPG strip was then placed gently down to the gel surface and sealed with melted agarose

Universiti Malaya

sealing solution (25 mM Tris base, 192 mM glycine, 0.1% w/v SDS, 0.5% agarose, 0.002% w/v bromophenol blue). Afterwards, SDS-PAGE was carried out in two phases of running condition at a constant temperature of 16°C (Table 3.3).

Table 3.3: Running condition for SDS-PAGE

Phase	Voltage (V)	Current (mA/gel)	Power (W/gel)	Duration (h:min)
1	50	17	2	0:30
2	600	25	15	2:30

3.3.3 Silver Staining

The silver staining procedure was performed according to a modified silver staining method that is compatible with the subsequent MS analysis (Shevchenko et al., 1996). After completion of SDS-PAGE, the gel was removed gently from the gel cassettes and immersed in fixing solution (40% v/v methanol, 10% v/v acetic acid, glacial, 0.05% v/v formaldehyde) for overnight incubation. The gel incubation and the subsequent staining procedure were carried out on an orbital shaker (IKA KS 4000 i control, IKA, Staufen, Germany) with gentle shaking to have better spreading and uniformity of the stain.

On the following day, the fixing solution was discarded, and the gel was washed 20 minutes with incubation solution (35% v/v ethanol) for three times. The gel was then immersed in sensitizing solution (0.025% sodium thiosulfate solution) for 3 minutes. Later, the gel was washed 5 minutes for three times with distilled water.

After washing, the gel was incubated with silver staining solution (0.2% w/v silver nitrate solution, 0.07% v/v formaldehyde) for 20 minutes. The gel was then quickly rinsed with distilled water for three times to remove excessive silver staining solution. This was followed by adding developing solution (235.9 mM sodium carbonate, 2% v/v sensitizing solution, 0.2% v/v formaldehyde) to the gel. The protein spots were developed and

Universiti Malaya

became visible. Once the development was completed, the developing solution was replaced with stopping solution (40% v/v methanol, 10% v/v acetic acid, glacial) to avoid overstaining on the protein spots. After that, the developed gel was washed with distilled water for two to three times and stored at 4°C.

3.3.4 Image Analysis

After silver staining, the images of 2-DE gels were scanned and captured using ImageQuant LAS 500 (GE Healthcare Biosciences, Uppsala, Sweden). All the gel images were analysed using Progenesis SameSpots v4.0 software (Totallab, Newcastle, UK). The image quality control was done to inspect all the gel images were in the correct configuration and there had no other problems that could interfere with the subsequent image analysis.

In brief, all the gel images were aligned and normalised automatically to allow all spots to be detected and matched on all gels accurately. Once aligned, the gel images were automatically analysed by the software. The gel images were basically separated into two groups for comparative analysis. After analysing background-corrected abundance, the normalised volume of each protein spot was calculated according to the ratio of the target volume and total image volume. Each protein spot of altered abundance was expressed as average normalised volume ± standard error mean (SEM).

Statistical analysis of protein expression level for each protein spot was determined based on average normalised volume and the differences in protein expression between the two groups were assessed by analysis of variance (ANOVA). The data were also checked manually to eliminate any error in matching pairs. Protein spots with a p-value less than 0.05 ($p < 0.05$) and fold change cut-off of 1.2 ($FC \geq 1.2$) were considered as statistically significant. The principal component analysis was performed on all the protein spots of the gels to determine whether any of the protein spots were outliers by

Universiti Malaya

producing a simplified graphical representation of the multidimensional data that clustered according to the group condition.

3.4 Glycoproteomics Study

Glycoproteomics study was divided into *N*-glycoprotein and *O*-glycoprotein studies. *N*-glycoprotein study was carried out using Concanavalin A (Con A) affinity chromatography coupled with 2-DE and silver staining, while *O*-glycoprotein study was performed using 2-DE coupled with Jacalin lectin blotting.

3.4.1 *N*-glycoprotein Study

Concanavalin A (Con A) lectin coupled to Sepharose 4B was used for the separation and purification of *N*-glycoproteins. Prior to chromatographic procedure, a total volume of 2 ml Con A Sepharose 4B (GE Healthcare Biosciences, Uppsala, Sweden) is pre-equilibrated with binding buffer (20 mM Tris-HCl pH7.4, 0.5 M NaCl, 1 mM $MnCl_2$, 1 mM $CaCl_2$) into conical polypropylene chromatography column (BioRad Laboratories, Hercules, CA, USA), and the column was washed with 20 ml of binding buffer to remove preservatives. After that, a total of 10 µl of serum sample was loaded onto the column with packed Con A sepharose 4B and incubated overnight on an SSM4 See-saw rocker (Stuart, Staffordshire, UK) at 4°C.

On the following day, the mixtures in the column were equilibrated with equilibration buffer (20 mM Tris-HCl pH7.4, 0.5 M NaCl, 1 mM $MnCl_2$, 1 mM $CaCl_2$). The column was then washed with 20 ml of equilibration buffer until no proteins appeared in the eluent (monitored by absorbance at 280 nm). The unbound fraction of proteins enriched with non-glycosylated and *O*-glycosylated proteins was washed away. The bound *N*-glycoproteins were eluted with elution buffer (0.3 M methyl-α-D-glucopyranoside, 20 mM Tris-HCl pH7.4, 0.5 M NaCl) based on gravity flow at 2 to 8°C. Each fraction of the eluent was then collected and monitored for its absorbance at 280 nm by Nanodrop

Universiti Malaya

spectrophotometer (ND-2000, NanoDrop Technologies, Rockland, Del, USA) to determine the protein concentration in the eluent along with the chromatographic process.

Subsequently, the eluted fractions were concentrated using Vivaspin Turbo 15 with a molecular weight cut-off of 10,000 (Sartorius Stedim Biotech GmbH, Goettingen, Germany) at 3,000 x g using a refrigerated centrifuge, Tomy MX-305 (Tomy Seiko, Tokyo, Japan). The protein concentrations of the concentrated fraction were then measured using the Pierce 660 nm protein assay (Thermo Fisher Scientific, Waltham, MA, USA). Once the concentrated fraction is confirmed containing protein, the concentrated fraction was further desalted using a 2-D clean-up kit (GE Healthcare BioSciences, Uppsala, Sweden). The proteins were pelleted by centrifugation at maximum speed using a refrigerated centrifuge, Eppendorf centrifuge 5415R (Eppendorf, Hamburg, Germany). The precipitate was washed to remove non-protein contaminants. The mixture was centrifuged again at maximum speed. The resultant pellet was resuspended into rehydration buffer and subjected to 2-DE as described in Section 3.3. After 2-DE and silver staining, the gels were scanned and captured using ImageQuant LAS 500 (GE Healthcare Biosciences, Uppsala, Sweden) for further image analysis.

3.4.2 *O*-glycoprotein Study

The *O*-glycoprotein study was performed using 2-DE coupled with Jacalin lectin blotting. A total of 3 µl serum sample was subjected to 2-DE as described in Section 3.3. Instead of silver staining, the 2-DE gel was equilibrated with transfer buffer (25 mM Tris base pH8.3, 192 mM glycine, 0.1% w/v SDS, 20% v/v methanol) for 15 minutes on an orbital shaker (IKA KS 4000 i control, IKA, Staufen, Germany) with gentle shaking. The 2-DE gel was then transferred onto nitrocellulose membrane, 0.45 µm (Thermo Fisher Scientific, Waltham, MA, USA) using Trans-Blot turbo blotting system (BioRad Laboratories, Hercules, CA, USA) at 15 V for 15 minutes. After transferred, the

Universiti Malaya

membrane was placed in SuperBlock blocking buffer (Thermo Fisher Scientific, Waltham, MA, USA) for another 15 minutes on an orbital shaker (IKA KS 4000 i control, IKA, Staufen, Germany) with gentle shaking at room temperature to block non-specific protein binding sites. After that, the membrane was washed 5 minutes for three times using Tris-buffered saline, pH7.4 with 0.1% Tween 20 (TBST). This was followed by overnight incubation with 20 μg/ml of biotinylated Jacalin (Vector Laboratories, Burlingame, CA, USA) at 4°C.

On the following day, the membrane was washed 5 minutes for three times with TBST. After completely washing, the membrane was incubated with 10 μg/ml of horseradish peroxidase streptavidin (Vector Laboratories, Burlingame, CA, USA) for an hour at room temperature. Later, the membrane was developed using chemiluminescent substrates, ECL Select Western blotting detection reagent (GE Healthcare Biosciences, Uppsala, Sweden) according to manufacturer's protocol. The membrane image was captured with ImageQuantLAS 500 (GE Healthcare Biosciences, Uppsala, Sweden) for further image analysis.

3.5 Mass Spectrometry (MS) Analysis

3.5.1 In-gel Trypsin Digestion

Prior to perform protein identification, the protein spots of interest were excised and subjected to in-gel trypsin digestion. Briefly, the excised gel plugs were destained with 100 µl of destaining buffer (15 mM potassium ferricyanide in 50 mM sodium thiosulfate) for 15 minutes with gentle shaking until all the stains become transparent. The destaining buffer was then discarded and the gel plugs were incubated with 150 µl of reducing solution (10 mM DTT in 100 mM ammonium bicarbonate) for 30 minutes at 60°C in a thermomixer (Eppendorf, Hamburg, Germany). After that, the reducing solution was

Universiti Malaya

discarded and followed by incubation with 150 µl of alkylation solution (55 mM IAA in 100 mM ammonium bicarbonate) for 20 minutes in the dark at room temperature.

Subsequently, the gel plugs were washed three times with 500 µl of 50% acetonitrile in 100 mM ammonium bicarbonate for 20 minutes with gentle shaking. The gel plugs were dehydrated by adding 50 µl of 100% acetonitrile and incubate for 15 minutes with gentle shaking. The gel plugs were then air dried for 5 to 10 minutes at room temperature after removing the supernatant carefully. After that, 25 µl of digestion solution (6 ng/µl trypsin in 50 mM ammonium bicarbonate) was added into the dried gel plugs for overnight incubation at 37°C in a thermomixer (Eppendorf, Hamburg, Germany).

The peptides from the gel plugs were extracted with 50% acetonitrile and 100% acetonitrile on the following day. A total of 50 µl of 50% acetonitrile was added into the gel plugs and incubated for 15 minutes with gentle shaking. All the peptide mixtures were carefully transferred into a new microcentrifuge tube. This was followed by adding 50 µl of 100% acetonitrile into the gel plugs and incubated for 15 minutes with gentle shaking. The peptide mixtures were transferred into the new microcentrifuge tube and concentrated in a vacuum concentrator (Eppendorf concentrator plus, Eppendorf, Hamburg, Germany). The tubes were stored at -20°C until further use for mass spectrometry analysis.

3.5.2 Sample Preparation for Mass Spectrometry

The trypsinised peptides were reconstituted with 0.1% formic acid solution and desalted using ZipTip pipette tip that contains C_{18} reversed-phase media (Millipore, Massachusetts, USA). The ZipTip pipette tip was pre-wetted and equilibrated with 10 µl of 50% acetonitrile and 10 µl of 0.1% formic acid solution, respectively. The reconstituted peptides were then carefully aspirated and dispensed into the ZipTip pipette tip for maximum binding of complex mixtures. After washing three times with 10 µl of 0.1%

Universiti Malaya

formic acid solution, the peptides that trapped inside the ZipTip pipette tip were dispensed into 3 µl of elution solution (0.1% formic acid in 50% acetonitrile).

Subsequently, the eluted peptides were mixed with 3 µl of α-cyano-4-hydroxy-cinnamic acid (CHCA) matrix solution (6 mg/ml α-cyano-4-hydroxy-cinnamic acid, 70% acetonitrile, 0.1% v/v trifluoroacetic acid aqueous) and spotted onto a 384-well MALDI target plate (AB Sciex, Foster City, CA, USA). The mixtures were air dried completely at room temperature. The dried mixtures were then analysed using the 4800 Plus MALDI-TOF/TOF Analyser (AB Sciex, Framingham, MA, USA). The MS results were acquired automatically with a trypsin auto-digest exclusion list and the 20 most intense precursor ions were selected for tandem mass spectrometry (MS/MS) analysis with a minimum signal-to-noise (S/N) of at least 10.

3.5.3 Protein Identification

For protein identification, mass spectra of the protein spots were carried out using the MASCOT search engine (Matrix Science, London, UK) against *Home sapiens* species in the UniProtKB database (last updated: December 2016). Searches were performed with fixed modification on carbamidomethylation of cysteines and variable modification of methionine oxidation. The following parameters were used in the MASCOT peptide mass fingerprint and ion search: a) enzyme trypsin is used, b) one missed cleavage is allowed, c) mass value: monoisotopic is included, d) peptide mass tolerance at 100 ppm, and e) fragment mass tolerance: ±0.2 Da. According to MASCOT search results, protein scores greater than 50 were considered significant ($p < 0.05$).

Universiti Malaya

3.6 Bioinformatics Analysis

The identified proteins were further determined using Database for Annotation, Visualization, and Integrated Discovery (DAVID) v6.8 software (https://david.ncifcrf.gov/home.jsp) for functional annotation analysis. DAVID bioinformatics resources provide a comprehensive functional algorithm with gene ontology terms to explore the biological meaning of interesting gene or protein lists (Huang da, Sherman, & Lempicki, 2009a, 2009b). Moreover, the identified proteins were analysed using Ingenuity Pathway Analysis (IPA) v7.1 (Qiagen Ingenuity System) software (https://www.ingenuity.com) to determine the pathway annotations and interaction networks. The interaction networks and predominant canonical pathways of the proteins were generated algorithmically by using the Ingenuity pathways knowledge base.

Additionally, *N*- and *O*-glycoprotein prediction analysis was performed to predict potential glycosylation sites of the identified glycoproteins. *N*-glycosylation sites were predicted using the NetNGlyc 1.0 server (http://www.cbs.dtu.dk/services/NetNGlyc), whereas *O*-glycosylation sites were predicted using the NetOGlyc 4.0 server (http://www.cbs.dtu.dk/services/NetOGlyc) (Gupta & Brunak, 2002; Steentoft et al., 2013). These servers are employed algorithms based on the neural networks to predict *N*-glycosylation and mucin-type-*O*-glycosylation sites, respectively. Only the sites with thresholds higher than 0.5 are predicted as glycosylated sites.

Furthermore, the proteins of interest were selected from the list of identified proteins for enzyme-linked immunosorbent assay (ELISA) and immunohistochemistry (IHC) analyses. The selection criteria of these proteins for the validation study are based on the protein score in MS analysis with a value of greater than 50 ($p < 0.05$), the protein role and interaction in the bioinformatics analysis as well as the literature searches.

Universiti Malaya

3.7 Enzyme-linked Immunosorbent Assay (ELISA)

A quantitative sandwich ELISA was used to validate the proteins of interest in this study. The protein expression levels of α1-antitrypsin (AAT) (Thermofisher Scientific, Frederick, MD, USA), α2-HS-glycoprotein (AHSG) (Thermofisher Scientific, Frederick, MD, USA), apolipoprotein A-I (APOA1) (Cusabio Biotech, Wuhan, China), clusterin (CLU) (Cusabio Biotech, Wuhan, China), and haptoglobin (HP) (Cusabio Biotech, Wuhan, China) were detected according to the manufacturer's protocol. In brief, 100 µl of standard, blank, and diluted samples were added into a 96-well plate and incubated for 1 to 2.5 hours. A total of 100 µl biotinylated antibody was added into each well and incubated for 1 hour.

After washing, the streptavidin-HRP antibody was added into each well and incubated for 45 minutes to 1 hour. This was followed by colour development with the addition of 3,3′,5,5′-tetramethylbenzidine (TMB) substrate into each well and incubated for 20 to 30 minutes in the dark after the final washing to remove any unbound substances. The reaction was stopped with 50 µl stop solution and the plate was gently tapped to ensure thorough mixing. After that, the optical density of each well was measure at 450 nm using MQX200 µQuant microplate reader (BioTek, Winooski, VT, USA). The specific condition of each ELISA kit is listed in Table 3.4.

Standard curves with a four-parameter logistic (4-PL) curve-fit were generated using Curve Expert software version 1.4 (Hyams D.G., Starkville, MS, USA). These standard curves were used to determine the concentration of target proteins in each sample. The level of protein expression between different groups of ELISA assays was analysed and compared using SPSS 20.0 statistical software package in which p-values less than 0.05 ($p < 0.05$) were considered statistically significant.

Universiti Malaya

Table 3.4: Specific condition of ELISA kit

	AAT	AHSG	APOA1	CLU	HP
Sample dilution factor	1:10	1:50,000	1:8,000	1:5,000	1:20,000
Standards/samples incubation	2.5 hours at RT with shaking condition	2.5 hours at RT with shaking condition	1 hour at 37ºC	2 hours at 37ºC	1 hour at 37ºC
Washing	4 times	4 times	-	-	-
Biotinylated antibody incubation	1 hour at RT with shaking condition	1 hour at RT with shaking condition	-	1 hour at 37ºC	-
Washing	4 times	4 times	3 times	3 times	3 times
Streptavidin-HRP incubation	45 min at RT with shaking condition	45 min at RT with shaking condition	1 hour at 37ºC	1 hour at 37ºC	1 hour at 37ºC
Washing	4 times	4 times	5 times	5 times	5 times
TMB substrate incubation	30 min at RT with shaking condition	30 min at RT with shaking condition	20 min at 37ºC	20 min at 37ºC	20 min at 37ºC
Optical density	450 nm	450 nm	450 nm	450 nm	450 nm

RT: room temperature

Universiti Malaya

3.8 Immunohistochemistry (IHC)

Immunohistochemistry (IHC) was performed on 4 µm thickness of formalin-fixed paraffin-embedded (FFPE) tissue sections. Briefly, FFPE tissue sections were deparaffined in xylene and rehydrated in decreasing the concentration of ethanol (100%, 95%, and 70%). Antigen was retrieved using 10 mM sodium citrate buffer (pH6.0) or target retrieval antigen buffer (pH9.0, Dako, Agilent, Santa Clara, CA, USA) for 20 minutes at 95°C in a water bath (Julabo SW22, Julabo Labortechnik GmbH, Seelbach, Germany). The sections were then cooled down at room temperature.

Subsequently, endogenous peroxidase activity was inhibited with the peroxidase-blocking solution (Dako, Agilent, Santa Clara, CA, USA) for 5 minutes at room temperature. The sections were then washed 5 minutes with Tris-buffered saline solution pH7.6 containing 0.05% Tween 20 (TBST) for two times at room temperature. Non-specific background staining was blocked with background sniper (Biocare Medical, Concord, CA, USA) for 10 minutes at room temperature. Again, the sections were washed 5 minutes with TBST for two times.

After washing, the sections were incubated with the primary antibodies for one hour at room temperature or overnight at 4°C. The primary antibodies used for IHC were AAT (rabbit monoclonal antibody, ab207303, Abcam, Cambridge, MA, USA), AHSG (mouse monoclonal antibody, ab89227, Abcam, Cambridge, MA, USA), APOA1 (rabbit monoclonal antibody, ab52945, Abcam, Cambridge, MA, USA), CLU (rabbit monoclonal antibody, ab92458, Abcam, Cambridge, MA, USA), and HP (rabbit polyclonal antibody, ab23100, Abcam, Cambridge, MA, USA). The sections were then washed 5 minutes with TBST for three times. This was followed by incubation of the secondary antibody, Dako Real EnVision streptavidin horseradish peroxidase (HRP) solution (Dako, Agilent, Santa Clara, CA, USA) for 30 minutes at room temperature.

Universiti Malaya

The immunoreaction was visualized with 3,3'-diaminobenzidine solution (DAB) chromogen substrate (Dako, Agilent, Santa Clara, CA, USA) for 10 minutes at room temperature after TBST washing. The sections were counterstained with Harris haematoxylin (Leica Biosystems, Concord, ON, Canada), dehydrated and mounted with xylene based medium (Thermo Shandon, Waltham, MA, USA). Both positive and negative controls were included in IHC to confirm the validity and non-reactivity of the antibody, respectively. The specific condition of each primary antibody that applied during IHC was listed in Table 3.5.

The images of immuno-stained slides were digitalized using the Pannoramic Desk slide scanner (3D Histech, Budapest, Hungary). The images were viewed using Pannoramic Viewer v1.15.3 (3D Histech, Budapest, Hungary) for a semi-quantitatively scoring assessment. The scoring assessment was performed by two independent observers who are blinded to the clinical data. Any discrepancies in the scoring assessment were discussed to achieve a consensus for a derivative of the final scores.

The intensity scores were quantified as: negative = 0, weak = 1, moderate = 2, and strong = 3. The percentage of immuno-positive staining were quantified as: 0 = ≤ 10%, 1 = 11-25%, 2 = 26-50%, 3 = 51-75%, and 4 = > 75%. The final immuno-reactive score was determined by multiplying the intensity and percentage of immuno-positive staining scores to obtain the immuno-reactive score ranging from 0 to 12 (Wang et al., 2018; Zhao, Ye, et al., 2015). Statistical analysis was performed using SPSS 20.0 statistical software package (SPSS Inc, Chicago, IL, USA) where p-values less than 0.05 ($p < 0.05$) were considered statistically significant.

Universiti Malaya

Table 3.5: Specific condition of primary antibodies

	AAT	AHSG	APOA1	CLU	HP
Antigen retrieval buffer	target retrieval antigen buffer (pH9.0)	target retrieval antigen buffer (pH9.0)	sodium citrate buffer (pH6.0)	target retrieval antigen buffer (pH9.0)	sodium citrate buffer (pH6.0)
Dilution factor	1:800	1:250	1:1,000	1:400	5 µg/ml
Incubation period	1 hour at RT	4°C at overnight	1 hour at RT	1 hour at RT	1 hour at RT
Positive control tissue	liver	placenta	liver	tonsil	liver

RT: room temperature

Universiti Malaya

CHAPTER 4: RESULTS

4.1 Study Subjects

A total of 60 serum samples from normal healthy volunteers (n = 10) as control and patients diagnosed with OPMD (n = 10), early stage (n = 20), and advanced stage (n = 20) of OSCC during the period from 2004 to 2011 were included in this study for protein and glycoprotein profiling. In this study, early stage is defined as stages I and II, whereas advanced stage is defined as stages III and IV based on the seventh edition of AJCC TNM staging. All the samples were obtained from MOCDTBS at OCRCC-UM. Most of the samples were in the older age group (45 years of age and over). The mean age of control, OPMD, early OSCC, and advanced OSCC were 55.70±16.21, 53.80±14.87, 57.90±15.76, and 63.10±12.10, respectively. Based on the demographic data, most of the OPMD and OSCC patients were females, where Indians were the highest ethnicity diagnosed with OSCC. In this study, the betel quid chewing habit had a higher percentage in OPMD and OSCC patients compared with the control. Buccal mucosa, gingiva, lip, and palate were the most common sites among all OPMD and OSCC cases. The demographic characteristics of the study samples are shown in Table 4.1.

Universiti Malaya

Table 4.1: Demographic characteristics of study samples

	Control (n = 10) (%)	OPMD (n = 10) (%)	Early OSCC (n = 20) (%)	Advanced OSCC (n = 20) (%)
Age (years)	55.70±16.21	53.80±14.87	57.90±15.76	63.10±12.10
< 45	3 (30)	1 (10)	3 (15)	2 (10)
≥ 45	7 (70)	9 (90)	17 (85)	18 (90)
Gender				
Male	5 (50)	1 (10)	5 (15)	6 (30)
Female	5 (50)	9 (90)	15 (75)	14 (70)
Ethnic group				
Malay, Chinese & Others	10 (100)	6 (60)	10 (50)	8 (40)
Indian	0 (0)	4 (40)	10 (50)	12 (60)
Smoking				
No	6 (60)	9 (90)	15 (75)	16 (80)
Yes	4 (40)	1 (10)	5 (15)	4 (20)
Alcohol drinking				
No	9 (90)	9 (90)	14 (70)	18 (90)
Yes	1 (10)	1 (10)	6 (30)	2 (10)
Betel quid chewing				
No	9 (90)	6 (60)	13 (65)	10 (50)
Yes	1 (10)	4 (40)	7 (35)	10 (50)
Site				
Tongue, floor of mouth	-	1 (10.0)	10 (50.0)	8 (40.0)
Buccal mucosa, gingiva, lip & palate	-	9 (90.0)	10 (50.0)	12 (60.0)

Universiti Malaya

4.2 2-DE Gel Analysis

4.2.1 Image Analysis of 2-DE Gel

In this study, unfractionated serum samples of control, OPMD, early OSCC, and advanced OSCC were subjected to 2-DE to separate the protein spots based on different isoelectric points and relative molecular masses. Hundreds of spots have been revealed on the 2-DE protein profiles. The 2-DE protein profiles were generally comparable with the SWISS-2D PAGE reference map because the protein profiles were preliminarily identified using the SWISS-2D PAGE database (Sanchez et al., 1995).

A total of 21 proteins were selected for analysis as they were well-resolved in the 2-DE gels. These proteins were included α1-antitrypsin (AAT), α2-HS-glycoprotein (AHSG), alpha-internexin (AINX), albumin (ALB), alpha-1-microglobulin/bikunin precursor (AMBP), apolipoprotein A1 (APOA1), tumour suppressor ARF (ARF), complement C3 (C3), clusterin (CLU), haptoglobin (HP), Ig alpha-2 chain region (IGHA2), Ig gamma-2 chain C region (IGHG2), Ig kappa chain C region (IGKC), leucine-rich alpha-2-glycoprotein (LRG1), peroxiredoxin-2 (PRDX2), DNA repair protein RAD50 (RAD50), retinol-binding protein 4 (RBP4), serum amyloid P-component (SAMP), serotransferrin (TF), vinculin (VCL), and vitamin D-binding protein (VDBP) (Figure 4.1). Among these proteins, AHSG, AAT, ALB, APOA1, HP, and IGKC appeared as clusters of protein spots in the protein profiles. MALDI-TOF/TOF with MASCOT search using UniProtKB database was used to further identify these proteins. The details of the identified proteins including the UniProtKB accession number, theoretical molecular mass, theoretical isoelectric point (pI), MASCOT score, and percentage of sequence coverage of the identified proteins in the MS analysis are shown in Table 4.2. Protein scores greater than 50 were indicated as significant ($p < 0.05$). The mass spectra of all identified proteins are shown in Appendix B.

Universiti Malaya

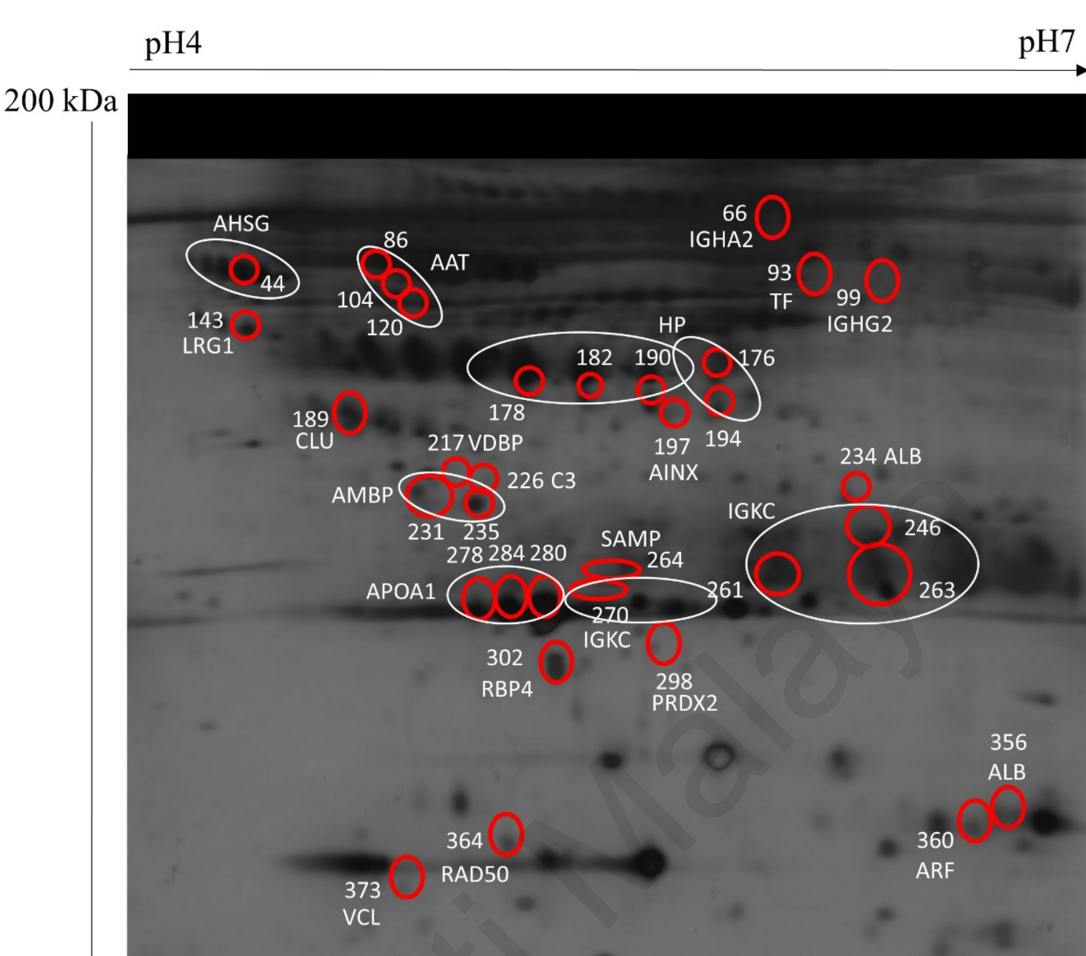


Figure 4.1: Typical representative of serum protein profile using 2-DE and silver staining

Unfractionated serum samples were subjected to 2-DE and silver staining. The individual protein spots were circled in red colour, whereas the clusters of protein spots were circled in white colour.

Universiti Malaya

Table 4.2: MS identification of serum proteins

Protein	Spot ID	UniprotKB Accession No.	Theoretical Mass (kDa)/pI	Sequence Coverage (%)	Mascot Score
AAT	86	P01009	46.88/5.37	38	236
	104	P01009	46.88/5.37	44	70
	120	P01009	46.88/5.37	29	82
AINX	197	Q16352	55.53/5.34	41	48
AHSG	44	P02765	40.10/5.43	31	137
ALB	234	P02768	71.32/5.92	24	124
	356	P02768	71.32/5.92	9	56
AMBP	231	P02760	39.89/5.95	29	85
	235	P02760	39.89/5.95	28	182
APOA1	278	P02647	30.76/5.56	48	249
	280	P02647	30.76/5.56	62	572
	284	P02647	30.76/5.56	46	208
ARF	360	Q8N726	14.95/12.41	21	17
C3	226	P01024	188.57/6.02	12	251
CLU	189	P10909	53.03/5.89	17	128
HP	176	P00738	45.86/6.13	33	225
	178	P00738	45.86/6.13	11	41
	182	P00738	45.86/6.13	21	293
	190	P00738	45.86/6.13	20	177
	194	P00738	45.86/6.13	20	69
IGHA2	66	P01877	37.30/5.71	12	72
IGHG2	99	P01859	36.51/7.65	17	60
IGKC	246	P01834	11.77/5.58	15	41
	261	P01834	11.77/5.58	54	169
	263	P01834	11.77/5.58	32	77
	270	P01834	11.77/5.58	32	93
LRG1	143	P02750	38.38/6.45	23	141
PRDX2	298	P32119	22.05/5.66	60	321
RAD50	364	Q92878	154.82/6.48	24	42
RBP4	302	P02753	23.34/5/76	47	346
SAMP	264	P02743	25.49/6.10	14	79
TF	93	P02787	79.29/7.66	14	71
VCL	373	P18206	124.29/5.50	32	43
VDBP	217	P02774	54.53/5.40	14	72

AAT: α1-antitrypsin, AINX: alpha-internexin, AHSG: α2-HS-glycoprotein, ALB: albumin, AMBP: alpha-1-microglobulin/bikunin precursor, APOA1: apolipoprotein A-I, ARF: tumour suppressor ARF, C3: complement C3, CLU: clusterin, HP: haptoglobin, IGHA2: Ig alpha-2 chain C region, IGHG2: Ig gamma-2 chain C region, IGKC: Ig kappa chain C region, LRG1: leucine-rich alpha-2-glycoprotein, PRDX2: peroxiredoxin-2, RAD50: DNA repair protein RAD50, RBP4: retinol-binding protein 4, SAMP: serum amyloid P-component, TF: serotransferrin, VCL: vinculin, VDBP: vitamin D-binding protein

Universiti Malaya

4.2.2 2-DE Protein Profile Analysis

More than hundreds of protein spots in each 2-DE gel from control, OPMD, early OSCC, and advanced OSCC were detected and analysed using Progenesis SameSpot v4.0 software. This comparative analysis revealed differentially expressed protein spots with p-values less than 0.05 ($p < 0.05$) and fold change cut-off of 1.2 ($FC \geq 1.2$). Each statistically significant spot was manually examined to ensure each spot was not an artefact and it can be excised from the gel for MS analysis. The typical representative 2-DE serum protein profiles of control, OPMD, early OSCC, and advanced OSCC are illustrated in Figure 4.2.

Universiti Malaya

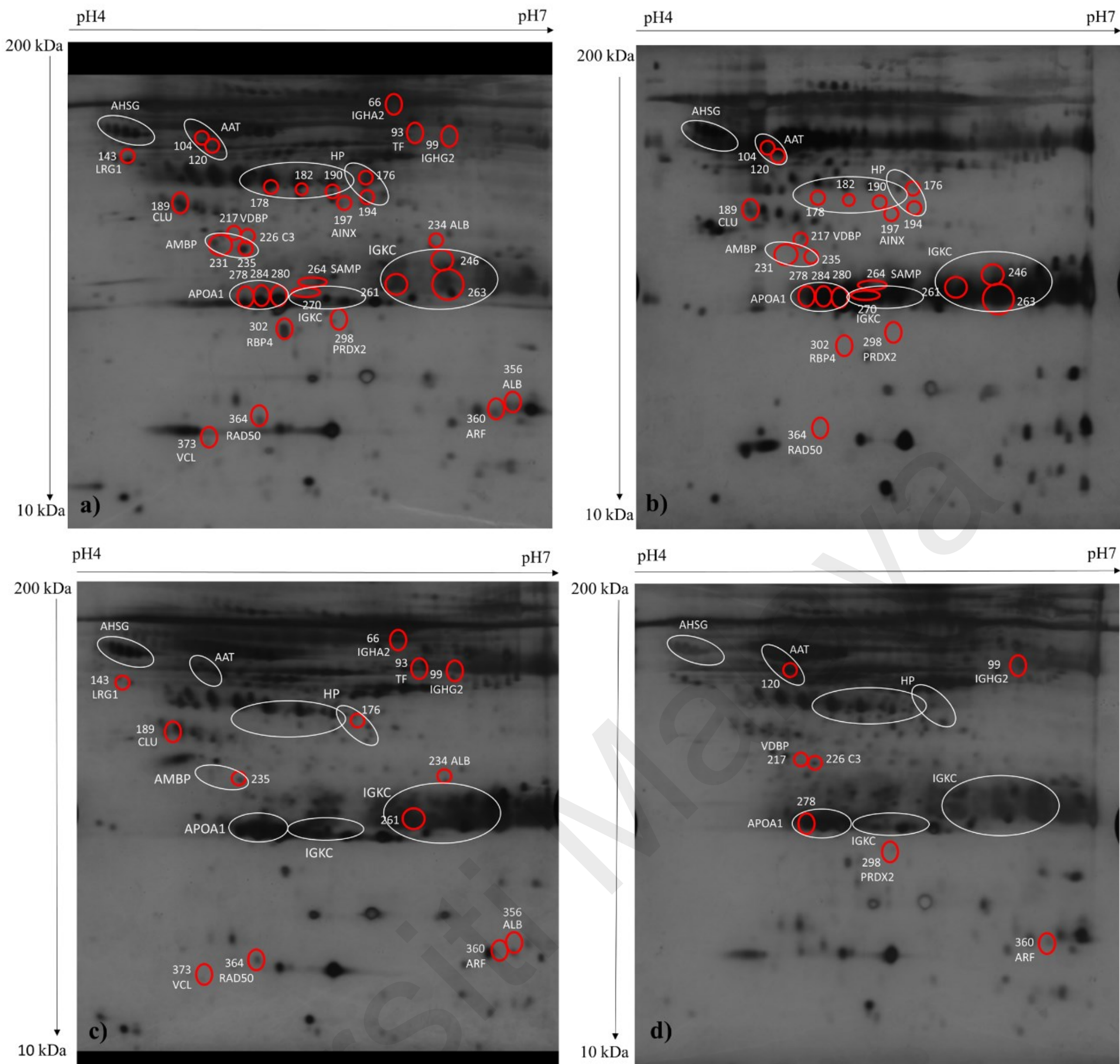


Figure 4.2: Typical representative 2-DE serum protein profiles of (a) control, (b) OPMD, (c) early OSCC, and (d) advanced OSCC

Unfractionated serum samples of control, OPMD, early OSCC, and advanced OSCC were subjected to 2-DE and silver staining. The individual protein spots were circled in red colour, whereas the clusters of protein spots were circled in white colour.

Universiti Malaya

A total of 23 protein spots that differentially expressed in OPMD compared with the control was detected using Progenesis SameSpot v4.0 software. There were 5 up-regulated proteins (AAT, APOA1, IGKC, SAMP, and VDBP) and 7 down-regulated proteins (AINX, AMBP, CLU, HP, PRDX2, RAD50, and RBP4) detected as differentially expressed proteins in OPMD (Table 4.3).

Table 4.3: Differentially expressed protein spots in OPMD compared with control

No.	Spot ID	Protein	Anova p-value	Fold change	Average Normalised Volumes	
					Normal	OPMD
1	284	APOA1	1.354e-08	2.0 (up)	6.653e+007	1.358e+008
2	280	APOA1	7.954e-08	1.9 (up)	2.615e+007	5.049e+007
3	278	APOA1	1.344e-07	2.8 (up)	2.263e+007	6.293e+007
4	178	HP	1.004e-06	2.5 (down)	6.153e+007	2.426e+007
5	182	HP	4.499e-06	2.8 (down)	9.464e+006	3.419e+006
6	194	HP	5.335e-06	2.5 (down)	8.475e+006	3.390e+006
7	190	HP	5.560e-06	3.4 (down)	1.347e+007	3.927e+006
8	176	HP	1.118e-05	2.6 (down)	4.767e+006	1.861e+006
9	263	IGKC	7.773e-06	1.6 (up)	9.751e+007	1.596e+008
10	246	IGKC	4.417e-05	1.7 (up)	3.156e+007	5.423e+007
11	270	IGKC	1.188e-04	2.3 (up)	9.820e+006	2.262e+007
12	261	IGKC	1.813e-04	2.0 (up)	2.942e+007	6.009e+007
13	235	AMBP	1.300e-05	2.9 (down)	6.882e+006	2.353e+006
14	231	AMBP	2.922e-05	2.1 (down)	6.008e+006	2.838e+006
15	197	AINX	2.054e-05	2.4 (down)	4.310e+006	1.827e+006
16	120	AAT	2.138e-05	1.9 (up)	1.066e+007	1.978e+007
17	104	AAT	6.366e-05	2.2 (up)	4.345e+006	9.344e+006
18	302	RBP4	3.642e-05	2.5 (down)	1.118e+007	4.528e+006
19	264	SAMP	8.523e-05	2.0 (up)	1.860e+007	3.648e+007
20	298	PRDX2	8.793e-05	2.3 (down)	3.019e+006	1.327e+006
21	364	RAD50	0.006	2.0 (down)	9.245e+006	4.583e+006
22	189	CLU	0.007	1.8 (down)	1.216e+007	6.281e+006
23	217	VDBP	0.034	2.3 (up)	1.050e+006	2.428e+006

Universiti Malaya

In early OSCC, a total of 13 protein spots were differentially expressed compared with control. There were 4 up-regulated proteins (IGKC, IGHA2, IGHG2, and TF) and 8 down-regulated proteins (ALB, AMBP, ARF, CLU, HP, LRG1, RAD50, and VCL) detected as differentially expressed proteins in early OSCC (Table 4.4). Whereas, a total of 7 protein spots were differentially expressed in advanced OSCC compared with the control. There were 5 up-regulated proteins (AAT, APOA1, C3, IGHG2, and VDBP) and 2 down-regulated proteins (ARF and PRDX2) detected as differentially expressed proteins in advanced OSCC (Table 4.5).

Table 4.4: Differentially expressed protein spots in early OSCC compared with control

No.	Spot ID	Protein	Anova p-value	Fold change	Average Normalised Volumes	
					Normal	Early OSCC
1	356	ALB	6.285e-05	1.9 (down)	1.550e+007	8.214e+006
2	234	ALB	0.006	1.6 (down)	3.772e+006	2.370e+006
3	261	IGKC	1.158e-04	1.7 (up)	2.942e+007	5.602e+007
4	360	ARF	2.658e-04	2.0 (down)	5.808e+006	2.833e+006
5	143	LRG1	2.792e-04	1.6 (down)	1.393e+007	8.755e+006
6	364	RAD50	3.609e-04	2.1 (down)	9.245e+006	4.378e+006
7	235	AMBP	5.606e-04	1.9 (down)	6.822e+006	3.546e+006
8	176	HP	0.003	1.6 (down)	4.767e+006	3.009e+006
9	189	CLU	0.004	1.6 (down)	1.216e+007	7.638e+006
10	99	IGHG2	0.006	1.8 (up)	9.727e+006	1.794e+007
11	373	VCL	0.013	1.9 (down)	3.254e+006	1.675e+006
12	93	TF	0.040	1.9 (up)	4.570e+006	8.468e+006
13	66	IGHA2	0.043	1.7 (up)	8.554e+006	1.448e+007

Universiti Malaya

Table 4.5: Differentially expressed protein spots in advanced OSCC compared with control

No.	Spot ID	Protein	Anova p-value	Fold change	Average Normalised Volumes	
					Normal	Advanced OSCC
1	120	AAT	4.447e-06	1.5 (up)	1.066e+007	1.632e+007
2	298	PRDX2	6.097e-04	1.8 (down)	3.019e+006	1.662e+006
3	226	C3	0.004	1.5 (up)	1.788e+006	2.682e+006
4	278	APOA1	0.005	1.6 (up)	2.263e+007	3.648e+007
5	99	IGHG2	0.011	1.6 (up)	9.727e+006	1.513e+007
6	360	ARF	0.014	1.6 (down)	5.808e+006	3.643e+006
7	217	VDBP	0.027	1.9 (up)	1.050e+006	1.968e+006

For enhanced visualisation, each differentially expressed protein spot that significantly detected in OPMD (Table 4.3), early OSCC (Table 4.4), and advanced OSCC (Table 4.5) were compared individually with the control. These spots were arranged according to the respective protein name (Table 4.6).

Table 4.6: Differentially expressed protein spots in OPMD, early OSCC, and advanced OSCC compared with control

No.	Spot ID	Protein	Control	OPMD	Early OSCC	Advanced OSCC
1	120	AAT				
2	104	AAT				
3	197	AINX				
4	356	ALB				

Universiti Malaya

Table 4.6 continued

No.	Spot ID	Protein	Control	OPMD	Early OSCC	Advanced OSCC
5	234	ALB				
6	235	AMBP				
7	231	AMBP				
8	284	APOA1				
9	280	APOA1				
10	278	APOA1				
11	360	ARF				
12	226	C3				
13	189	CLU				
14	178	HP				
15	182	HP				
16	194	HP				

University of Malaya

Table 4.6 continued

No.	Spot ID	Protein	Control	OPMD	Early OSCC	Advanced OSCC
17	190	HP				
18	176	HP				
19	66	IGHA2				
20	99	IGHG2				
21	263	IGKC				
22	246	IGKC				
23	270	IGKC				
24	261	IGKC				
25	143	LRG1				
26	298	PRDX2				
27	364	RAD50				
28	302	RBP4				

Universiti Malaya

Table 4.6 continued

No.	Spot ID	Protein	Control	OPMD	Early OSCC	Advanced OSCC
29	264	SAMP				
30	93	TF				
31	373	VCL				
32	217	VDBP				

Additionally, principal component analysis (PCA) was also performed on all the differentially expressed protein spots as quality control to inspect the separation distant between the groups using Progenesis SameSpot v4.0 software. All the differentially expressed protein spots that were detected in the OPMD, early OSCC, and advanced OSCC showed a distinct separation with the control (Figure 4.3). Nevertheless, there were no distant separation of the differentially expressed protein spots between OPMD and OSCC as well as early and advanced OSCC in the PCA plot (Figure 4.4).

Universiti Malaya

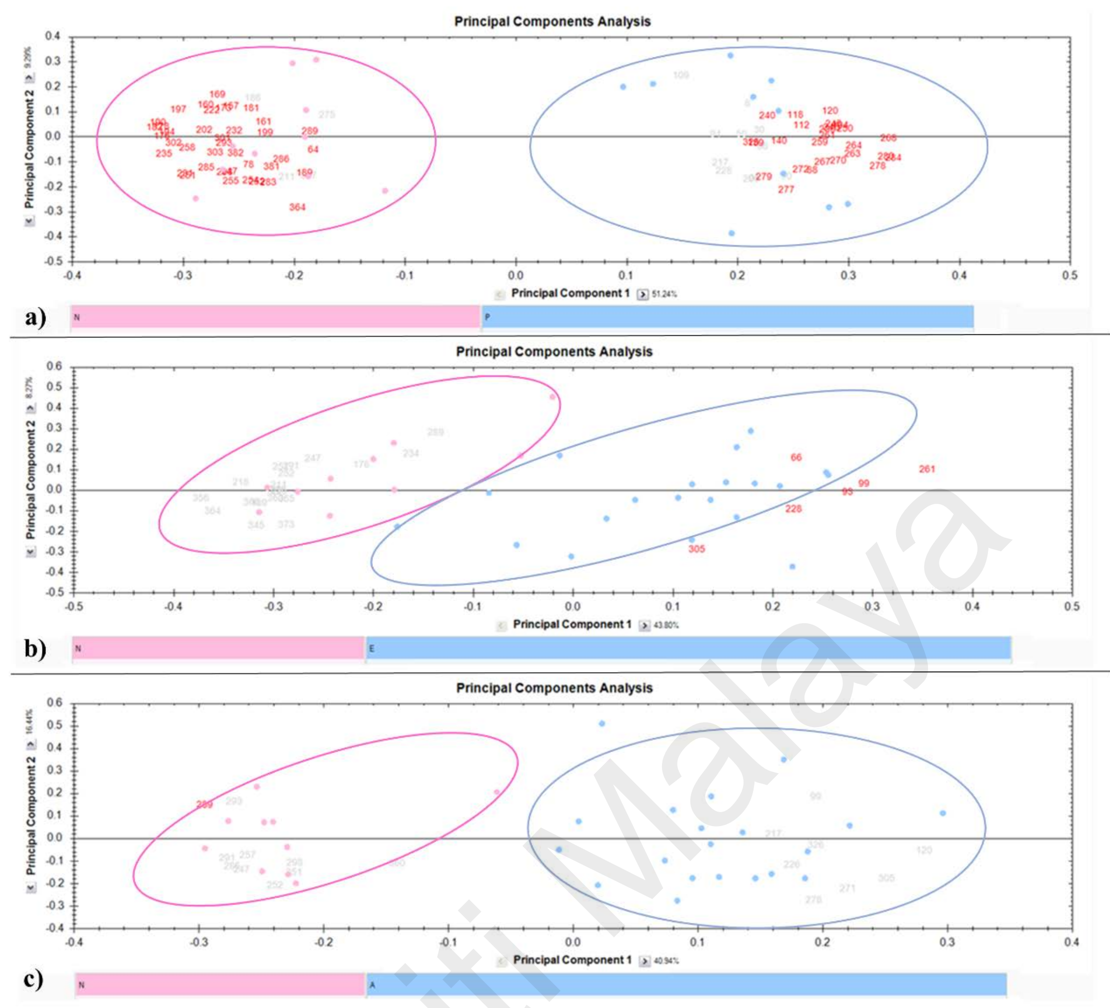


Figure 4.3: Principal component analysis (PCA) plot of (a) OPMD, (b) early OSCC, and (c) advanced OSCC compared with control

PCA plot between the study groups of (a) OPMD and control, (b) early OSCC and control, and (c) advanced OSCC and control.

Universiti Malaya

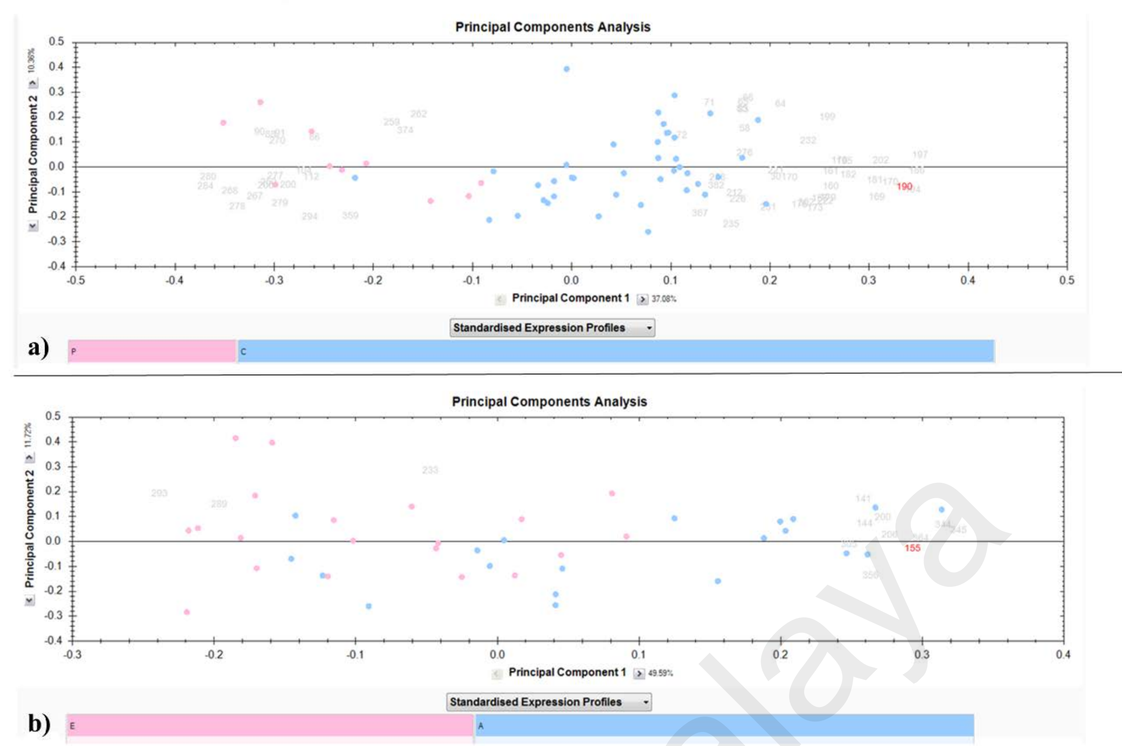


Figure 4.4: Principal component analysis (PCA) plot of OPMD and OSCC

PCA plot between the study groups of (a) OPMD and OSCC, and (b) early OSCC and advanced OSCC.

Universiti Malaya

The relative protein abundance was plotted as fold change to inspect the changes of the differentially expressed proteins in OPMD and OSCC compared with the control (Figure 4.5). This fold change was calculated based on the average normalised volume of each protein spot. A total of 9 proteins (AAT, APOA1, C3, IGHA2, IGHG2, IGKC, SAMP, TF, and VDBP) were up-regulated and 11 proteins (AINX, ALB, AMBP, ARF, CLU, HP, LRG1, PRDX2, RAD50, RBP4, and VCL) were down-regulated in the study groups.

Taken together of the 2-DE protein profiles and MS analysis, AAT, APOA1, IGKC, SAMP, and VDBP were up-regulated, while AMBP, CLU, HP, PRDX2, and RBP4 were down-regulated in OPMD compared with the control. In early OSCC, IGHA2, IGHG2, IGKC, and TF were up-regulated and ALB, AMBP, CLU, HP, and LRG1 were down-regulated when compared with the control. Whereas AAT, APOA1, C3, IGHG2, and VDBP were up-regulated and PRDX2 was down-regulated in advanced OSCC when compared with the control.

Universiti Malaya

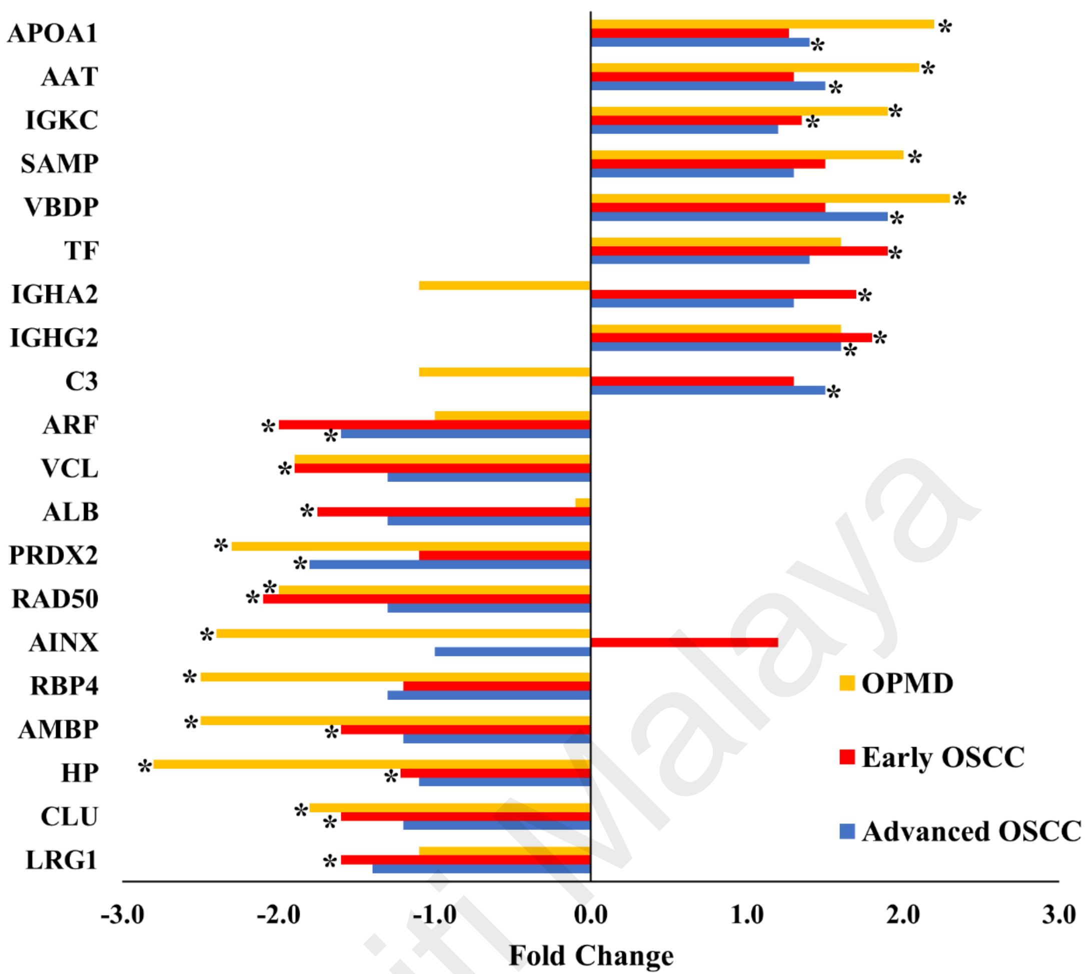


Figure 4.5: The fold change (relative protein abundance) level of identified proteins in OPMD, early OSCC, and advanced OSCC compared with control

Proteins marked with * were significantly different compared with the control ($p < 0.05$).

Universiti Malaya

4.3 Glycoproteomics Analysis

The combination of 2-DE coupled with lectin-based approaches was used to generate glycoprotein profiles of OPMD and OSCC. Concanavalin A (Con A) and Jacalin lectins were used to study *N*- and *O*-glycoprotein profiles according to their sugar-binding specificity, respectively. The *N*- and *O*-glycoprotein profiles showed distinct glycoprotein patterns.

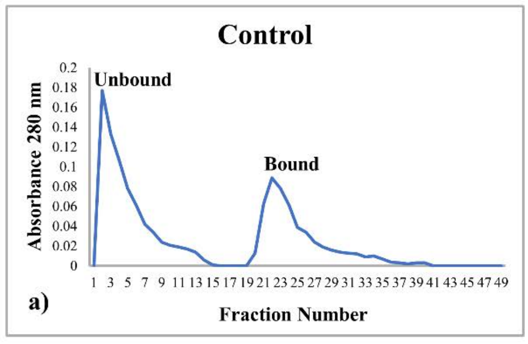
4.3.1 *N*-glycoprotein Analysis

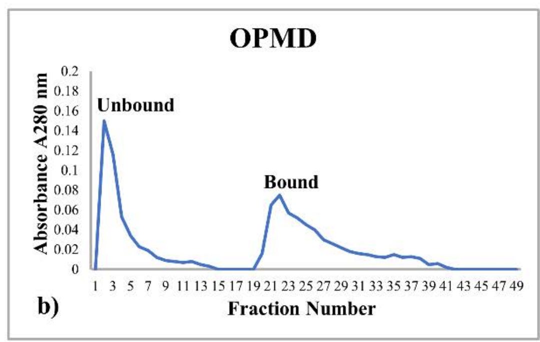
4.3.1.1 Con A affinity chromatography

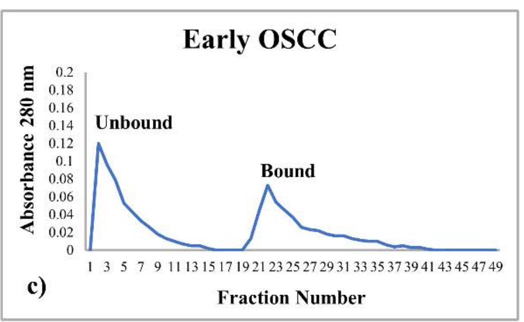
N-glycoproteins from each serum sample were enriched using Con A affinity column chromatography prior subjected to 2-DE. The absorbance of the unbound fractions and bound fractions of Con A lectin with 3 M methyl-α-D-glucopyranoside was monitored at 280 nm. Fraction 1 to 15 contained the unbound fractions. Once the absorbance reading reached the baseline, elution of bound fractions was started and monitored from fraction 19 to 49 until it reached the baseline again. The representative elution profiles of control, OPMD, early OSCC, and advanced OSCC are shown in Figure 4.6.

The average yield of protein from eluted fractions of control, OPMD, early OSCC, and advanced OSCC after Vivaspin concentration was approximately 853.7 µg/ml, 797.5 µg/ml, 743.4 µg/ml, and 829.7 µg/ml, respectively. The concentrated fractions with bound *N*-glycoproteins are desalted to non-protein contaminants prior subjected to 2-DE to obtain the *N*-glycoprotein profile.

Universiti Malaya







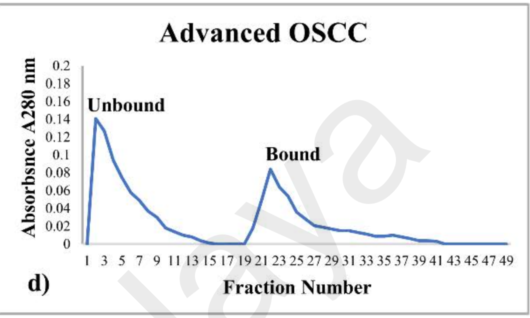


Figure 4.6: Representative elution profiles of (a) control, (b) OPMD, (c) early OSCC, and (d) advanced OSCC

4.3.1.2 Image analysis of *N*-glycoprotein profile

A total of four glycoprotein clusters were consistently detected by Con A lectin in the study groups. *N*-glycosylated AAT and HP were mainly detected in control. *N*-glycoproteins, including AAT, AHSG, APOA1, and HP were consistently detected in OPMD. In OSCC, *N*-glycosylated AAT, AHSG, APOA1, and HP were detected in early OSCC, whereas *N*-glycosylated AAT, APOA1, and HP were detected in advanced OSCC. The typical representative serum *N*-glycoprotein profiles of control, OPMD, early OSCC, and advanced OSCC are depicted in Figure 4.7.

Universiti Malaya

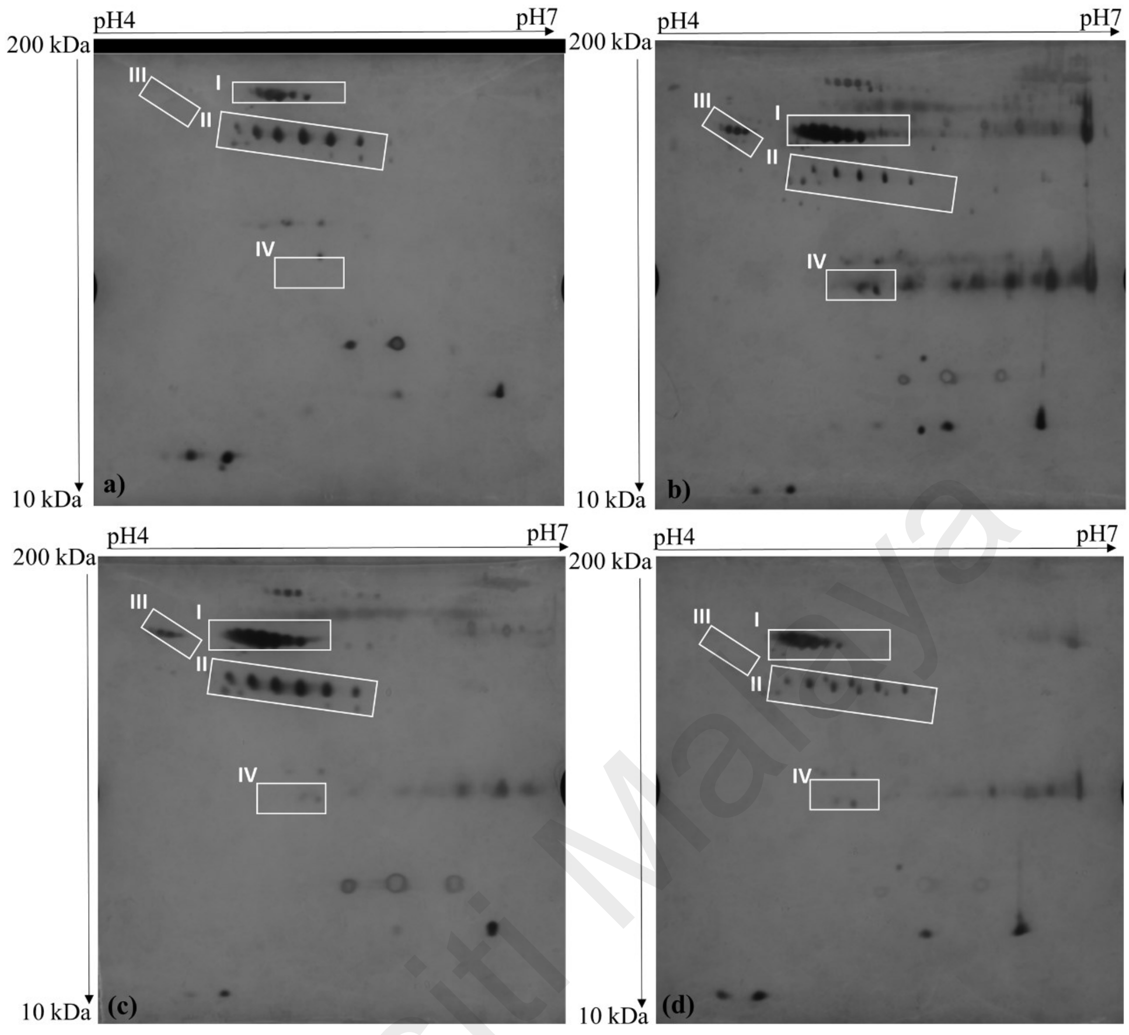


Figure 4.7: Typical representative serum *N*-glycoprotein profiles of (a) control, (b) OPMD, (c) early OSCC, and (d) advanced OSCC

Unfractionated serum samples of control, OPMD, early OSCC, and advanced OSCC were subjected to Con A affinity column chromatography coupled with 2-DE and silver staining. The spot clusters are I: AAT, II: HP, III: AHSG, and IV: APOA1.

Universiti Malaya

4.3.1.3 *N*-glycoprotein profile analysis

When the comparative analysis was performed on the *N*-glycoprotein profiles of control, OPMD, early OSCC, and advanced OSCC, 2 *N*-glycoproteins, which were AAT and APOA1 showed significant differences in the study groups. AAT and APOA1 were also significantly up-regulated in early OSCC and OPMD, respectively. The *N*-glycoproteins detected in each study group are summarized in Table 4.7.

Table 4.7: List of detected *N*-glycoproteins in OPMD, early OSCC, and advanced OSCC when compared with control

No.	Protein	OPMD		Early OSCC		Advanced OSCC	
		p-value	Fold change	p-value	Fold change	p-value	Fold Change
1	AAT	0.080	1.2 (up)	0.004	1.3 (up)	0.368	1.2 (up)
2	HP	0.054	1.3 (down)	0.661	1.0 (up)	0.558	1.1 (down)
3	AHSG	0.226	1.2 (up)	0.994	1.0 (up)	0.625	1.1 (down)
4	APOA1	2.68E-04	2.2 (up)	0.132	1.3 (up)	0.120	1.9 (up)

Universiti Malaya

4.3.2 *O*-glycoprotein Analysis

4.3.2.1 Image analysis of *O*-glycoprotein profile

To obtain *O*-glycoprotein profiles, the unfractionated serum samples were subjected to 2-DE coupled with Jacalin lectin blotting. Only 2 protein clusters were consistently detected by Jacalin lectin in the study groups. The *O*-glycosylated CLU was the only *O*-glycoprotein consistently detected in control and OPMD. Whereas *O*-glycosylated ASHG and CLU were detected in both early and advanced OSCC. The typical representative serum *O*-glycoprotein profiles of control, OPMD, early OSCC, and advanced OSCC are illustrated in Figure 4.8.

Universiti Malaya

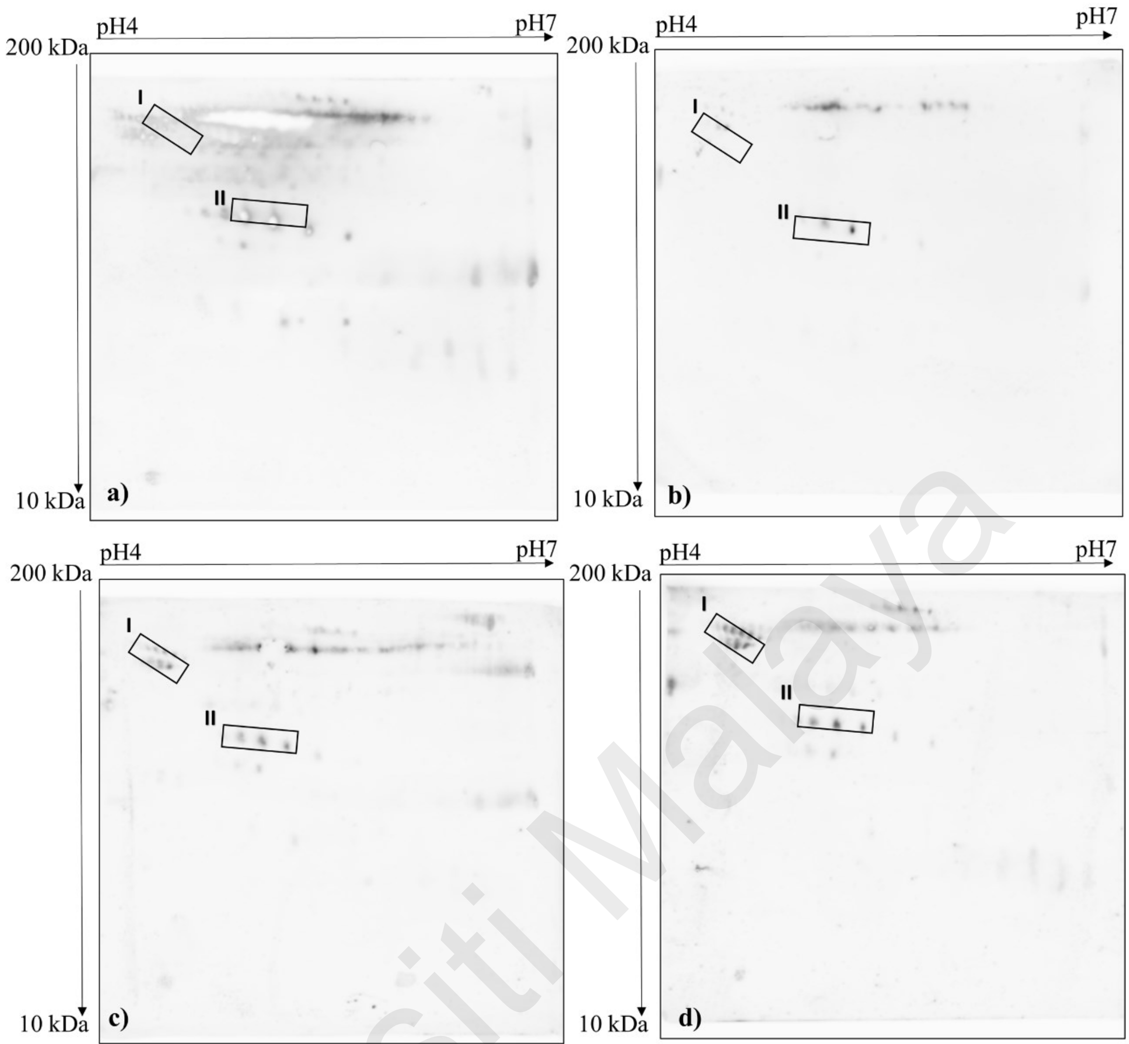


Figure 4.8: Typical representative serum *O*-glycoprotein profiles of (a) control, (b) OPMD, (c) early OSCC, and (d) advanced OSCC

Unfractionated serum samples of control, OPMD, early OSCC, and advanced OSCC were subjected to 2-DE coupled with Jacalin lectin blotting. The labelled spot clusters are I: AHSG and II: CLU.

Universiti Malaya

4.3.2.2 *O*-glycoprotein profile analysis

There were no significant differences between the expression levels of all detected glycoproteins in the *O*-glycoprotein profiles of control, OPMD, early OSCC, and advanced OSCC. The *O*-glycoproteins detected in the study groups are summarized in Table 4.8.

Table 4.8: List of detected *O*-glycoproteins in OPMD, early OSCC, and advanced OSCC when compared with control

No.	Protein	OPMD		Early OSCC		Advanced OSCC	
		p-value	Fold change	p-value	Fold change	p-value	Fold Change
1	AHSG	0.886	1.6 (up)	0.753	1.6 (up)	0.525	2.0 (up)
2	CLU	0.530	1.9 (down)	0.954	1.2 (down)	0.950	1.1 (up)

Universiti Malaya

4.4 Bioinformatics Analysis

The 21 identified proteins and glycoproteins (AAT, AHSG, AINX, ALB, AMBP, APOA1, ARF, C3, CLU, HP, IGHA2, IGHG2, IGKC, LRG1, PRDX2, RAD50, RBP4, SAMP, TF, VCL, and VDBP) were analysed using DAVID v6.8 to understand their functional roles in the development and progression of OSCC. The gene ontology (GO) analysis has classified these proteins into a total of 39 gene ontology (GO) annotation terms and were divided into three functional categories, including biological process, molecular function, and cellular component ($p < 0.05$). In the biological process category, most of the identified proteins were involved in platelet degranulation, receptor-mediated endocytosis, complement activation, and acute phase response. As for molecular function category, these proteins were associated with the binding activities, including antioxidant, immunoglobulin receptor, and serine-type endopeptidase binding. While in the cellular component group, most of the proteins are belonged to blood microparticle and expressed in the extracellular space or region. The significant results of GO analysis are shown in Figure 4.9.

Universiti Malaya

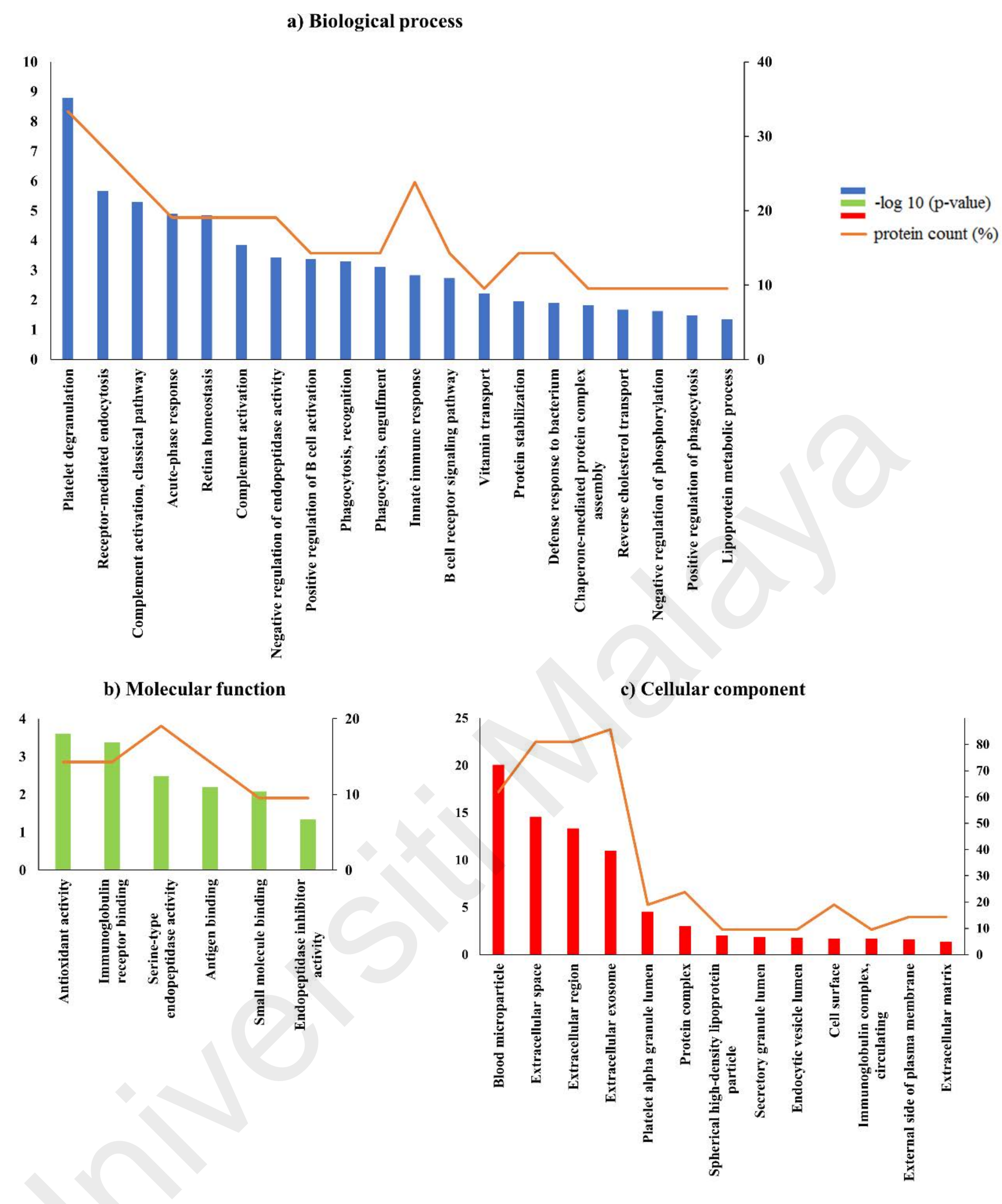


Figure 4.9: Gene ontology analysis of identified proteins

University Malaya

To further infer the functional roles of these identified proteins, functional annotation clustering analysis was used to assess the relationship of these identified proteins based on their annotation terms. Based on the enrichment score, these proteins were clustered into four groups (Table 4.9). Approximately 81.0% of the proteins were enriched in the extracellular space or region. This functional annotation clustering analysis revealed that AAT, AHSG, ALB, APOA1, CLU, TF, and VCL are involved in the platelet degranulation, whereas C3, CLU, IGHA2, IGHG2, and IGKC are related to the activation of classical complement pathway. Furthermore, the molecular function of C3, IGHG2, IGKC, and HP was associated with the serine-type endopeptidase activity.

Table 4.9: Functional annotation clustering analysis of identified proteins using DAVID v6.8

GO group	Term name	Protein count (%)	Protein name	Fold enrichment	p-value
Enrichment score: 16.00					
Cellular Component	Blood microparticle	13 (61.90)	AHSG, ALB, AMBP, APOA1, C3, CLU, HP, IGHA2, IGHG2, IGKC, SAMP, TF, VDBP	74.22	8.63E-21
Cellular Component	Extracellular space	17 (80.95)	AAT, AHSG, AINX, ALB, AMBP, APOA1, C3, CLU, HP, IGHA2, IGHG2, IGKC, LRG1, RBP4, SAMP, TF, VDBP	10.95	2.61E-15
Cellular Component	Extracellular region	17 (80.95)	AAT, AHSG, ALB, AMBP, APOA1, C3, CLU, HP, IGHA2, IGHG2, IGKC, LRG1, RBP4, SAMP, TF, VCL, VDBP	9.16	4.42E-14
Enrichment score: 3.64					
Biological Process	Platelet degranulation	7 (33.33)	AAT, AHSG, ALB, APOA1, CLU, TF, VCL	54.34	1.66E-09

Universiti Malaya

Table 4.9 continued

GO group	Term name	Protein count (%)	Protein name	Fold enrichment	p-value
Cellular Component	Platelet alpha granule lumen	4 (19.05)	AAT, AHSG, ALB, CLU	63.11	2.86E-05
Enrichment score: 3.14					
Biological Process	Complement activation, classical pathway	5 (23.81)	C3, CLU, IGHA2, IGHG2, IGKC	40.38	5.12E-05
Biological Process	Complement activation	4 (19.05)	C3, CLU, IGHG2, IGKC	36.76	1.44E-04
Molecular Function	Serine-type endopeptidase activity	4 (19.05)	C3, IGHG2, IGKC, HP	12.61	0.003
Enrichment score: 3.09					
Molecular Function	Immunoglobulin receptor binding	3 (14.29)	IGHA2, IGHG2, IGKC	92.75	4.26E-04
Biological Process	Positive regulation of B cell activation	3 (14.29)	IGHA2, IGHG2, IGKC	92.26	4.31E-04
Biological Process	Phagocytosis, recognition	3 (14.29)	IGHA2, IGHG2, IGKC	85.67	5.00E-04
Biological Process	Phagocytosis, engulfment	3 (14.29)	IGHA2, IGHG2, IGKC	68.54	7.83E-04
Biological Process	Innate immune response	5 (23.81)	CLU, IGHA2, IGHG2, IGKC, VDBP	9.30	0.001
Biological Process	B cell receptor signalling pathway	3 (14.29)	IGHA2, IGHG2, IGKC	44.42	0.002
Molecular Function	Antigen binding	3 (14.29)	IGHA2, IGHG2, IGKC	23.41	0.007
Cellular Component	External side of plasma membrane	3 (14.29)	IGHA2, IGHG2, IGKC	12.22	0.022

Universiti Malaya

To explore the pathway annotations and interaction network of the identified proteins, the identified proteins were analysed using IPA v7.1. Liver X receptor/retinoid X receptor (LXR/RXR) activation, famesoid X receptor/retinoid X receptor (FXR/RXR) activation, acute phase response signalling, clathrin-mediated endocytosis signalling, and atherosclerosis signalling were the top five canonical pathways that found in the canonical pathway analysis (Figure 4.10). Among the canonical pathways, LXR/RXR activation and acute phase response signalling pathways were activated, whereas the production of nitric oxide and reactive oxygen species in macrophages was inhibited in OSCC. The relevant proteins of these canonical pathways are shown in Table 4.10.

Universiti Malaya

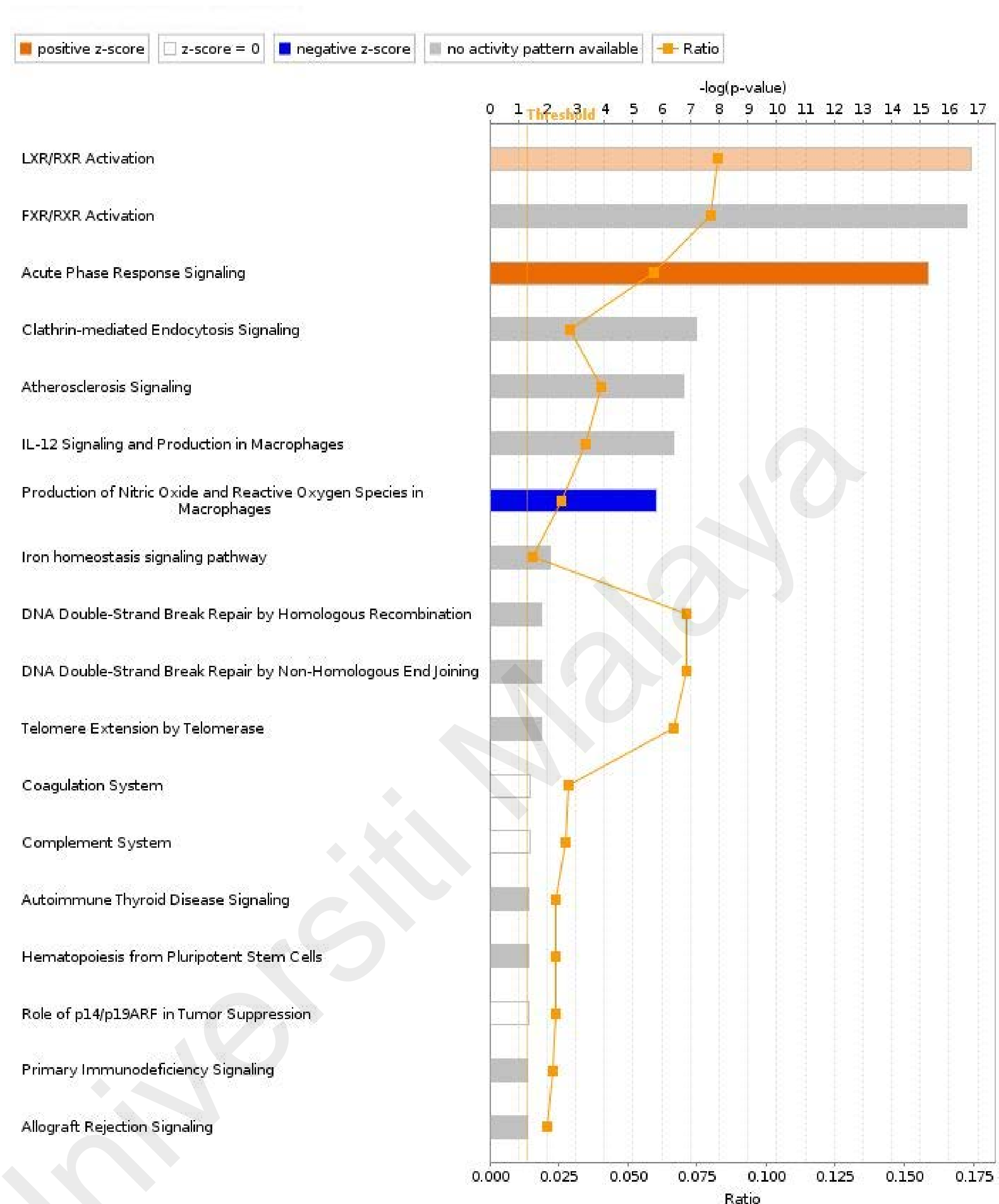


Figure 4.10: Canonical pathway analysis of identified proteins

Universiti Malaya

Table 4.10: List of the proteins in the activation/inhibition canonical pathways

No.	Canonical pathway	-log(p-value)	z-score	Protein name
1	LXR/RXR activation	16.7	0.632	AAT, AHSG, ALB, AMBP, APOA1, C3, CLU, RBP4, TF, VDBP
2	Acute phase response signalling	15.3	1.633	AAT, AHSG, ALB, AMBP, APOA1, C3, HP, SAMP, RBP4, TF
3	Production of nitric oxide and reactive oxygen species in macrophages	5.81	-0.447	AAT, ALB, APOA1, CLU, RBP4

IPA revealed that these proteins were also involved in lipid metabolism, small molecule biochemistry, cellular compromise, cell cycle, and cell-to-cell signalling and interaction. The top five significant molecular and cellular functions of the identified proteins are listed in Table 4.11.

Table 4.11: Top five significant molecular and cellular functions of identified proteins

No.	Molecular and cellular functions	p-value	Number of proteins	Protein name
1	Lipid metabolism	1.18E-02 – 8.57E-07	6	AAT, ALB, APOA1, C3, CLU, PRDX2
2	Small molecule biochemistry	1.18E-02 – 8.57E-07	6	AAT, ALB, APOA1, C3, CLU, PRDX2
3	Cellular compromise	8.59E-03 – 2.45E-06	7	AAT, ALB, ARF, C3, CLU, IGHA2, SAMP
4	Cell cycle	1.18E-02 – 4.96E-06	6	ALB, ARF, CLU, HP, RAD50, TF
5	Cell-to-cell signalling and interaction	1.18E-02 – 9.62E-06	10	AAT, ALB, ARF, C3, IGHA2, IGKC, PRDX2, SAMP, VCL, VDBP

Universiti Malaya

Moreover, a total of 17 identified proteins (AAT, AHSG, ALB, AMBP, APOA1, ARF, C3, CLU, HP, IGHG2, PRDX2, RAD50, RBP4, SAMP, TF, VCL, and VDBP) were related to inflammatory response, organismal injury and abnormalities, which is the major associated network functions in the associated network functions analysis (Table 4.12).

Table 4.12: Associated network functions analysis

No.	Top disease and functions	Score	Focus Molecules	Molecules in network
1	Inflammatory response, organismal injury and abnormalities, cardiovascular disease	30	17	AAT, ABCA1, ACTA2, AGT, AHSG, AIMP2, ALB, AMBP, ANGPTL3, ANXA2, APOA1, APOA2, APOC2, APOH, ARF, ATM, ATP7A, BAK1, BCL2, BTG3, C3, C5AR2, CANX, CASP9, Caspase 3/7, Cbp/p300, CCAR2, CCL18, CD40, CD44, CD163, CD209, CDC25B, CDC25C, CEACAM8, CETP, CFB, CLU, Collagen Alpha1, Collagen type II, COMMD1, CR2, CXCL12, CYBA, DAB2IP, DOHH, Dynein, EIF4G1, elastase, ERK1/2, FABP4, FCER2, FGFR3, Fibrinogen, FTH1, FURIN, GFI1, GHRL, GLI2, HABP2, hemoglobin, HNF1A, HNF1B, HP, HSF2, Ige, IgG4, IGHE, IGHG2, Igm, Integrin, IRS1, ITGAX, LDL, LGALS3, LIPC, LPL, MAP2K2, MAPK8, MDM4, mir-21, mir-26, mir-27, mir-146, mir-181, miR-146a-5p (and other miRNAs w/seed GAGAACU), miR-24-3p (and other miRNAs w/seed GGCUCAG), MMP3, MRE11, MYB, MYL2, NDUFA13, NFkB (complex), Nr1h, NR1H3, NR1H4, NR5A2, OSCAR, OSM, OSMR, PAK1, PDIA3, Pkg, PLAA, PLAU, PLSCR1, PPARA, PRDX2, PXN, RAD17, RAD50, RBP4, RBPJ, RNF31, S100A11, SAMP, SELP, SMC3, SMC1A, SMPD2, SRC (family), TERF2, TERF2IP, TF,

Universiti Malaya

Table 4.12 continued

No	Top disease and functions	Score	Focus Molecules	Molecules in network
				THBD, TNF, Tnf receptor, TNFSF13B, TP53I3, TSG101, TTR, TXN, USP11, VCAM1, VCL, VCP, VDBP, VHL, VTN, ZEB1
2	Cell-to-cell signalling and interaction, cellular compromise, cellular function and maintenance	2	1	FCAR, IGHA2
3	Cardiovascular disease, cell morphology, development disorder	2	1	CDK5, LRG1
4	Endocrine system development and function, lipid metabolism, small molecule biochemistry	2	1	AGTR1, AINX, REST, STAU1

The protein interaction network was also generated to delineate the interacting proteins involved in the relevant network. The molecule activity predictor was overlaid with inflammatory response, organismal injury and abnormalities, and cardiovascular disease network to predict the upstream and downstream effects (activation or inhibition) of neighbouring molecules of the identified proteins. The network analysis showed direct and indirect interaction of the identified proteins (AAT, AHSG, ALB, AMBP, APOA1, ARF, C3, CLU, HP, IGHG2, PRDX2, RAD50, RBP4, SAMP, TF, VCL, and VDBP) according to the Ingenuity pathways knowledge base. These differentially expressed proteins had led to predicted activation of neighbouring molecules like B cell lymphoma 2 (BCL2), cell division cycle 25B (CDC25B), cluster of differentiation 44 (CD44), CD163, extracellular signal-regulated kinase 1 and 2 (ERK1/2), and tumour necrosis factor (TNF) as well as predicted inhibition of neighbouring molecules like ataxia telangiectasia mutated (ATM), caspase 3/7, chemokine (C-X-C motif) ligand 2 (CXCL12), CD40, and cell cycle checkpoint protein RAD17 (RAD17) (Figure 4.11).

Universiti Malaya

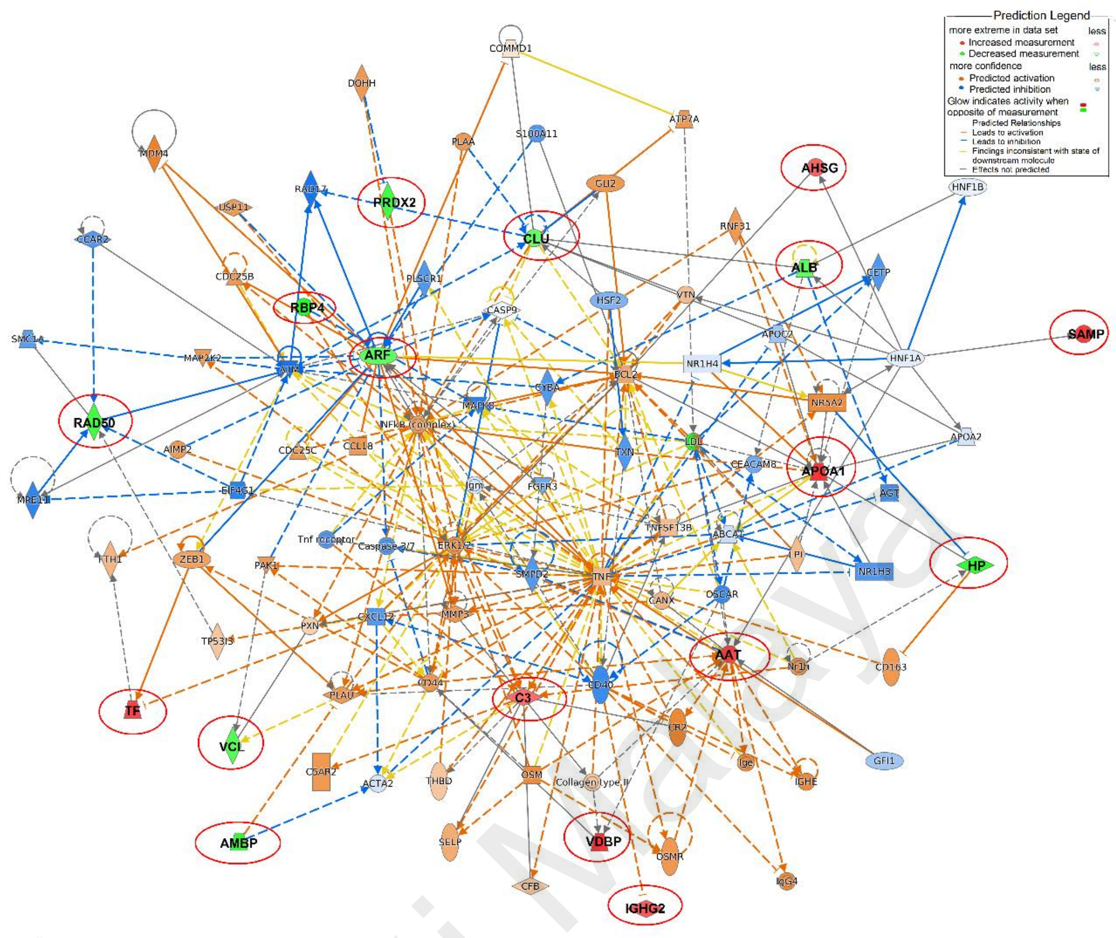


Figure 4.11: Network analysis of identified proteins with molecule activity prediction

Direct and indirect relationships are indicated by solid lines and dashed lines, respectively

Universiti Malaya

Additionally, the NetNGlyc 1.0 and NetOGlyc 4.0 servers were used to predict the potential glycosylation sites of the identified glycoproteins in this study. AAT, AHSG, AHSG, and HP are detected as *N*-glycoproteins, whereas AHSG and CLU are detected as *O*-glycoproteins. Most of these identified glycoproteins had both predicted *N*- and *O*-glycosylation sites. However, no sites of *N*-glycosylation are predicted by the NetNGlyc 1.0 server for APOA1. The predicted glycosylation sites of the identified glycoproteins are shown in Table 4.13, and the details of the analysis are shown in Appendix C.

Table 4.13: Predicted glycosylation sites of the identified glycoproteins

Protein	Predicted glycosylation sequence (amino acid position)	
	***N*-glycosylation site**	***O*-glycosylation site**
AAT	Positive: 70, 107, 271, 414	Positive: 35, 37, 38, 46, 337
AHSG	Positive: 156, 176	Positive: 252, 256, 257, 270, 280, 293, 297, 319, 325, 328, 330, 334, 339, 341, 345
APOA1	no	Positive: 14, 185, 221, 228
CLU	Positive: 86, 103, 145, 291, 354, 374	Positive: 170, 210, 264, 292, 394
HP	Positive: 184, 207, 211, 241	Positive: 67, 126, 316, 317, 323

Universiti Malaya

4.5 Enzyme-linked Immunosorbent Assay (ELISA) Analysis

A total of 120 serum samples from normal healthy volunteers (n = 35) as control and patients diagnosed with OPMD (n = 12), early stage (n = 34), and advanced stage (n = 39) of OSCC from the period from 2004 to 2015 were included in this study for ELISA analysis. Out of the 120 serum samples, 60 samples are from the proteomics and glycoproteomics study set. These samples were obtained from MOCDTBS at OCRCC-UM. Most of the samples in the study group were in the older age group (45 years of age and over). The mean age of control, OPMD, early OSCC, and advanced OSCC were 52.00±17.32, 53.92±13.59, 56.26±17.88, and 60.62±13.29, respectively. In this ELISA analysis, most of the patients diagnosed with OPMD and OSCC were Indian females. Among the OPMD and OSCC patients, betel quid chewing had a higher percentage than smoking and alcohol drinking habits. Additionally, buccal mucosa, gingiva, lip, and palate were the most common sites among all OPMD cases. While tongue and floor of mouth were the most common sites among all OSCC cases. The demographic characteristics of the study samples are shown in Table 4.14.

Universiti Malaya

Table 4.14: Demographic characteristics of study samples in ELISA analysis

	Control (n = 35) (%)	OPMD (n = 12) (%)	Early OSCC (n = 34) (%)	Advanced OSCC (n = 39) (%)
Age (years)	52.00±17.32	53.92±13.59	56.26±17.88	60.62±13.29
< 45	12 (34.3)	1 (8.3)	9 (26.5)	6 (15.4)
≥ 45	23 (65.7)	11 (91.7)	25 (73.5)	33 (84.6)
Gender				
Male	18 (51.4)	1 (8.3)	12 (35.3)	14 (35.9)
Female	17 (48.6)	11 (91.7)	22 (64.7)	25 (64.1)
Ethnic group				
Malay, Chinese & Others	32 (91.4)	7 (58.3)	19 (55.9)	21 (53.8)
Indian	3 (8.6)	5 (41.7)	15 (44.1)	18 (46.2)
Smoking				
No	26 (74.3)	11 (91.7)	24 (70.6)	27 (69.2)
Yes	9 (25.7)	1 (8.3)	10 (29.4)	12 (30.8)
Alcohol drinking				
No	33 (94.3)	11 (91.7)	26 (76.5)	32 (82.1)
Yes	2 (5.7)	1 (8.3)	8 (23.5)	7 (17.9)
Betel quid chewing				
No	32 (91.4)	8 (66.7)	24 (70.6)	23 (59.0)
Yes	3 (8.6)	4 (33.3)	10 (29.4)	16 (41.0)
Site				
Tongue, floor of mouth	-	2 (16.7)	19 (55.9)	21 (53.8)
Buccal mucosa, gingiva, lip & palate	-	10 (83.3)	15 (44.1)	18 (46.2)

Universiti Malaya

In ELISA analysis, the expression levels of AAT, AHSG, and APOA1 were higher in OPMD and OSCC as compared with control. Whereas the expression level of CLU was lower in OPMD and OSCC, while the expression level of HP was also lower in OSCC when compared to control. But, the expression level of HP was higher in OPMD than in control. The expression levels of AAT, AHSG, APOA1, CLU, and HP in the study groups were detailed in Table 4.15 and Figure 4.12.

Table 4.15: The expression level of AAT, AHSG, APOA1, CLU, and HP in control, OPMD, early OSCC, and advanced OSCC

	Control	OPMD	Early OSCC	Advanced OSCC
AAT Mean ± SEM (ng/ml)	596.67 ± 81.97	765.18 ± 106.69	816.14 ± 103.69	1280.51 ± 141.41
AHSG Mean ± SEM (µg/ml)	1045.27 ± 112.77	1839.73 ± 341.08	1390.65 ± 236.30	1217.93 ± 114.76
APOA1 Mean ± SEM (µg/ml)	481.56 ± 62.85	930.69 ± 139.69	1100.46 ± 248.11	1411.58 ± 236.31
CLU Mean ± SEM (µg/ml)	193.71 ± 15.11	72.34 ± 14.29	63.82 ± 8.46	62.36 ± 6.23
HP Mean ± SEM (µg/ml)	1125.62 ± 151.36	1244.92 ± 223.53	617.51 ± 62.91	844.59 ± 94.19

Universiti Malaya

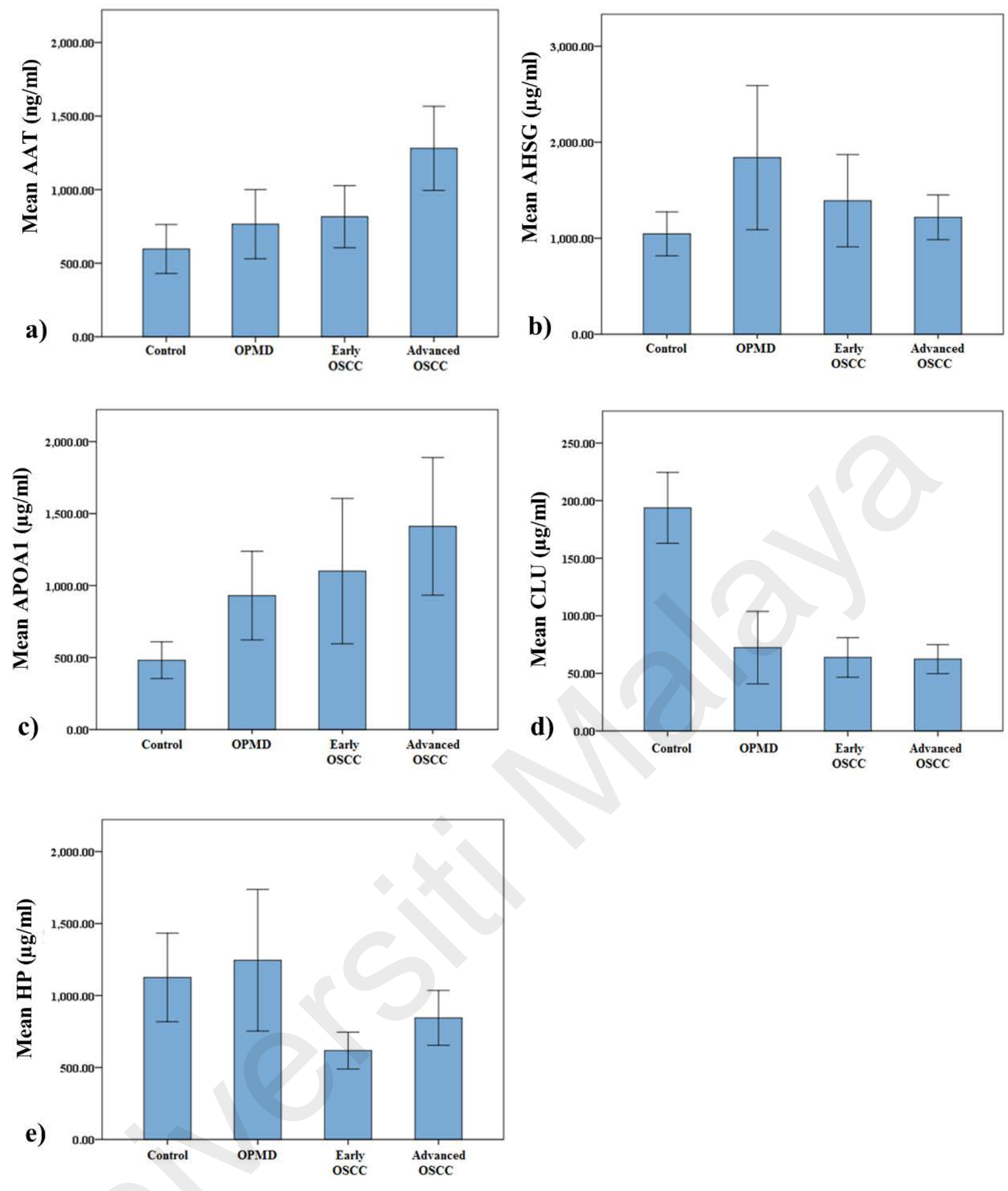


Figure 4.12: Mean expression level of (a) AAT, (b) AHSG, (c) APOA1, (d) CLU, and (e) HP in control, OPMD, early OSCC, and advanced OSCC

Universiti Malaya

Receiver operator characteristic (ROC) analysis was performed to determine the optimal cut-off value and the diagnostic performance given by the area under the curve (AUC) with 95% confidence interval (95%CI). The low and high expression levels of AAT, AHSG, APOA1, CLU, and HP for each study group were distinguished based on the optimal cut-off value. The results of ROC analysis for AAT, AHSG, APOA1, CLU, and HP are shown in Table 4.16.

Table 4.16: Optimal cut-off value and area under the curve (AUC) in ROC analysis

	Optimal cut-off value	Sensitivity (%)	Specificity (%)	AUC (95%CI)	p-value
AAT	283.19 ng/ml	91.8	37.1	0.693 (0.59-0.80)	0.001
AHSG	1013.82 µg/ml	56.2	71.4	0.617 (0.59-0.73)	0.049
APOA1	561.77 µg/ml	60.3	68.6	0.690 (0.59-0.80)	0.001
CLU	111.29 µg/ml	94.5	89.0	0.945 (0.91-0.99)	< 0.001
HP	1360.29 µg/ml	34.3	90.4	0.617 (0.51-0.73)	0.049

Univariate and multivariate logistic regression analyses were carried out to examine the association between the expression level of AAT, AHSG, APOA1, CLU, and HP with the study groups (Table 4.17 and Table 4.18). The expression level of AAT was significantly higher in OSCC compared with control after further strengthened by multivariate logistic regression analysis that were adjusted with variables including age, gender, ethnic group, smoking, alcohol drinking, and betel quid chewing habit. The odds ratio (OR) for patients with serum AAT cut-off value of > 283.19 ng/ml being associated with OSCC was 6.23 (95% CI: 1.68-23.14), and this cut-off value is also being associated

Universiti Malaya

with early OSCC where the OR was 6.68 (95%CI: 1.23-36.20). Moreover, the expression level of AHSG was significantly higher in advanced OSCC when compared with control, and the OR for patients with serum AHSG cut-off value of > 1013.82 µg/ml being associated with advanced OSCC was 4.84 (95%CI: 1.30-17.96). The results indicated AAT and AHSG are independent factors that have an impact on OSCC.

The expression level of APOA1 was significantly higher in OPMD and OSCC compared with control which further strengthened by multivariate logistic regression analysis that were adjusted with variables including age, gender, ethnic group, smoking, alcohol drinking, and betel quid chewing habit. The results indicated APOA1 is an independent factor that impacted on OPMD and OSCC. The OR for patients with serum APOA1 cut-off value of > 561.77 µg/ml being associated with OPMD was 11.43 (95%CI: 1.38-94.50), and this cut-off value is also being associated with OSCC where the OR was 3.60 (95%CI: 1.35-9.62).

The expression level of CLU was significantly lower in OSCC compared with control after further strengthened by multivariate logistic regression analysis that were adjusted with variables including gender, ethnic group, smoking, alcohol drinking, and betel quid chewing habit. The results indicated CLU is an independent factor that impacted on OSCC with serum CLU cut-off value of < 111.29 µg/ml. The logistic regression analyses also revealed the expression level of HP was significantly lower in OSCC compared with control after further strengthened by multivariate logistic regression analysis. It was also significantly higher in early OSCC compared with OPMD. The results indicated HP is an independent factor that impacted on OSCC and the OR for patients with serum HP cut-off value of < 1360.29 µg/ml being associated with OSCC was 8.00 (95%CI: 1.85-34.63).

Universiti Malaya

Table 4.17: Association analysis between the expression level of AAT, AHSG, and APOA1 with the study groups

	AAT				AHSG				APOA1			
	Univariate		Multivariate*		Univariate		Multivariate*		Univariate		Multivariate*	
	OR (95%CI)	p-value	OR (95%CI)	p-value	OR (95%CI)	p-value	OR (95%CI)	p-value	OR (95%CI)	p-value	OR (95%CI)	p-value
OPMD vs Control+	6.50 (0.75-56.30)	0.089	17.33 (0.36-824.67)	0.148	5.00 (1.23-20.41)	**0.025**	4.82 (0.64-36.15)	0.126	6.55 (1.48-29.10)	**0.013**	11.43 (1.38-94.50)	**0.024**
OSCC vs Control+	6.60 (2.24-19.44)	**0.001**	6.23 (1.68-23.14)	**0.006**	3.203 (1.35-7.62)	**0.008**	2.49 (0.93-6.67)	0.070	3.31 (1.41-7.78)	**0.006**	3.60 (1.35-9.62)	**0.011**
Early OSCC vs Control+	6.11 (1.55-24.01)	**0.010**	6.681 (1.23-36.20)	**0.028**	2.50 (0.92-6.76)	0.071	1.72 (0.54-5.49)	0.358	2.50 (0.92-6.55)	0.073	2.56 (0.82-8.04)	0.107
Advanced OSCC vs Control+	7.09 (1.82-27.71)	**0.005**	4.17 (0.89-19.52)	0.070	4.00 (1.51-10.62)	**0.005**	4.84 (1.30-17.96)	**0.019**	4.36 (1.64-11.58)	**0.003**	4.03 (1.24-13.10)	**0.020**
OSCC vs OPMD+	1.02 (0.11-9.26)	0.989	0.51 (0.51-5.47)	0.577	0.64 (0.18-2.32)	0.497	0.69 (0.18-2.72)	0.597	0.51 (0.13-2.03)	0.336	0.66 (0.13-3.43)	0.618
Early OSCC vs OPMD+	0.94 (0.09-10.00)	0.959	0.75 (0.05-10.94)	0.833	0.50 (0.13-1.98)	0.323	0.45 (0.10-2.10)	0.311	0.375 (0.09-1.63)	0.191	0.49 (0.07-2.77)	0.388
Advanced OSCC vs OPMD+	1.09 (0.10-11.57)	0.942	0.48 (0.03-7.06)	0.589	0.80 (0.21-3.13)	0.748	1.24 (0.24-6.46)	0.798	0.67 (0.15-2.89)	0.588	0.61 (0.09-4.15)	0.613
Advanced OSCC vs Early OSCC+	1.16 (0.22-6.17)	0.861	0.94 (0.15-5.78)	0.949	1.60 (0.63-4.06)	0.323	1.69 (0.63-4.49)	0.296	1.18 (0.69-4.58)	0.234	1.76 (0.63-4.89)	0.278

*Adjusted p-value and OR (odd ratio) for age, gender, ethnic group, smoking, alcohol drinking, and betel quid chewing, +Reference category, CI: confidence interval, bold: p-value < 0.05

Universiti Malaya

Table 4.18: Association analysis between the expression level of CLU and HP with the study groups

	CLU				HP			
	Univariate		Multivariate[-]		Univariate		Multivariate*	
	OR (95%CI)	**p-value**	**OR (95%CI)**	**p-value**	**OR (95%CI)**	**p-value**	**OR (95%CI)**	**p-value**
OPMD vs Control[+]	82.50 (10.27-662.85)	**< 0.001**	N/C	0.999	1.04 (0.26-4.18)	0.952	1.66 (0.23-11.96)	0.617
OSCC vs Control[+]	134.06 (26.93-667.43)	**< 0.001**	386.73 (36.73-4065.39)	**< 0.001**	4.92 (1.73-14.00)	**0.003**	8.00 (1.85-34.63)	**0.005**
Early OSCC vs Control[+]	95.70 (17.24-531.26)	**< 0.001**	N/C	0.995	17.22 (2.09-141.78)	**0.008**	21.93 (1.64-293.00)	**0.020**
Advanced OSCC vs Control[+]	198.00 (31.11-1259.87)	**< 0.001**	N/C	0.996	2.87 (0.94-8.75)	0.064	6.67 (1.10-40.55)	**0.039**
OSCC vs OPMD[+]	1.63 (0.30-8.78)	0.573	2.357 (0.38-14.79)	0.360	0.21 (0.05-0.89)	**0.034**	4.13 (0.87-19.66)	0.075
Early OSCC vs OPMD[+]	1.16 (0.19-6.95)	0.871	1.326 (0.17-10.25)	0.787	16.50 (1.62-168.48)	**0.018**	21.12 (1.19-376.05)	**0.038**
Advanced OSCC vs OPMD[+]	2.40 (0.35-16.39)	0.372	9.615 (0.42-219.60)	0.156	0.17 (0.02-1.46)	0.181	2.15 (0.40-11.67)	0.375
Advanced OSCC vs Early OSCC[+]	2.07 (0.456-9.39)	0.346	2.060 (0.43-9.92)	0..368	6.00 (0.68-52.62)	0.106	0.16 (0.02-1.51)	0.111

[-]Adjusted p-value and OR (odd ratio) for gender, ethnic group, smoking, alcohol drinking, and betel quid chewing (number of confounding factors is excluded due to cells containing number of observations of less than 5), N/C: OR (95%CI) not calculated due to cells containing number of observations of less than 5

*Adjusted p-value and OR (odd ratio) for age, gender, ethnic group, smoking, alcohol drinking, and betel quid chewing, [+]Reference category, CI: confidence interval, bold: p-value < 0.05

Universiti Malaya

Chi-square analysis was carried out to examine the association of socio-demographic and clinicopathological characteristics of AAT, AHSG, APOA1, CLU, and HP expression in OSCC (Appendix D). Generally, there was no significant association between AAT, AHSG, APOA1, CLU, and HP expression and the characteristics. However, the level of AAT expression was higher in socio-demographic characteristics of the older age group (91.4%), females (87.2%), and advanced OSCC (92.3%). While, the level of AHSG expression was observed higher in the older age group (58.6%), female (59.6%), Indian (66.7%), and advanced OSCC (61.5%)

The level of APOA1 expression was higher in the older age group (63.8%), female (61.7%), and advanced OSCC (66.7%). Moreover, the high expression of APOA1 was significantly correlated with Broders' grading ($p = 0.019$). Contrary to the expression levels of AAT, AHSG, and APOA1, the level of CLU was noticed lower in the older age group (87.9%), female (99.4%), and advanced OSCC (92.3%). Whereas the level of HP expression was only lower in the older age group (89.7%) and female (89.4%).

Universiti Malaya

4.6 Immunohistochemistry (IHC) Analysis

A total of 70 FFPE tissue samples from normal oral mucosa (n = 10) as control and patients diagnosed with OPMD (n = 11), early stage (n = 18), and advanced stage (n = 21) of OSCC during the period from 2004 to 2018 were included in this study for IHC analysis. The FFPE tissue samples that were obtained from MOCDTBS at OCRCC-UM were independent sample cohorts and did not have any matched serum sample in this study. Most of the sample in OPMD and OSCC were in the older age group (45 years of age and over). The mean age of control, OPMD, early OSCC, and advanced OSCC were 24.30±3.59, 60.00±9.77, 58.22±11.28, and 58.23±12.60, respectively. In this IHC analysis, most of the patients diagnosed with OPMD and OSCC were Indian females. The betel quid chewing habit had a higher percentage in OPMD and OSCC patients compared with control. Furthermore, buccal mucosa, gingiva, lip, and palate were the most common sites among all OPMD and OSCC cases. The demographic characteristics of the study samples are shown in Table 4.19.

Universiti Malaya

Table 4.19: Demographic characteristics of study samples in IHC analysis

	Control (n = 10) (%)	OPMD (n = 11) (%)	Early OSCC (n = 18) (%)	Advanced OSCC (n = 31) (%)
Age (years)	24.30±3.59	60.00±9.77	58.22±11.28	58.23±12.60
< 45	10 (100.0)	0 (0.0)	2 (11.1)	4 (12.9)
≥ 45	0 (0.0)	11 (100.0)	16 (88.9)	27 (87.1)
Gender				
Male	3 (30.0)	2 (18.2)	5 (27.8)	7 (22.6)
Female	7 (70.0)	9 (81.8)	13 (83.3)	24 (77.4)
Ethnic group				
Malay & Chinese	9 (90.0)	2 (18.2)	3 (16.7)	2 (6.5)
Indian	1 (10.0)	9 (81.8)	15 (83.3)	29 (93.5)
Smoking‡				
No	7 (87.5)	9 (90.0)	15 (83.3)	27 (87.1)
Yes	1 (12.5)	1 (10.0)	3 (16.7)	4 (12.9)
‡Data missing	2	1		
Alcohol drinking‡				
No	7 (87.5)	9 (90.0)	12 (66.7)	19 (61.3)
Yes	1 (12.5)	1 (10.0)	6 (33.3)	12 (38.7)
‡Data missing	2	1		
Betel quid chewing‡				
No	8 (100.0)	6 (60.0)	6 (33.3)	8 (25.8)
Yes	0 (0.0)	4 (40.0)	12 (66.7)	23 (74.2)
‡Data missing	2	1		
Site				
Tongue, floor of mouth	-	2 (18.2)	4 (22.2)	8 (25.8)
Buccal mucosa, gingiva, lip & palate	10 (100.0)	9 (81.8)	14 (77.8)	23 (74.2)

Universiti Malaya

The protein expression of AAT, AHSG, APOA1, CLU, and HP was examined by IHC (Figure 4.13-4.17). Normal liver tissue was used as the positive control in IHC for AAT, APOA1, and HP protein. Whereas normal placenta and tonsil tissue were used as the positive control in IHC of AHSG protein and CLU protein, respectively. The positive control tissue in IHC of AAT, AHSG, APOA1, and HP displayed strong staining in the cytoplasm. Whereby, the positive control tissue in IHC of CLU displayed strong membranous staining. All the negative control tissue of AAT, AHSG, APOA1, CLU, and HP showed an absence of cytoplasmic and nucleus staining.

Generally, AAT, AHSG, and APOA1 proteins were expressed in all the tissue samples. The epithelial cells of control, OPMD, and OSCC tissues exhibited strong membranous and granular cytoplasmic staining of AAT. The epithelial cells of OPMD and OSCC tissues demonstrated a stronger membranous and granular cytoplasmic staining of AHSG and APOA1 than in control tissues. Whereas the epithelial cells of control, OPMD, and OSCC tissues exhibited membranous and granular cytoplasmic staining of CLU and HP. Nevertheless, some of the OPMD and OSCC tissue samples displayed weak or no staining of CLU and HP.

Universiti Malaya

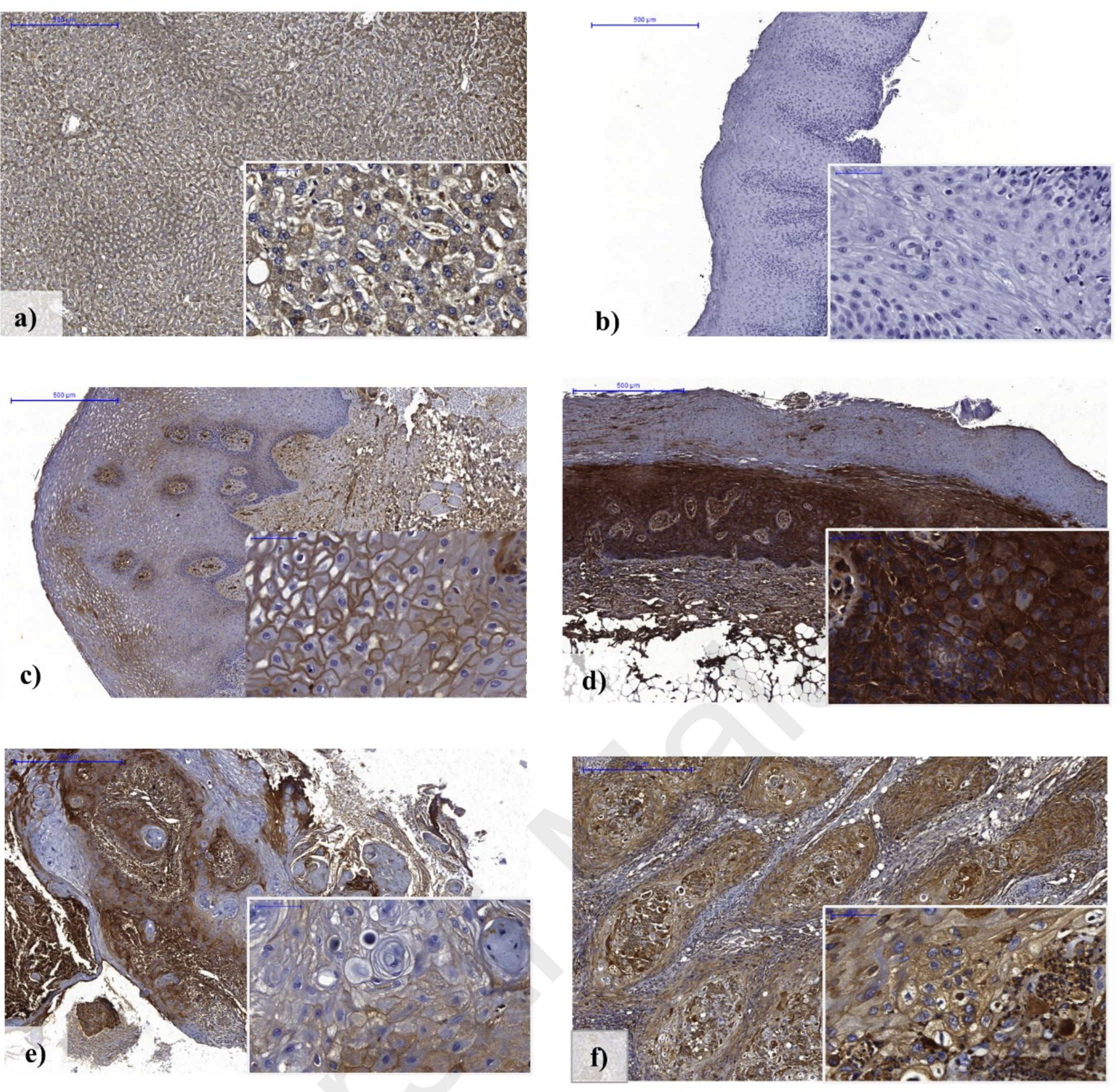


Figure 4.13: Photomicrograph of AAT expression in (a) positive control, (b) negative control, (c) normal oral mucosa, (d) OPMD, (e) early OSCC, and (f) advanced OSCC

Original magnification, 50x (a-f) and 400x (inserts in a-f).

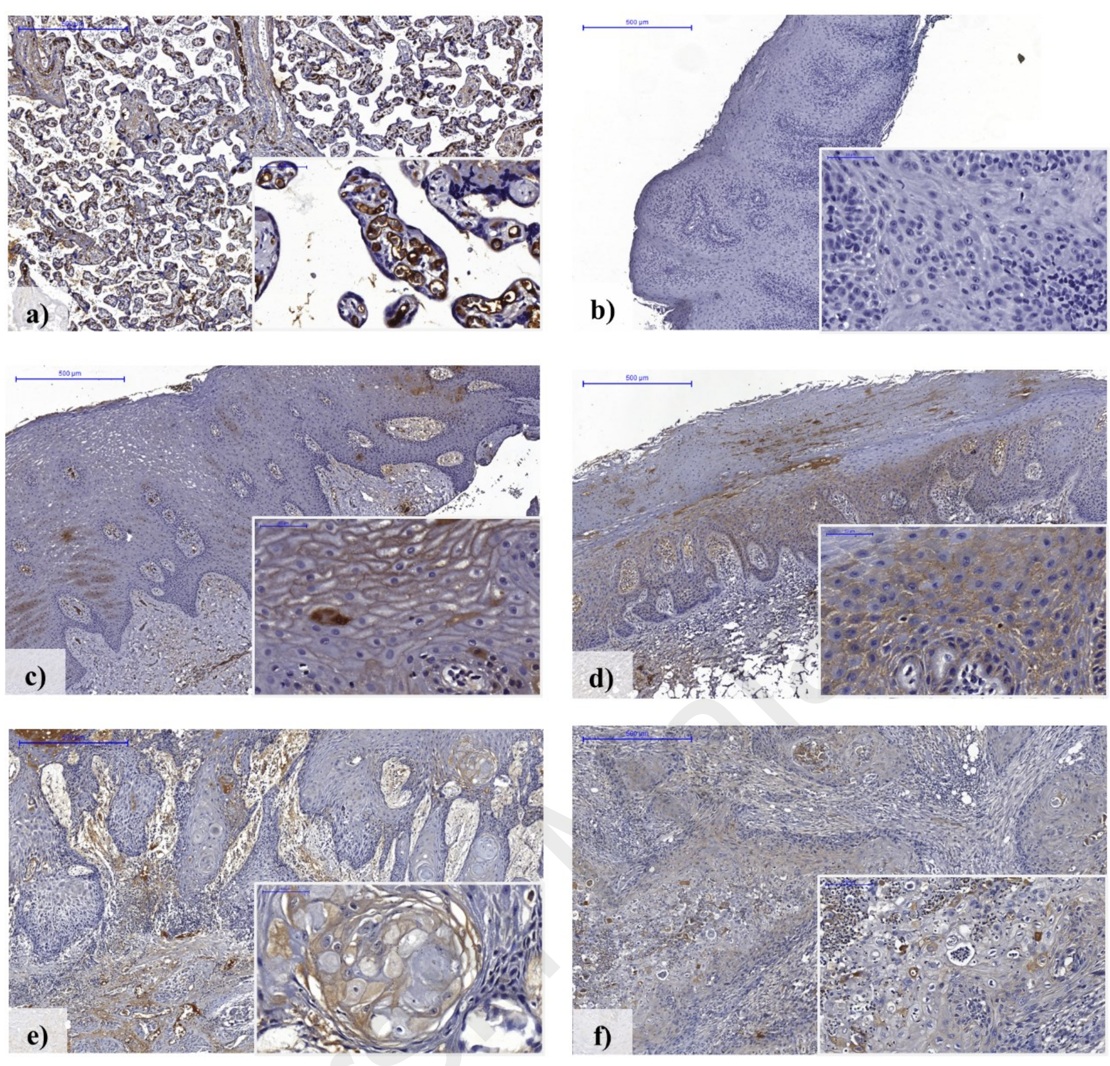


Figure 4.14: Photomicrograph of AHSG expression in (a) positive control, (b) negative control, (c) normal oral mucosa, (d) OPMD, (e) early OSCC, and (f) advanced OSCC

Original magnification, 50x (a-f) and 400x (inserts in a-f).

University

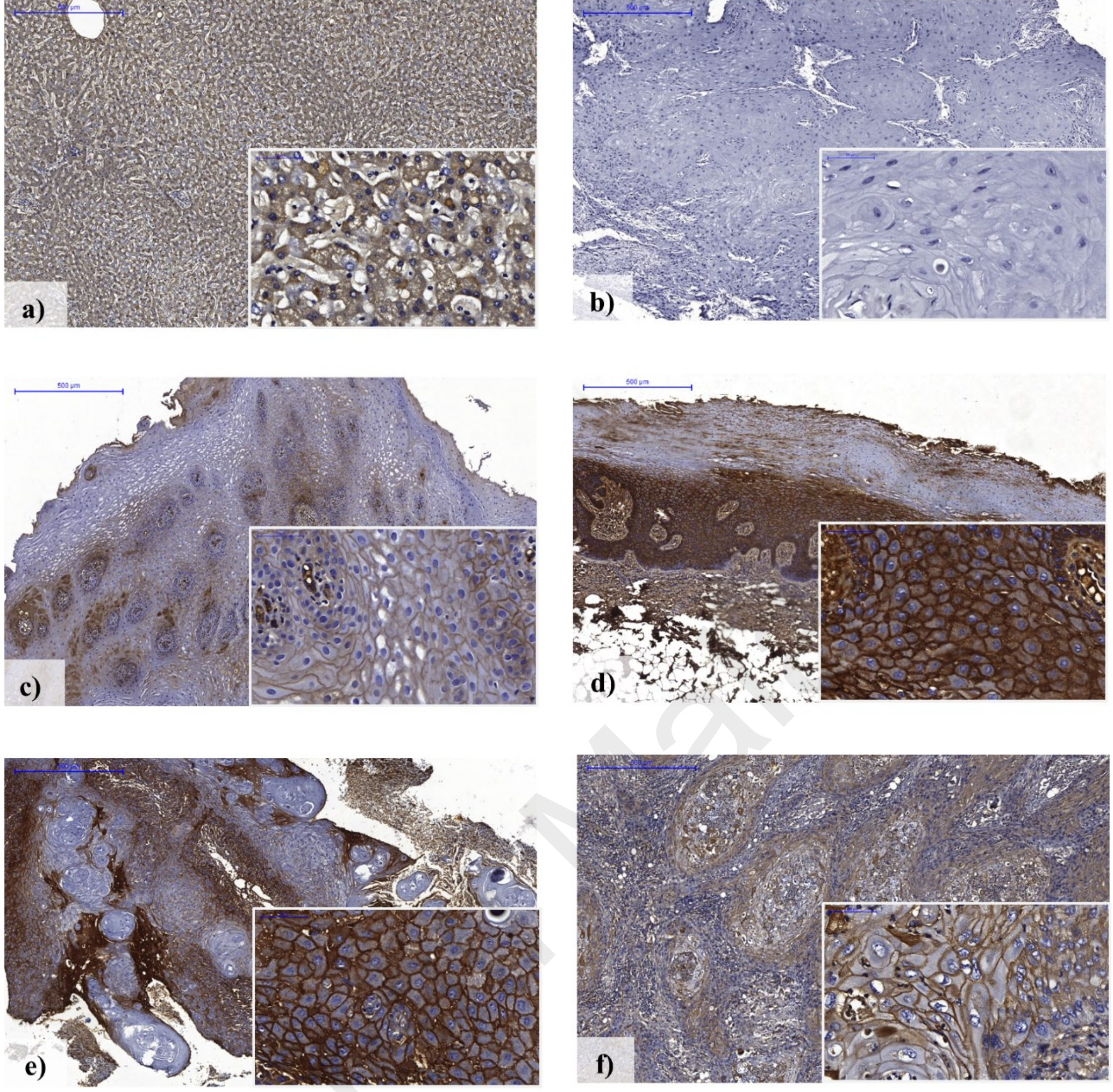


Figure 4.15: Photomicrograph of APOA1 expression in (a) positive control, (b) negative control, (c) normal oral mucosa, (d) OPMD, (e) early OSCC, and (f) advanced OSCC

Original magnification, 50x (a-f) and 400x (inserts in a-f).

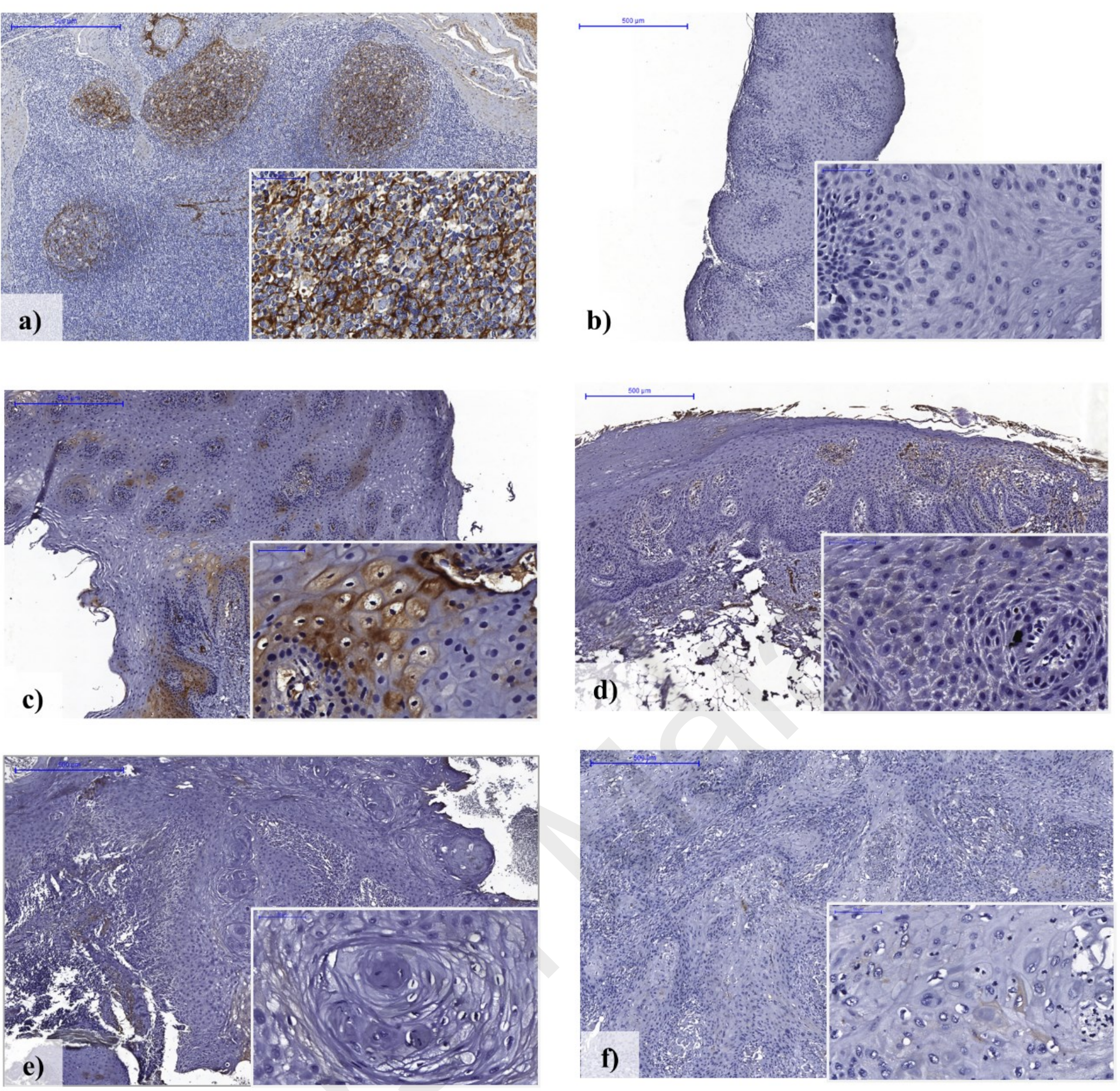


Figure 4.16: Photomicrograph of CLU expression in (a) positive control, (b) negative control, (c) normal oral mucosa, (d) OPMD, (e) early OSCC, and (f) advanced OSCC

Original magnification, 50x (a-f) and 400x (inserts in a-f).

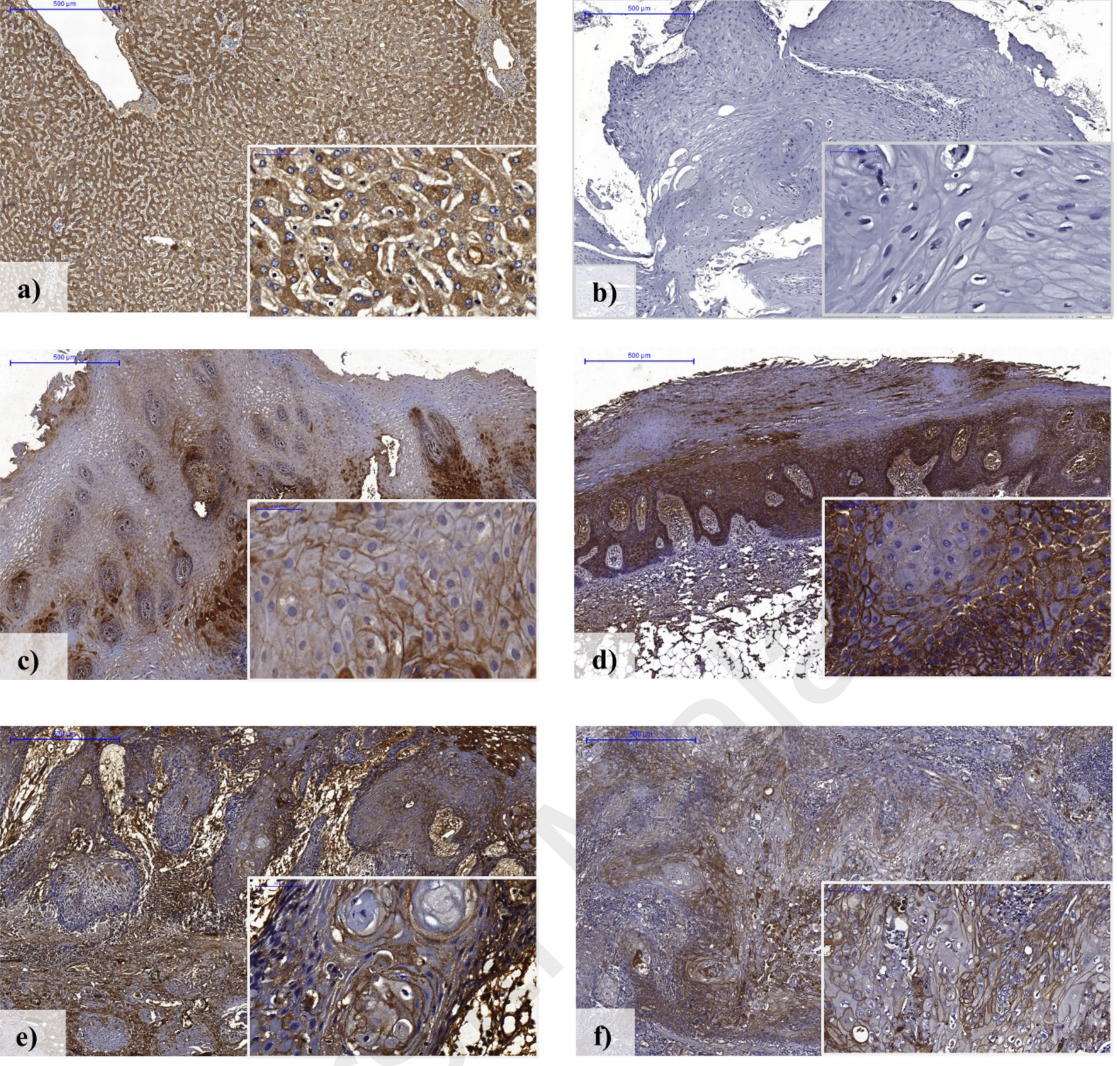


Figure 4.17: Photomicrograph of HP expression in (a) positive control, (b) negative control, (c) normal oral mucosa, (d) OPMD, (e) early OSCC, and (f) advanced OSCC

Original magnification, 50x (a-f) and 400x (inserts in a-f).

ROC analysis was performed to establish the cut-off value with the IHC immuno-reactive score. The low and high expression levels of AAT, AHSG, APOA1, CLU, and HP for each study group were determined with optimal cut-off value. The results of ROC analysis for AAT, AHSG, APOA1, CLU, and HP are shown in Table 4.20.

Table 4.20: Optimal cut-off value and area under the curve (AUC) in ROC analysis

	Optimal cut-off score	Sensitivity (%)	Specificity (%)	AUC (95%CI)	p-value
AAT	7.50	55.0	70.0	0.592 (0.38-0.80)	0.107
AHSG	2.50	55.1	70.0	0.641 (0.46-0.82)	0.092
APOA1	9.75	67.3	60.0	0.617 (0.43-0.81)	0.245
CLU	1.50	70.0	95.9	0.833 (0.66-1.00)	0.001
HP	10.50	60.0	89.8	0.804 (0.66-0.95)	0.003

The association between the expression level of AAT, AHSG, APOA1, CLU, and HP with the study groups was conducted using univariate and multivariate logistic regression analyses (Table 4.21 and Table 4.22). There was no significant difference in AAT protein expression of OPMD ($p = 0.262$) and OSCC ($p = 0.159$) compared with control. Nonetheless, the expression of AAT protein was noticed to be visually highest in OSCC, followed by OPMD and control, which were 55.1%, 54.5%, and 30.0%, respectively.

There was also no significant difference in AHSG protein expression of OPMD ($p = 0.131$) and OSCC ($p = 0.159$) as well as APOA1 protein expression of OPMD ($p = 0.507$) and OSCC ($p = 0.114$) compared with control. However, the expression of AHSG protein

Universiti Malaya

was noticed visually highest in OPMD, followed by OSCC and control, which were 63.6%, 55.1%, and 30.0%, respectively. Whereas the expression of APOA1 protein was also noticed visually highest in OSCC, followed by OPMD and control, which were 67.3%, 54.5%, and 40.0%, respectively.

There was a significant difference in CLU protein expression of OSCC ($p < 0.001$) compared with control as well as OSCC compared with OPMD ($p = 0.002$). The expression of CLU protein was noticed visually highest in control, followed by OPMD and OSCC, which were 70.0%, 45.5%, and 4.1%, respectively. Similarly, there was a significant difference in HP protein expression of OSCC compared with control ($p = 0.001$). The expression of HP protein was noticed visually highest in control, followed by OPMD and OSCC, which were 60.0%, 18.2%, and 4.1%, respectively.

Universiti Malaya

Table 4.21: Association analysis between the expression level of AAT, AHSG, and APOA1 with the study groups

	AAT				AHSG				APOA1			
	Univariate		Multivariate*		Univariate		Multivariate*		Univariate		Multivariate*	
	OR (95%CI)	p-value	OR (95%CI)	p-value	OR (95%CI)	p-value	OR (95%CI)	p-value	OR (95%CI)	p-value	OR (95%CI)	p-value
OPMD vs Control+	2.80 (0.46-16.93)	0.262	N/C	N/C	4.08 (0.66-25.38)	0.131	N/C	N/C	1.80 (0.32-10.20)	0.507	N/C	N/C
OSCC vs Control+	2.86 (0.66-12.39)	0.159	N/C	1.000	2.86 (0.66-12.39)	0.159	N/C	0.993	3.09 (0.76-12.53)	0.114	N/C	0.995
Early OSCC vs Control+	2.33 (0.45-12.00)	0.311	N/C	1.000	4.67 (0.88-24.80)	0.071	N/C	0.996	1.88 (0.39-9.01)	0.433	N/C	0.997
Advanced OSCC vs Control+	3.23 (0.70-14.91)	0.133	N/C	1.000	2.19 (0.48-10.05)	0.314	N/C	1.000	4.31 (0.96-19.31)	0.056	N/C	1.000
OSCC vs OPMD+	1.02 (0.28-3.80)	0.973	1.02 (0.19-5.31)	0.986	0.70 (0.18-2.71)	0.607	0.320 (0.05-2.24)	0.251	1.73 (0.46-6.50)	0.424	1.162 (0.22-6.07)	0.859
Early OSCC vs OPMD+	0.83 (0.19-3.75)	0.812	N/C	0.988	1.14 (0.24-5.50)	0.868	0.00 (0.00-N/C)	0.998	1.04 (0.23-4.70)	0.958	0.42 (0.05-3.39)	0.410
Advanced OSCC vs OPMD+	1.15 (0.29-4.61)	0.839	0.87 (0.11-6.71)	0.890	0.54 (0.13-2.21)	0.388	0.10 (0.01-1.60)	0.102	2.40 (0.57-10.05)	0.232	1.32 (0.18-9.70)	0.784
Advanced OSCC vs Early OSCC+	1.39 (0.43-4.45)	0.565	1.25 (0.35-4.41)	0.730	0.47 (0.14-1.57)	0.219	0.48 (0.13-1.75)	0.269	2.30 (0.67-7.86)	0.184	2.39 (0.59-9.60)	0.22

*Adjusted p-value and OR (odd ratio) for age, gender, ethnic group, smoking, alcohol drinking, and betel quid chewing, +Reference category, CI: confidence interval, bold: p-value < 0.05, N/C: OR (95%CI) and/or p-value not calculated due to cells containing number of observations of less than 5

Universiti Malaya

Table 4.22: Association analysis between the expression level of CLU and HP with the study groups

	CLU*				HP			
	Univariate		Multivariate-		Univariate		Multivariate*	
	OR (95%CI)	p-value	OR (95%CI)	p-value	OR (95%CI)	p-value	OR (95%CI)	p-value
OPMD vs Control+	0.36 (0.06-2.16)	0.262	N/C	N/C	0.15 (0.02-1.08)	0.952	N/C	N/C
OSCC vs Control+	0.02 (0.00-0.13)	**< 0.001**	N/C	0.998	0.08 (0.02-0.36)	**0.001**	N/C	0.999
Early OSCC vs Control+	0.03 (0.00-0.28)	**0.003**	N/C	0.997	0.08 (0.01-0.58)	**0.012**	N/C	0.998
Advanced OSCC vs Control+	0.01 (0.00-0.16)	**0.001**	N/C	0.999	0.07 (0.01-0.41)	**0.003**	N/C	1.000
OSCC vs OPMD+	0.05 (0.01-0.32)	**0.001**	0.01 (0.00-0.17)	**0.001**	0.51 (0.09-3.06)	**0.034**	0.16 (0.02-1.57)	0.116
Early OSCC vs OPMD+	0.07 (0.01-0.73)	**0.026**	0.02 (0.00-0.45)	**0.013**	0.56 (0.07-4.70)	0.595	0.13 (0.00-2.18)	0.154
Advanced OSCC vs OPMD+	0.04 (0.00-0.41)	**0.007**	0.01 (0.00-0.19)	**0.003**	0.48 (0.07-3.36)	0.461	0.15 (0.01-1.64)	0.119
Advanced OSCC vs Early OSCC+	0.57 (0.03-9.65)	0.695	0.48 (0.13-1.75)	0.269	0.87 (0.13-5.86)	0.873	0.78 (0.11-5.64)	0.809

-Adjusted p-value and OR (odd ratio) for gender, ethnic group, smoking, alcohol drinking, and betel quid chewing (number of confounding factors is excluded due to cells containing number of observations of less than 5), N/C: OR (95%CI) not calculated due to cells containing number of observations of less than 5

*Adjusted p-value and OR (odd ratio) for age, gender, ethnic group, smoking, alcohol drinking, and betel quid chewing, +Reference category, CI: confidence interval, bold: p-value < 0.05, OR (95%CI) and/or p-value not calculated due to cells containing number of observations of less than 5

Universiti Malaya

Chi-square analysis was carried out to examine the association of socio-demographic and clinicopathological characteristics with AAT, AHSG, APOA1, CLU, and HP expression in OSCC (Appendix E). There was no significant association between AAT, AHSG, APOA1, CLU, and HP expression and the characteristics. However, the level of AAT expression was higher in the older age group (58.1%), Indian (59.1%), betel quid chewing habit (57.1%), and advanced OSCC (58.1%). Whereas the level of AHSG expression was noticed higher in the older age group (53.5%), Indian (54.5%), and betel quid chewing habit (60.0%).

The level of APOA1 expression was also higher in the older age group (62.8%), Indian (68.2%), betel quid chewing habit (71.4%), and advanced OSCC (74.2%). In contrast to the expression level of AAT, AHSG, and APOA1, the expression level of CLU was lower in the older age group (95.3%), Indian (95.5%), betel quid chewing habit (94.3%), and advanced OSCC (96.8%). Similarly, the level of HP expression was also lower in the older age group (88.4%), Indian (88.6%), betel quid chewing habit (85.7%), and advanced OSCC (90.3%).

The follow-up time for OSCC patients in this study was used for overall survival analysis, which ranged from 2 to 89 months (mean: 29 months). However, there was no significant difference between the low and high expression levels of AAT, AHSG, APOA1, CLU, and HP in Kaplan-Meier overall survival analysis (Figure 4.18).

Universiti Malaya

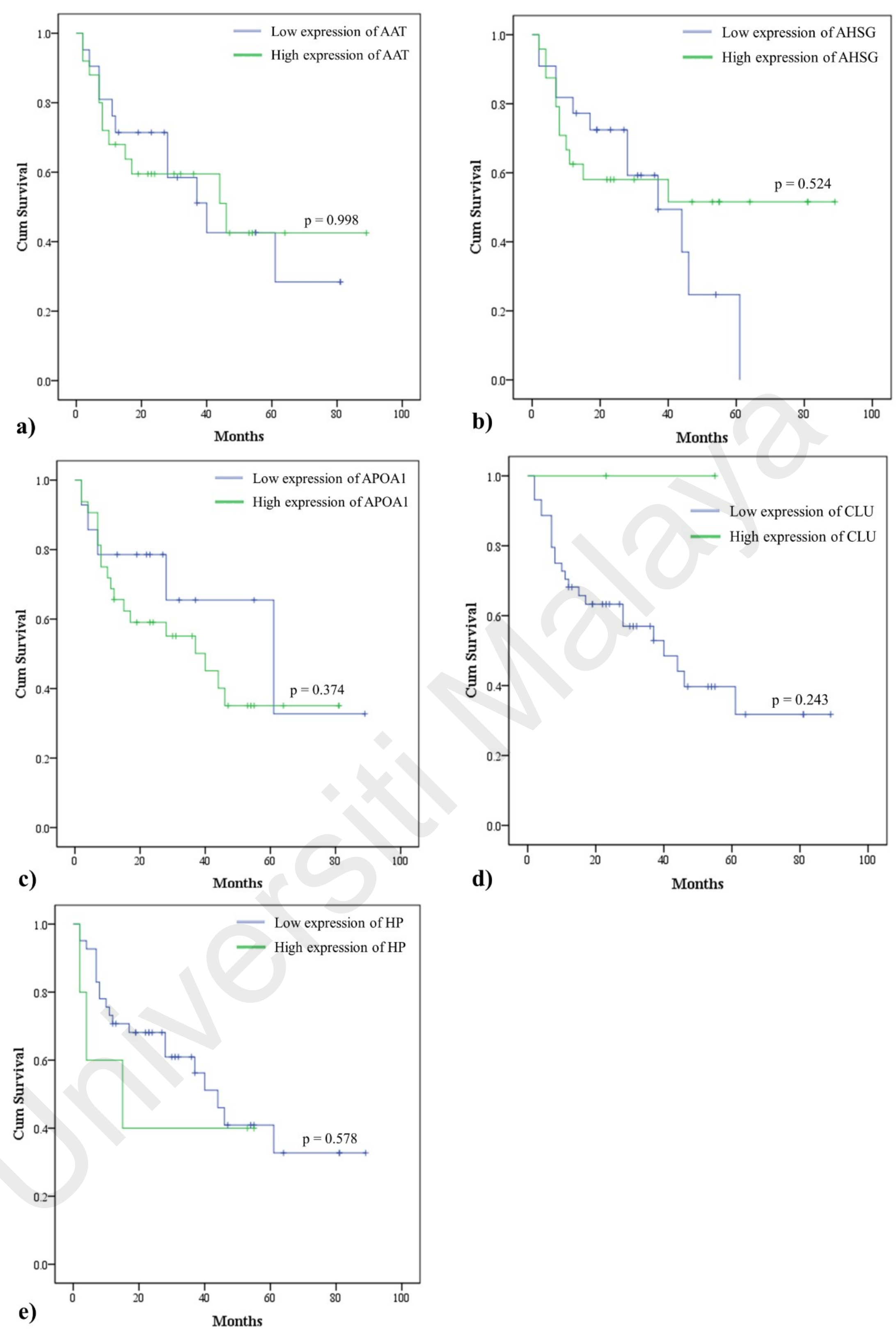


Figure 4.18: Kaplan-Meier overall survival curve for the expression of (a) AAT, (b) AHSG, (c) APOA1, (d) CLU, and (e) HP in OSCC

Universiti Malaya

CHAPTER 5: DISCUSSION

OSCC remains to be one of the major health concerns accounting for an estimated 354,864 new cases and 177,384 deaths worldwide (Ferlay et al., 2018). The incidence and prevalence rate of OSCC has been increasing over the years globally (Gupta et al., 2016). It is the most common type of cancer in South and Southeast Asian region due to the high incidence rates (Gupta et al., 2016; Johnson et al., 2011). Furthermore, the global incidence rate of OSCC has been predicted an increase of 48.5% by 2035 compared to those reported in 2012 due to the changes in the demographics (Shield et al., 2017). Therefore, detecting OSCC at an early stage is vital to improve the quality of life as well as to increase the survival rate of patients.

To date, physical examination of the oral cavity, followed by tissue biopsy remains the gold standard for OSCC diagnosis. However, there are some limitations associated with tissue biopsies such as the risk of sampling bias, resource requirement, and patient morbidity (Yang et al., 2018). Although many non-invasive screening tests have been developed, such as toluidine blue staining, oral brush biopsy, and optical biopsy, these examinations showed poor diagnostic accuracy and limited by false positive results (Yang et al., 2018). Thus, there is an urgent need for more effective biomarkers, which may assist in the screening, diagnosing, and/or monitoring OSCC. Numerous studies of proteomics and glycoproteomics have demonstrated that these fields can make notable contributions to the discovery of cancer-associated biomarkers (Drake et al., 2010; Li et al., 2017; Pan et al., 2011; Shah et al., 2015). Therefore, serum proteomics and glycoproteomics analyses were performed simultaneously in control (healthy volunteers), OPMD, and OSCC groups to identify potential biomarkers of OSCC.

Universiti Malaya

5.1 Analysis of Serum Protein and Glycoprotein Profiles Using Proteomics and Glycoproteomics Approaches

5.1.1 Analysis of Serum Protein Profiles Using Proteomics Approach

Serum remains the commonly used biological fluid for biomarker identification due to its minimally invasive nature, and it reflects the pathological condition of a disease. It is known that serum contains a mixture of proteins with a wide dynamic range (Anderson & Anderson, 2002). When depletion strategy is applied to remove high abundance proteins in serum, it may induce biases that can impair technical reproducibility and cause the loss of some non-targeted proteins that might interact strongly with the high abundance proteins (Bellei et al., 2011; Polaskova, Kapur, Khan, Molloy, & Baker, 2010; Tu et al., 2010).

It has also been reported that albumin depletion could cause loss of low abundance cytokine proteins that may affect subsequent proteomics analysis (Granger, Siddiqui, Copeland, & Remick, 2005). Additionally, albumin depletion could also lead to the loss of 35 albumin-associated proteins and most of them have been previously recognized as potential biomarkers (Gundry, Fu, Jelinek, Van Eyk, & Cotter, 2007). At the same time, different isoforms of high abundance proteins such as albumin and immunoglobulins were detected after the removal of the targeted high abundance proteins due to the incomplete depletion of the target proteins (Echan, Tang, Ali-Khan, Lee, & Speicher, 2005; Stempfer, Kubicek, Lang, Christa, & Gerner, 2008). Therefore, profiling of serum proteins on OSCC was conducted without the removal of high abundance proteins. This profiling study can be valuable in understanding the disease pathology and provide clues for the identification of potential biomarkers.

Analysis of protein profiles using 2-DE and MS analysis revealed a total of 20 differentially expressed proteins in OPMD and OSCC when compared with control.

Universiti Malaya

These 20 proteins (AAT, AINX, ALB, AMBP, APOA1, ARF, C3, CLU, HP, IGHA2, IGHG2, IGKC, LRG1, PRDX2, RAD50, RBP4, SAMP, TF, VCL, and VDBP) were selected for analysis because they were well-resolved in the 2-DE gels. Among those significantly identified proteins in MS analysis, a total of 9 proteins (AAT, APOA1, C3, IGHA2, IGHG2, IGKC, SAMP, TF, and VDBP) were up-regulated and 7 proteins (ALB, AMBP, CLU, HP, LRG1, PRDX2, and RBP4) were down-regulated in OPMD and OSCC when compared with control.

Based on the PCA plots, OPMD and OSCC were well differentiated from the control group. Although there were no differences found between OPMD and OSCC in the PCA plot, it had shown that OPMD was associated with the development of OSCC. Most cases of OSCC developed from malignant transformation of OPMD (Dionne et al., 2015; Speight et al., 2018; Warnakulasuriya et al., 2008). The present findings found a total of 10 differentially expressed proteins were significantly detected in OPMD when compared with control. These proteins included 5 up-regulated proteins (AAT, APOA1, IGKC, SAMP, and VDBP) and 5 down-regulated proteins (AMBP, CLU, HP, PRDX2, and RBP4). While a total of 15 differentially expressed proteins were significantly detected in OSCC when compared with control. There were 4 up-regulated proteins (IGHA2, IGHG2, IGKC, and TF) and 5 down-regulated proteins (ALB, AMBP, CLU, HP, and LRG1) found in early OSCC. Whereas 5 proteins (AAT, APOA1, C3, IGHG2, and VDBP) were up-regulated and only one protein (PRDX2) was down-regulated in advanced OSCC.

The detected proteins in this study were compared with those that have been reported from previous studies to ensure the accuracy of the protein profiles analysis. Some of the differentially expressed proteins, such as C3, LRG1, TF, and AMBP that were detected in the present findings were consistent with the previous OSCC proteomics studies (Chen

Universiti Malaya

et al., 2014; Csosz et al., 2018; Jessie et al., 2013; Sekikawa et al., 2018). Protein C3 is a key mediator of the innate immune system and plays a role in the activation of the complement system. Previous studies have demonstrated similar findings of C3 up-regulation in the serum and saliva samples of OSCC patients (Chen et al., 2014; Csosz et al., 2018; Jessie et al., 2013; Kawahara et al., 2016). The up-regulation of C3 was also found in pancreatic, lung, and skin cancers (Chen, Wu, et al., 2013; Lin, He, et al., 2014; Riihila et al., 2017). The activation of complement system promotes tumour growth by increasing cell proliferation and regulating the immune response to the tumour (Afshar-Kharghan, 2017).

Protein LRG1 belongs to the leucine-rich repeat family that is commonly known to be implicated in protein-protein interaction, signal transduction, and cell adhesion. In accordance with the present findings, LRG1 was also found to be down-regulated in the serum samples of OSCC (Chen et al., 2014) and tissue samples of HNSCC patients (Wang et al., 2017) as well as in the milk samples of breast cancer patients (Aslebagh, Channaveerappa, Arcaro, & Darie, 2018). However, contrary to the present findings, LRG1 was reported to be up-regulated in the plasma and saliva of OSCC patients (Chang, Lin, et al., 2019; Kawahara et al., 2016). The disparity between the present findings with the previous OSCC studies could be due to differences in the biological sample sources and the detection methods. Furthermore, the up-regulation of LRG1 was also observed in several cancers such as ovarian, pancreatic, colorectal, and oesophageal (Andersen et al., 2010; Furukawa et al., 2015; Zhang, Huang, et al., 2018; Zhao, Fan, et al., 2015). Thus, the alterations of LRG1 is believed to play a role in the development of OSCC. Nonetheless, the involvement of LRG1 in the development and progression of OSCC remains unclear and needs further investigation.

Universiti Malaya

Protein TF is the iron-binding transport protein that is responsible for the regulation of iron and haem synthesis in the biological fluids. Iron promotes DNA synthesis and is essential for cell growth and proliferation. Moreover, TF is known to be involved in stimulating the proliferation of cancer cells (Lee, Liu, & Huang, 2014). The up-regulation of TF has been reported in lung, breast, and ovarian cancers (Linxweiler et al., 2014; Milioli et al., 2015; Swiatly et al., 2018). The up-regulation of TF has also previously been observed in the saliva and tissue samples of OSCC patients (Jessie et al., 2013; Jou et al., 2010; Roman et al., 2013). Furthermore, the increase of the salivary TF level in OSCC patients is strongly associated with the size and stage of the tumour (Jou et al., 2010). Thus, the up-regulation of TF is believed to have an important role in oral carcinogenesis.

Protein AMBP is a precursor that proteolytically processed into two different functional proteins (alpha-1-microglobulin and bikunin) and involved in cell adhesion, immune response as well as in the regulation of biological processes. Recently, the up-regulation of AMBP was reported in the saliva samples of OPMD and OSCC patients (Shan et al., 2019). Nevertheless, the present findings are in line with a recent study that showed AMBP was down-regulated in the OSCC cell lines as well as in the tissue samples of OSCC patients (Sekikawa et al., 2018). These variations could be due to the variability of biological sample sources and differences in the methods of detection. Despite the dysregulation, the down-regulation of AMBP was suggested to be involved in the cell invasion and metastasis of OSCC because it was significantly related to cervical lymph node metastasis and poor overall survival (Sekikawa et al., 2018). Several studies have shown that AMBP was down-regulated in prostate and bladder cancers as well (Fan et al., 2011; Haj-Ahmad, Abdalla, & Haj-Ahmad, 2014; Su, Sheng, Chen, Tsay, & Wu, 2016).

Universiti Malaya

Protein VDBP that belongs to albumin family is the major transport protein in vitamin D metabolism. The present findings found VDBP was up-regulated in OPMD and advanced OSCC when compared with control. Conversely, VDBP was down-regulated in the plasma samples of OSCC patients (Tung et al., 2013). One of the possible disparities in this study could be due to the possibilities of OSCC having different regulatory mechanism since the mentioned study used plasma samples of OSCC patients with betel nut chewing habit for at least 5 years. Nevertheless, VDBP has been shown to be up-regulated in serum from mice bearing orthotopic human OSCC (Bijian et al., 2009). To date, the up-regulation of VDBP was notably identified in the serum samples of OPMD patients from the present findings, which has not been reported before in OPMD. Furthermore, VDBP was found to be up-regulated in ovarian, kidney, breast, and lung cancers (Huang et al., 2018; Kratzer, Weinstein, Albanes, & Mondul, 2019; Pawlik et al., 2006; Wang et al., 2012). Thus, VDBP may facilitate the carcinogenesis and cancer progression via its involvement in bone homeostasis, fatty acid binding, actin scavenging, and the modulation of immune and inflammatory responses (Speeckaert, Speeckaert, van Geel, & Delanghe, 2014). VDBP could enhance cancer cell aggressiveness via the activation of the insulin-like growth factor-1/insulin-like growth factor binding protein-2/Akt (IGF-1/IGFBP-2/Akt) pathway (Huang et al., 2018).

Additionally, several proteins, which include SAMP and RBP4 were found uniquely expressed in OPMD. Protein SAMP is a member of the pentraxin family that regulates several aspects of the immune response. For the first time, the up-regulation of SAMP was detected in the serum samples of OPMD patients. Previously, the up-regulation of SAMP was observed in the saliva samples of OSCC patients (Chen et al., 2017). Furthermore, the up-regulation of SAMP has also been detected in ovarian, lung, and breast cancers (Shield-Artin et al., 2012; Uribarri et al., 2014; Zeng et al., 2011). However, one previous study has demonstrated the down-regulation of SAMP in the

Universiti Malaya

tissue samples of OSCC when compared with oral leukoplakia using the web-based tool (Wang et al., 2009). This has indirectly shown that SAMP could involve in the malignant transformation of OSCC. Therefore, SAMP was observed as one of the cancer-associated proteins in the mentioned studies.

Protein RBP4 belongs to the lipocalin family and is the major transporter for retinol in the blood. It has been reported that RBP4 was up-regulated in the serum and saliva samples of OSCC patients (Chen et al., 2014; Csosz et al., 2018). Some studies have shown RBP4 was up-regulated in cancers including breast and lung cancers (Jiao, Cui, Ma, Li, & Si, 2016; Saleem, Raza, & S, 2019). Whereas other studies had found RBP4 was down-regulated in colorectal, liver, and ovarian cancers (Fei et al., 2017; Kataria et al., 2016; Lorkova et al., 2012). Likewise, the role of RBP4 in cancers including OSCC remains uncertain. Based on the present findings, although RBP4 was not detected in OSCC, RBP4 was found to be down-regulated in OPMD. A previous study indicated that several methylated genes including *RBP4* were associated with a high risk of malignant transformation of oral leukoplakia into OSCC (Abe et al., 2016). Therefore, the down-regulation of RBP4 in OPMD could be due to methylation at the gene level.

5.1.2 Analysis of Serum Glycoprotein Profiles Using Glycoproteomics Approach

The glycoproteomics analysis was also performed in the present study to detect glycoproteins with alteration in the post-translational modifications. Glycosylation is one of the most common forms of PTM that regulates many biological processes. *N*- and *O*-glycosylation are the major types of glycosylation involved in the maintenance of protein folding and conformation, protein stability, protein mobility as well as protein interaction and signal transduction (Roth et al., 2012). Furthermore, aberrant glycosylation has been associated with the hallmarks of cancer (Meany & Chan, 2011; Munkley & Elliott, 2016; Pinho & Reis, 2015). Many cancer-associated glycoproteins with aberrant glycosylation

Universiti Malaya

have been related in promoting malignant transformation, migration, invasion, and metastasis of cancer cells (Hakomori, 2002; Meany & Chan, 2011; Pinho & Reis, 2015). Con A and Jacalin lectins were used in the glycoproteomics analysis to understand the OSCC development and progression as well as to identify potential biomarkers since these lectins have high affinity to *N*-glycoproteins and *O*-glycoproteins, respectively (Dan et al., 2016; Lis & Sharon, 1998).

When the serum samples were subjected to Con A lectin affinity chromatography coupled with 2-DE analysis, a total of 4 *N*-glycoproteins were detected, namely AAT, AHSG, APOA1, and HP. Nevertheless, only AAT and APOA1 were found to be significantly up-regulated in the study groups. Whereas only 2 *O*-glycoproteins (AHSG and CLU) were detected when the serum samples were subjected to 2-DE coupled with Jacalin lectin blotting. There was no significant difference found between the expression levels of AHSG and CLU in OPMD and OSCC when compared with control.

Correspondingly, bioinformatics analysis was performed on the detected glycoproteins (AAT, AHSG, APOA1, CLU, and HP) in this study to ensure the reliability of the glycoproteomics analysis. The potential *N*- and *O*-glycosylation sites of the glycoproteins were predicted using the NetNGlyc 1.0 and NetOGlyc 4.0 servers, respectively (Gupta & Brunak, 2002; Steentoft et al., 2013). The *in silico* analysis demonstrated that AAT, AHSG, HP, and CLU possessed both potential *N*- and *O*-glycosylation sites. Although APOA1 was detected as *N*-glycoprotein in this study, the analysis only showed the potential *O*-glycosylation sites of APOA1. Nevertheless, each glycoprotein is heterogeneous with varying proportions of glycosylation sites (Apweiler et al., 1999; Roth et al., 2012).

More than half of all the proteins in human serum are glycosylated and it has been reported that nearly 90% of the glycoproteins are *N*-glycoproteins (Apweiler et al., 1999;

Universiti Malaya

Drake et al., 2006). Furthermore, many of these *N*-glycoproteins had been used as cancer biomarkers for diagnosis, monitoring, and prognostics purposes (Drake et al., 2006; Kirwan, Utratna, O'Dwyer, Joshi, & Kilcoyne, 2015). On the other hand, most of the glycoproteins detected in the present study were also shown to be well-resolved in the 2-DE gels. The expression level of AAT, APOA1, CLU, and HP was found to be up-regulated in both proteomics and glycoproteomics analyses, whereas the expression level of AHSG was up-regulated in the glycoproteomics analysis. Although most of the changes observed in the glycoprotein were generally associated with the differential protein expression, the changes at the glycosylation sites of the glycoproteins in some circumstances may not directly be associated with the changes in the protein level. The reason could be due to the complication and microheterogeneity of glycosylation sites in the glycoproteins (Ahn, Kim, & Yoo, 2015; Drake et al., 2010).

Most of the proteins and glycoproteins detected in the proteomics and glycoproteomics analyses were acute-phase proteins, which includes AAT, AHSG, ALB, APOA1, C3, CLU, HP, and LRG1. Acute-phase proteins had been defined as proteins that were primarily synthesized by the liver and released into circulation with increasing or decreasing the plasma concentrations by at least 25% in response to inflammation (Gabay & Kushner, 1999). Inflammation has been associated with many cancers, of which it promotes the initiation and progression of the cancer cells (Coussens & Werb, 2002). The aberrant glycosylated acute-phase proteins were also found to be related to the cancer-associated inflammation (Sarrats et al., 2010). These acute-phase proteins could have an impact on the tumour microenvironment, thereby facilitates cancer cell survival, proliferation, and metastasis (Dempsey & Rudd, 2012). Therefore, studies on protein and glycoprotein profiles are important to recognize the disease-associated changes in OSCC.

Universiti Malaya

AAT which is known as serine proteinase inhibitor A1 (SERPINA1), is the most abundant serine protease inhibitor in human serum. It is primarily synthesized by the liver. The main function of AAT is to regulate the proteolytic activity, and it has a role in immune-modulatory system and homeostasis (Guttman et al., 2015). AAT deficiency is a common inherited cause of liver and lung diseases (Greulich et al., 2016). Nevertheless, it has been reported that AAT was up-regulated in many cancers including colorectal, pancreatic, and breast cancers (Perez-Holanda, Blanco, Menendez, & Rodrigo, 2014; Wang et al., 2011; Zhou et al., 2010). Moreover, recent cancer studies had shown the up-regulation of AAT leads to cancer cell migration and invasion, which is associated with poor clinical prognosis (Ercetin et al., 2019; Zhao et al., 2018).

The altered glycosylation of AAT has also been demonstrated in lung, breast, and bladder cancers (Liang et al., 2015; Semaan et al., 2012; Yang et al., 2011). In these studies, changes in the expression levels of *N*-glycosylated AAT could be used in the monitoring and detection of cancer. This present finding concurs with the previous studies, in which AAT was up-regulated with aberrant glycosylation in OPMD and OSCC. Furthermore, the up-regulation of AAT was reported in the saliva samples of OSCC patients and it was also found to be related to the OSCC development (Csosz et al., 2018; Jessie et al., 2013; Kawahara et al., 2016).

AHSG or Fetuin A, is a multifunctional glycoprotein that involved in numerous normal and pathological processes such as vascular calcification, bone metabolism and mineralization, insulin resistance as well as cancer initiation and progression (Mori, Emoto, & Inaba, 2011). The up-regulation of AHSG has been reported previously in cervical, gastric, hypopharyngeal, and oesophageal cancers (Boichenko et al., 2014; Shi et al., 2018; Tian et al., 2015; Zhao, Fan, et al., 2015). *In vitro* studies shown that AHSG could germane to the progression and metastasis of breast cancer and HNSCC by

Universiti Malaya

promoting adhesion, proliferation, migration, and invasion of the cancer cells (Sakwe, Koumangoye, Goodwin, & Ochieng, 2010; Thompson et al., 2014).

AHSG was identified as a glycosylated protein in several diseases, including cancers. The *N*-glycosylated AHSG was found to be down-regulated in lung cancer (Rho et al., 2009; Tran et al., 2008). Whereas the *N*-glycosylated AHSG was up-regulated in cholangiocarcinoma, liver, and pancreatic cancers as well as chronic pancreatitis (Betesh et al., 2017; Comunale et al., 2006; Sarrats et al., 2010; Zhao, Qiu, Simeone, & Lubman, 2007). These studies revealed that the expression level of glycosylated AHSG is dependent on the regulation of other molecules in the pathological condition. On the other hand, the *O*-glycosylated AHSG was also detected in different types of cancer, including bladder, prostate, pancreatic, renal cell, and cholangiocarcinoma (Belczacka et al., 2019). Interestingly, both *N*- and *O*-glycosylated AHSG were detected in the present study. Indeed, AHSG contains both *N*- and *O*-glycosylation sites (Clerc et al., 2016). Although AHSG was not detected significantly in the present study, the expression level of AHSG and *N*-glycosylated AHSG was found to be up-regulated in the serum samples of mice bearing orthotopic OSCC as well as in the plasma and saliva samples of OSCC patients (Bijian et al., 2009; Chang, Lin, et al., 2019; Csosz et al., 2018). Thus, these findings have indicated that AHSG is associated with the development and progression of OSCC.

Recently, accumulating evidence suggests that alterations in apolipoproteins and lipid metabolism have been associated with cancer development and progression (Borgquist et al., 2016; Zamanian-Daryoush & DiDonato, 2015). APOA1 is the major protein component of high-density lipoprotein (HDL) that serves as an enzyme cofactor as well as responsible for cholesterol transportation and cellular cholesterol homeostasis (Mangaraj, Nanda, & Panda, 2016). The expression level of APOA1 varies in different types of cancer (Ren et al., 2019). Most studies have shown an inverse association of

Universiti Malaya

APOA1 level with the development and progression of cancers such as breast, ovarian, lung, and oesophageal cancers (Borgquist et al., 2016; Kim, Bae, Lim, Kim, & Ahn, 2012; Lihong et al., 2014; Wang, Li, et al., 2016). Meanwhile, few studies had revealed APOA1 is up-regulated in patients with bladder, colorectal, and gastric cancers (Li et al., 2014; Lim et al., 2016; Shi et al., 2018).

APOA1 was detected as *N*-glycoprotein in the study although no potential sites of *N*-glycosylation for APOA1 were predicted (Table 4.14). However, APOA1 was identified as *N*-glycoprotein in a previous study (Bunkenborg, Pilch, Podtelejnikov, & Wisniewski, 2004). Furthermore, the up-regulation of *N*- and *O*-glycosylated APOA1 has been observed in liver cancer (Chaerkady et al., 2008) and gastric cancer (Campos et al., 2015), respectively. This evidence showed that APOA1 contains both *N*- and *O*-glycosylation sites.

Previous studies have noted that the expression level of APOA1 and *N*-glycosylated APOA1 was down-regulated in plasma, tissue, and saliva samples of OSCC patients (Chang, Lin, et al., 2019; Liao et al., 2011; Tung et al., 2013). Nevertheless, APOA1 was found to be up-regulated in the saliva samples of oral leukoplakia patients as well as in the serum, saliva, and tissue samples of OSCC patients (Camisasca et al., 2017; Csosz et al., 2018; Lo et al., 2007), which is in line with the present findings. Additionally, a precursor of APOA1, proapolipoprotein was also up-regulated in OSCC and recurrent HNSCC (Chen et al., 2014; Gourin, Zhi, & Adam, 2009). The up-regulation of APOA1 and its precursor proapolipoprotein may indirectly promote tumour survival through kinase activation in recurrent HNSCC (Gourin et al., 2009). The inconsistent results of the APOA1 expression level in OSCC was also observed in breast cancer. Many studies had shown that the down-regulation of APOA1 was inversely associated with breast cancer, but the up-regulation of APOA1 was also positively associated with breast cancer

Universiti Malaya

in few other studies (Borgquist et al., 2016; His et al., 2014; Lin, Hong, et al., 2017; Martin et al., 2015; Ren et al., 2019). There is a possibility that the up-regulation of APOA1 was related to the inflammatory or immune response to cancers (Borgquist et al., 2016; Georgila, Vyrla, & Drakos, 2019). Thus, additional studies on the roles of APOA1 in cancer need to be conducted for further elucidation.

Another glycoprotein that was detected in the study is CLU, which is also known as apolipoprotein J. CLU is a heterodimeric disulfide-linked glycoprotein that found in various tissues and body fluids. It has been associated with numerous biological processes, including cell adhesion, cell cycle regulation, DNA repair, lipid transportation, tissue remodelling and apoptosis, membrane recycling, and immune system regulation (Shannan et al., 2006). The expression of CLU has been reported to be either up- or down-regulated in various cancers and was found to be involved in tumour growth and carcinogenesis. The up-regulation of CLU was reported in colorectal, breast, pancreatic, and bladder cancers (Bertuzzi et al., 2015; Flanagan, Whyte, Chatterjee, & Tenniswood, 2010; Jin et al., 2012; Makridakis et al., 2010; Redondo et al., 2010). Also, CLU was found to be down-regulated in different types of cancer such as non-small-cell lung cancer (squamous cell carcinoma type), testicular seminoma, and prostate cancers (Jin, Chen, Han, Luo, & Li, 2017; Liu, Han, et al., 2013; Scaltriti et al., 2004). These altered expressions of CLU in cancers had further shown that CLU could have a dual role as a tumour suppresser gene and oncogene (Kadam & Teni, 2016).

The expression level of glycosylated CLU was also observed to be either up- or down-regulated in cancers. The up-regulation of both *N*- and *O*-glycosylated CLU was previously reported in ovarian cancer (Miyamoto et al., 2016; Mu, Lim, Hashim, & Shuib, 2013). However, in the present study, *O*-glycosylated CLU was observed to be down-regulated in the serum of OPMD and OSCC patients compared with control. In line with

the present findings, numerous studies have shown that the expression level of glycosylated CLU was down-regulated in liver, colorectal, ovarian, and oesophageal cancers (Comunale et al., 2011; Rodriguez-Pineiro et al., 2004; Wu, Xie, Nie, Buckanovich, & Lubman, 2013; Zhang et al., 2003). Furthermore, the down-regulation of CLU has also been detected in the serum and saliva samples of OSCC patients (Chen et al., 2014; Hu et al., 2008). The discrepancy in the expression of CLU in cancers may be due to the presence of various glycosylation forms. It is known that CLU has two isoforms: secreted CLU and nuclear CLU, where both of these isoforms have a role in anti-apoptotic and pro-apoptotic function, respectively (Shannan et al., 2006). Nevertheless, the specific role of CLU in cancer development and progression remains ambiguous.

HP is a secreted acute-phase protein that is mainly produced by the liver. It is also a haemoglobin-binding protein, which plays a role in the clearance of free haemoglobin and the prevention of oxidative stress (Levy et al., 2010). Previous studies indicated that HP could act as an independent prognostic factor in various types of cancer. The up-regulation of HP is frequently detected in lung, colorectal, breast, and ovarian cancers, and is closely associated with the poor prognosis of the diseases (Lu et al., 2016; Sun et al., 2016; Tabassum, Reddy, & Mukherjee, 2012; Zhao, Annamalai, et al., 2007). Whereas the down-regulation of HP has been observed in thyroid, gastric, and liver cancers (Farrokhi Yekta, Arefi Oskouie, Rezaei Tavirani, Mohajeri-Tehrani, & Soroush, 2018; Liu et al., 2007; Tai et al., 2017). Thus, it is noted that the expression of HP is dysregulated in cancer development and progression. Recent evidence points that HP could induce the cancer cell migration, and cause the changes in the cell morphology and actin cytoskeleton structure (Garibay-Cerdenares, Hernandez-Ramirez, Osorio-Trujillo, Gallardo-Rincon, & Talamas-Rohana, 2015).

Universiti Malaya

Additionally, the aberrant glycosylation of HP has been discovered in various inflammatory disorders and cancers. The *N*-glycosylated HP has been revealed to be frequently up-regulated in pancreatic, liver, prostate, lung, ovarian, colorectal, breast, and gastric cancers (Zhang, Shang, Li, Qin, & Liu, 2016). The down-regulation of the *N*-glycosylated HP was found in liver cancer, which is consistent with the present findings (Qi et al., 2014). It is noted that HP consists of two α and β chains, of which only the β chains are glycosylated as it contains four *N*-glycosylation sites (Clerc et al., 2016). Therefore, an altered glycosylation feature of HP was observed as well. For example, HP with significantly increased fucosylation and decreased sialyl Lewis X (sLe^x) was detected in the serum samples of non-small cell lung cancer patients (Ferens-Sieczkowska, Kratz, Kossowska, Passowicz-Muszynska, & Jankowska, 2013). Also, the glycosylation status of HP may have been involved in the interaction with several receptors, such as CD163 and chemokine receptor 2 (CCR2) which are highly expressed on tumour-associated macrophages, and influence tumour growth and angiogenesis (Dempsey & Rudd, 2012).

The up-regulation of HP and *N*-glycosylated HP have been frequently detected in the serum, plasma, and saliva samples of OSCC patients (Chang, Lin, et al., 2019; Chen et al., 2014; Csosz et al., 2018; Jessie et al., 2013; Saraswat et al., 2018; Tung et al., 2013). However, the present findings appeared contradictory to those from previous studies. These conflicting results may be due to the different methodologies of measurement, sample sources, or stages of the disease. Nevertheless, HP was reported to be down-regulated in the tissue samples of OSCC patients (Thiel et al., 2011). The down-regulation of the HP α1 chain was also observed in HNSCC patients (Chen et al., 2008). Furthermore, it was noted that the expression level of HP and *N*-glycosylated HP was up-regulated in OSCC through the MS and Western blot analysis, while the expression level of HP was found to be down-regulated in OSCC especially in the advanced stage when

Universiti Malaya

subjected to ELISA analysis (Chang, Lin, et al., 2019). This has demonstrated the distinct and contradictory role of HP in OSCC. Nonetheless, little is known regarding the biological functions of HP in OSCC.

5.2 Functional Annotation and Protein Interaction Analysis

A list of proteins and glycoproteins that were associated with the development and progression of OSCC had been identified using integrated proteomics and glycoproteomics analyses. With the advancement of bioinformatics tools and algorithms, the dysregulated biological processes and pathways in OSCC were elucidated by exploring the identified proteins and glycoproteins in the biological domain knowledge resources. This could further improve the current understanding of the molecular basis in the development and progression of OSCC.

The functional annotation analysis using DAVID v6.8 was carried out to explicate the functional roles of the identified proteins and glycoproteins. The results revealed that these dysregulated proteins possessed various biological processes, molecular functions, and cellular components. Further examination found that most of these proteins were enriched in the extracellular space or region. The secreted proteins from the cancer cells could liberate into the microenvironment, and enter the blood and lymphatic circulation system, eventually contributing to cancer cell invasion, migration, and cell motility (Karagiannis, Pavlou, & Diamandis, 2010). Thus, the presence of these serum proteins could be secreted by the cancer cells within the extracellular environment. A total of 13 identified proteins and glycoproteins were also recognized as circulating blood microparticles. These blood microparticles could be involved in cancer progression and play a role in the promotion of angiogenesis, enhancement of metastatic potential, cellular survival, and evasion of immune surveillance (Gong, Jaiswal, Dalla, Luk, & Bebawy, 2015).

Universiti Malaya

The functional annotation clustering analysis revealed that platelet degranulation and activation of classical complement pathway are the prominent biological processes involved in the development and progression of OSCC. Generally, platelets have a functional role in haemostasis and thrombosis. However, studies have shown that platelets were also found to be associated with cancer progression. Several studies have revealed that platelet activation and degranulation, which were induced by cancer cells could promote tumour growth with the increase of angiogenesis, vascular permeability, and metastasis of cancer cells (Egan et al., 2011; Kopp, Placke, & Salih, 2009).

The activation of classical complement pathway is known to be involved in immune responses and homeostasis. Furthermore, the complement system is known as a central part of the immune system, which is the first defence against the pathogens. These present findings concur with previous studies, in which the activation of the complement system was found to be related to the progression of OSCC (Chen et al., 2014; Gallenkamp et al., 2018). Evidence also revealed that the development and progression of many cancers are accompanied by the activation of the classical complement pathway (Kwak et al., 2018; Markiewski et al., 2008; Vadrevu et al., 2014). The activation of the classical complement pathway could enhance the tumour growth by inducing angiogenesis, promoting cell proliferation and invasion as well as modulating the immune response to the tumour (Afshar-Kharghan, 2017).

The serine-type endopeptidase activity was found to be the most enriched molecular function in the functional annotation clustering analysis. Serine endopeptidase or serine protease is an enzyme that cleaves the peptide bonds in the proteins at the extracellular compartment. Under normal conditions, it involves in many physiological functions such as protein digestion, blood coagulation, complement activation, and immunity. Additionally, serine endopeptidase is implicated in cell proliferation, cell signalling,

Universiti Malaya

apoptosis, invasion, and metastasis of cancer cells (Poddar, Maurya, & Saxena, 2017). The serine-type endopeptidase activity has been reported previously in HNSCC, prostate, and bladder cancers (Islam et al., 2018; Keam et al., 2018; Lei et al., 2013). The dysregulation of serine-type endopeptidase activity may lead to the degradation of extracellular matrix and the imbalance of the cellular homeostasis in cancers (Poddar et al., 2017).

Apart from exploring the functional roles of the identified proteins and glycoproteins, IPA v7.1 was performed to elucidate the interaction networks and predominant canonical pathways of the proteins associated with OSCC. In the present study, IPA revealed numerous biological pathways associated with the development and progression of OSCC. Based on these findings, LXR/RXR activation and acute phase response signalling pathways were activated in OSCC. LXR, which classified as a transcription factor, belongs to the nuclear receptor family and is depends on RXR, its binding partner to form heterodimers. The activation of LXR/RXR pathway is generally implicated in cholesterol, lipid, and glucose metabolism as well as in the regulation of immune responses (A-Gonzalez & Castrillo, 2011; Schulman, 2017). The LXR/RXR activation pathway has also been previously described in thyroid, colorectal, and breast cancers (Li et al., 2019; Lim et al., 2016; Torres-Luquis et al., 2019). This had indicated that the LXR/RXR activation pathway is directly involved in the development and progression of cancer. Additionally, there is evidence indicating that the activation of LXR was related to the modification of phosphoinositide 3-kinases (PI3K)/Akt transduction pathway, which plays a pivotal role in cell proliferation, migration, and invasion (Dufour et al., 2013).

As mentioned above, the activation of LXR/RXR pathway is related to lipid metabolism. Intriguingly, IPA also revealed lipid metabolism was the major biological

Universiti Malaya

function under the "molecular and cellular functions" category. Hence, the activation of LXR/RXR pathway could lead to the alteration of lipid metabolism in OSCC as well. Lipid metabolism is a complex physiological process, involving the oxidation of fatty acids to synthesize lipids or produce energy in the cells. The alteration of lipid metabolism is related to carcinogenesis and cancer-associated cachexia (Santos & Schulze, 2012). Numerous studies had also revealed the implication of lipid metabolism in cervical, breast, prostate, and thyroid cancers (Guo et al., 2015; His et al., 2014; Li et al., 2019). The changes in the lipid metabolism are essential for malignant transformation as the cancer cells require a high fatty acid and cholesterol turnover rate to facilitate the survival and proliferation of these cancer cells (Maan, Peters, Dutta, & Patterson, 2018; Santos & Schulze, 2012). Nevertheless, the molecular mechanism of lipid metabolism in cancer progression has yet to be determined.

The activation of the acute phase signalling pathway was probably indicated the mechanism or relationship among inflammation, innate immunity system, and cancer (Coussens & Werb, 2002). The acute phase response is a rapid and non-specific response to local and systemic inflammation. Furthermore, chronic inflammation could promote cancer progression and has been well-known as a hallmark and cause of cancer (Hanahan & Weinberg, 2011). On the other hand, the major associated network functions of the identified proteins were found to be linked to the inflammatory response, organismal injury and abnormalities. The inflammatory response is in correspondence to the activation of signalling pathways that regulate inflammatory cells, such as neutrophils and macrophages at the site of tissue injury. Additionally, the dysregulation of nuclear factor kappa-light-chain-enhancer of activated B cells (NF-kB), mitogen-activated protein kinase (MAPK), and Janus kinase (JAK)/signal transducers and activators of transcription (STAT) signalling pathways are commonly related to inflammation and acute phase response, which may lead to carcinogenesis (Chen et al., 2018).

Universiti Malaya

The altered expression levels of acute-phase proteins and the activation of acute phase response signalling pathway had also been reported in many cancers including lung, cervical, colorectal, and pancreatic cancers (Ciereszko et al., 2019; Guo et al., 2015; Lim et al., 2016; Pang, Abdul-Rahman, Wan-Ibrahim, & Hashim, 2010; Saraswat et al., 2017). Most of the proteins and glycoproteins detected in this present study were acute-phase proteins since the inflammatory response often appears in cancer. Correspondingly, acute phase response was related with the modulation of the interleukin 6 (IL-6)/JAK/signal transducers and activators of transcription 3 (STAT3) signalling pathway, which is involved in the proliferation, progression, and metastasis of cancer cells (Johnson, O'Keefe, & Grandis, 2018). Thus, the dysregulation of these identified proteins revealed the activation of acute phase response signalling pathway is implicated in the development and progression of OSCC.

IPA had also revealed that the production of nitric oxide and reactive oxygen species in macrophages was inhibited in OSCC. The nitric oxide (NO) and reactive oxygen species (ROS) are relatively unstable nitrogen- or oxygen-reactive molecules and free radicals that contain one or more unpaired electrons. They can be released by the inflammatory cells such as macrophages and neutrophils in response to an inflammatory stimulus. Both NO and ROS have a role in cell signalling and homeostasis. The inhibition or activation of NO and ROS can yield different results in cancer cells. They can trigger altered activation of macrophages that promote tumour growth and immunosuppression in the tumour microenvironment (Brown, Recht, & Strober, 2017). On the other hand, studies have revealed that the oxidative and nitrosative stresses were closely related to the inflammation-associated carcinogenesis, and thus the production of NO and ROS in low level could trigger carcinogenesis by promoting malignancy transformation and cell proliferation (Aivaliotis et al., 2012; Reuter, Gupta, Chaturvedi, & Aggarwal, 2010).

Universiti Malaya

It has been reported that the production of NO was decreased in OSCC and it could increase tumour growth and proliferation (Avci et al., 2009). Additionally. the decrease in ROS production was found to be associated with the up-regulation of peroxidasin (PXDN) protein level that may promote cell proliferation and invasion in OSCC (Kurihara-Shimomura, Sasahira, Shimomura, & Kirita, 2020). Furthermore, accumulating evidence has revealed that NO and ROS were implicated in the development of HNSCC including OSCC (Hsieh, Chen, Lin, Chang, & Li, 2019; Kesarwala, Krishna, & Mitchell, 2016; Korde, Basak, Chaudhary, Goyal, & Vagga, 2011). Although the mechanism of NO and ROS affecting OSCC has not been fully explored, it is believed that NO and ROS could activate the signalling pathway of transcription factors including NF-kB and STAT3 during oxidative stress which can induce cancer cell survival, invasion, and angiogenesis (Aivaliotis et al., 2012; Reuter et al., 2010).

Other than the canonical pathways, IPA had also demonstrated the protein interaction of the identified proteins through the network analysis. The changes in the biological processes and functions were mediated through the protein interaction network that involved the signalling pathways regulation as well as the upstream and downstream effectors. The central nodes of the protein interaction network were involved in the regulation of NF-κB and ERK 1/2 signalling pathways. The transcription factor NF-κB plays an important role in various biological processes, such as inflammatory, immune response, cell proliferation, survival, and apoptosis (Chen et al., 2018). The activation of NF-κB signalling pathway, mediated by tumour necrosis factor-α (TNF-α) was related to the invasion, migration, and metastasis of OSCC (Tang et al., 2017). ERK1/2 are members of the MAPK family, in which ERK1 and ERK2 are also known as MAPK3 and MAPK1, respectively. MAPK is a family of serine/threonine protein kinases that can regulate cell proliferation, differentiation, cell survival, and apoptosis (Chen et al., 2018).

Universiti Malaya

Studies have shown that the activation of ERK1/2 signalling pathway, mediated by C-X-C chemokine receptor type 4 (CXCR4) and its ligand, CXCL12, together with crosstalk of STAT3 signalling pathway, could induce the regulation of cell cycle, cell proliferation, migration, and invasion in OSCC (Gkouveris, Nikitakis, Karanikou, Rassidakis, & Sklavounou, 2014; Yu, Wu, et al., 2011).

It is noted that tumour-associated macrophages play a major role in the development and progression of cancers. The protein interaction network generated by IPA has also revealed the involvement of tumour-associated macrophages in the development and progression of OSCC, where the macrophage markers, CD44 and CD163 were the predicted neighbouring molecules of the identified proteins. The up-regulation of CD163 was found to be associated with cancer stem cell markers, including CD40, and poor overall survival rate in OSCC (He et al., 2014). Additionally, IPA has predicted several neighbouring molecules in the protein interaction network such as ATM, BCL2, CDC25B, and caspase 3/7, which are associated with the regulation of the cell cycle and apoptosis. If DNA is damaged, ATM and checkpoint kinase 1 (CHEK1) will be activated, which then inhibits CDC25 activity that regulates the activation of cyclin-dependent kinases, and thus cell cycle progression is arrested for enabling the DNA repair (Medema & Macurek, 2012). Whereas BCL2 and caspase family of proteins that have a role in the tissue homeostasis and elimination of unwanted harmful cells were involved in apoptosis or programme cell death (Ola, Nawaz, & Ahsan, 2011). The down-regulation of ATM expression that leads to more aggressive clinicopathological features was found approximately in 31.0% of OSCC (He, Chen, & Li, 2008). Moreover, recent evidence has revealed the implication of apoptosis with the presence of BCL2 and caspase 3/7 in the development and progression of OSCC (Arya, Singh, & Daniel, 2016; Coutinho-Camillo et al., 2017). Hence, cell cycle and apoptosis have a significant function in the development and progression of OSCC.

5.3 Validation of Potential OSCC Biomarkers

In this study, proteomics and glycoproteomics analyses were employed for the discovery of potential biomarkers of OSCC. Taken together with the bioinformatics analysis, the identified proteins and glycoproteins were found to be associated with the development and progression of OSCC.

Based on the results obtained from the bioinformatics analysis and literature searches, five proteins, namely AAT, AHSG, APOA1, CLU, and HP were selected for further validation using antibody-based methods (ELISA and IHC). There were not many additional studies on these proteins although some of them have been reported previously in some OSCC studies. Both ELISA and IHC methods are also an important platform in proteomics research, especially in the validation of biomarkers to detect the targeted proteins (Solier & Langen, 2014).

ELISA remains a gold standard for the validation of biomarkers in body fluids such as serum and plasma. It allows quantitative measurement of these proteins (Solier & Langen, 2014). According to the ELISA analysis, AAT, AHSG, APOA1, CLU, and HP had shown significant differences across the study groups. The expression levels of AAT, AHSG, APOA1, CLU, and HP in the ELISA analysis were in concordance with the proteomics and glycoproteomics analyses. ROC analysis was also performed to provide further insight into the utility of these proteins. Among the proteins, CLU with 94.5% of sensitivity and 89.0% of specificity has a higher AUC value of 0.945, further indicating its diagnostic potential in OSCC. These findings were in line with a previous study that suggests a potential diagnostics role of CLU in liver cancer (Wang et al., 2010). Univariate and multivariate logistic regression analyses further predicted that AAT, APOA1, CLU, and HP could distinguish OSCC from the control significantly. Based on the analyses, these proteins were independent factors that have an impact on OSCC.

Universiti Malaya

Intriguingly, APOA1 and HP were found to be associated with the development of OPMD and early OSCC. Previous studies had shown that the dysregulation of APOA1 and HP was linked to a higher risk of developing OSCC (Camisasca et al., 2017; Chen et al., 2014; Thiel et al., 2011).

Concurrently, IHC was also performed to validate the presence of AAT, AHSG, APOA1, CLU, and HP in this study. IHC uses antibodies to detect specific proteins in FFPE tissues. It is a classical method that involves labelling of cellular proteins and microscopic examination for semi-quantitative assessment (Solier & Langen, 2014). In contrast with the ELISA analysis, only CLU and HP had shown significant differences across the study groups based on the IHC analysis. On the other hand, AAT, AHSG, APOA1, CLU, and HP in the IHC analysis have similar expression levels with the proteomics and glycoproteomics analyses. Among the proteins, CLU with 70.0% of sensitivity and 95.9% of specificity had a higher AUC value of 0.833, which indicates that it could have a potential diagnostic performance in OSCC. These findings further suggested CLU has a key role in the development and progression of cancers (Kadam & Teni, 2016). However, these five proteins were not correlated with any socio-demographic and clinicopathological characteristics or even the survival status of OSCC.

Both ELISA and IHC analyses showed that OSCC patients in the older age group and/or with betel quid chewing habit has high expression levels of AAT, AHSG, and APOA1 as well as low expression levels of CLU and HP. These findings may suggest that people older than 45 years or with betel quid chewing habit have a higher risk of developing OSCC. These findings concur with a recent case-control study that reported the age group of more than 40 years old and/or with betel quid chewing habit are the main risk factors of OSCC for the Malaysian population, especially in the Indian community (Ghani, Razak, et al., 2019).

Universiti Malaya

Importantly, the findings of ELISA were comparable with those results from the proteomics and glycoproteomics analyses. However, the results from ELISA exhibited disparity when compared with IHC analysis. The expression level of CLU and HP was detected significantly in the serum and tissue samples of OSCC. Whereas the expression level of AAT, AHSG, and APOA1 was only detected significantly in the serum samples of OSCC. This discrepancy indicated that these two antibody-based methods are not directly compatible, and they have different assessment on the targeted proteins. Moreover, the protein expression in tissue might not necessarily change in the same direction or correspond with the circulating level of the protein. It is known that cancer generally displays tumour heterogeneity that consists of diverse subpopulations of cancer cells (Marusyk & Polyak, 2010). OSCC has also exhibited tumour heterogeneity with site-wise variation in tissue composition (Sarode, Sarode, Tupkari, & Patil, 2017). Thus, tumour heterogeneity could affect the expression and distribution of certain proteins. Furthermore, the discrepancy of the results in the proteomics, glycoproteomics, and ELISA analyses with IHC analysis could be attributed to the use of different sample sources and different efficiency of the detection methods.

Although AAT, AHSG, APOA1, CLU, and HP had been identified and validated in this study, these differentially expressed proteins and glycoproteins had been shown in other types of cancers or diseases as aforementioned. Nevertheless, the combination of these differentially expressed proteins and glycoproteins might uniquely reflect the alteration in protein and glycoprotein expression levels during the development and progression of OSCC. To date, the combination of these proteins and glycoproteins is yet to be seen in other types of cancers or diseases. Therefore, it will be remarkably important to investigate whether the combination of these proteins and glycoproteins can be detected in other types of cancers when considering a panel of biomarker for OSCC.

Universiti Malaya

On the other hand, there were some limitations in this study. The main limitation of this study is related to the sample size. Overall, a total of 60 serum samples from control, OPMD, early, and advanced OSCC groups were included in the proteomics and glycoproteomics analyses. Whereas a total of 120 serum samples and 70 FFPE tissue samples were employed for validation purposes. The sample size for OPMD and early OSCC was small due to the limited availability of the sample. Despite the random sampling, it found that there were no samples chosen with Indian ethnicity in the control group for the proteomics and glycoproteomics analyses. Thus, the comparison between ethnic groups and OSCC cannot be made, and this is considered as a limitation of this study. A follow-on study could be performed on the data, with ethnicity being one of the sample selection criteria.

Both proteomics and glycoproteomics used in the present study can detect differentially expressed proteins and glycoproteins, respectively. However, these approaches were difficult to predict the structural changes of the proteins and glycoproteins during the development and progression of cancer. Additionally, the use of bioinformatics tools such as DAVID and IPA can explore the functional roles and biological pathways of the proteins and glycoproteins. These bioinformatics tools can also be used to investigate the interaction networks that are associated with the development and progression of cancer. Nevertheless, the IPA database contains many different protein-protein interactions derived from various experimental results. The interaction networks may not reflect the actual functional relationship between the proteins in the disease state.

Universiti Malaya

CHAPTER 6: CONCLUSION

The incidence and mortality rates of OSCC have remained high over the decades. This could reflect the delay in diagnosis since the signs and symptoms are only manifested at the advanced stage of OSCC and there are no reliable biomarkers for early screening. Therefore, this study was aimed to identify potential biomarkers that are involved in the development and progression of OSCC using integrated proteomics and glycoproteomics analyses. The identification of biomarkers would lead to the improvement in the diagnosis, prognosis, and treatment of OSCC.

The protein profiles of OPMD, OSCC, and control were established using 2-DE followed by silver staining. The proteomics analysis revealed 9 proteins (AAT, APOA1, C3, IGHA2, IGHG2, IGKC, SAMP, TF, and VDBP) that were significantly up-regulated and 7 proteins (ALB, AMBP, CLU, HP, LRG1, PRDX2, and RBP4) that were significantly down-regulated in OPMD and OSCC when compared with control. Concurrently, the *N*- and *O*-glycoprotein profiles of OPMD, OSCC, and control were also successfully generated using 2-DE coupled with Con A and Jacalin lectins, respectively. However, only *N*-glycosylated AAT and APOA1 were significantly detected in the glycoproteomics analysis. Of note, these differentially expressed proteins and glycoproteins may have a functional role in the development and progression of OSCC that can be potentially used for clinical applications.

In this study, it was also demonstrated that these differentially expressed proteins and glycoproteins may associate with the regulation of biological processes and interaction networks involved in OSCC. Based on the bioinformatics analysis, platelet degranulation, activation of classical complement pathway, LXR/RXR activation, and acute phase response signalling pathways are related to the development and progression of OSCC.

Universiti Malaya

These cellular processes and pathways indicate the involvement of chronic inflammation and the innate immunity system in OSCC.

Further validation using ELISA and IHC has revealed the potential use of AAT, AHSG, APOA1, CLU, and HP as the biomarkers of OSCC. In the ELISA analysis, AAT, APOA1, CLU, and HP were predicted as independent factors of OSCC. Whereas only CLU and HP showed a statistical significance in the IHC analysis.

Taken together, the findings of the present study have expanded the current understanding of the development and progression of OSCC. Although AAT, AHSG, APOA1, CLU, and HP are acute-phase proteins, which limit their specificity in OSCC, these proteins could be potentially used as complementary biomarkers to improve the detection of OSCC. Hence, further exploration of the utility of these biomarkers in OSCC is required. The changes in the biological mechanism of these proteins are also an important aspect for additional inspection. Using more robust and higher throughput approaches such as multiple reaction monitoring (MRM), selected reaction monitoring (SRM), and reverse-phase protein array (RPPA) in a larger cohort could be useful to determine the roles of these biomarkers in OSCC. Ultimately, there is a need for more prospective studies to translate the findings into cost-effective clinical practice.

Universiti Malaya

REFERENCES

A-Gonzalez, N., & Castrillo, A. (2011). Liver X receptors as regulators of macrophage inflammatory and metabolic pathways. *Biochim Biophys Acta, 1812*(8), 982-994. doi:10.1016/j.bbadis.2010.12.015

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. doi:10.1002/pmic.200900537

Abe, M., Yamashita, S., Mori, Y., Abe, T., Saijo, H., Hoshi, K., . . . Takato, T. (2016). High-risk oral leukoplakia is associated with aberrant promoter methylation of multiple genes. *BMC Cancer, 16*, 350. doi:10.1186/s12885-016-2371-5

Afshar-Kharghan, V. (2017). The role of the complement system in cancer. *Journal of Clinical Investigation, 127*(3), 780-789. doi:10.1172/Jci90962

Ahn, Y. H., Kim, J. Y., & Yoo, J. S. (2015). Quantitative mass spectrometric analysis of glycoproteins combined with enrichment methods. *Mass Spectrom Rev, 34*(2), 148-165. doi:10.1002/mas.21428

Aivaliotis, I. L., Pateras, I. S., Papaioannou, M., Glytsou, C., Kontzoglou, K., Johnson, E. O., & Zoumpourlis, V. (2012). How do cytokines trigger genomic instability? *J Biomed Biotechnol, 2012*, 536761. doi:10.1155/2012/536761

Akhter, M., Hossain, S., Rahman, Q. B., & Molla, M. R. (2011). A study on histological grading of oral squamous cell carcinoma and its co-relationship with regional metastasis. *J Oral Maxillofac Pathol, 15*(2), 168-176. doi:10.4103/0973-029X.84485

Andersen, J. D., Boylan, K. L., Jemmerson, R., Geller, M. A., Misemer, B., Harrington, K. M., . . . Skubitz, A. P. (2010). Leucine-rich alpha-2-glycoprotein-1 is upregulated in sera and tumors of ovarian cancer patients. *J Ovarian Res, 3*, 21. doi:10.1186/1757-2215-3-21

Anderson, N. L., & Anderson, N. G. (2002). The human plasma proteome: history, character, and diagnostic prospects. *Mol Cell Proteomics, 1*(11), 845-867.

Anneroth, G., Batsakis, J., & Luna, M. (1987). Review of the literature and a recommended system of malignancy grading in oral squamous cell carcinomas. *Scand J Dent Res, 95*(3), 229-249.

Apweiler, R., Hermjakob, H., & Sharon, N. (1999). On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochim Biophys Acta, 1473*(1), 4-8.

Arellano-Garcia, M. E., Li, R., Liu, X. J., Xie, Y. M., Yan, X. F., Loo, J. A., & Hu, S. (2010). Identification of Tetranectin as a Potential Biomarker for Metastatic Oral Cancer. *International Journal of Molecular Sciences, 11*(9), 3106-3121. doi:10.3390/ijms11093106

Universiti Malaya

Arya, V., Singh, S., & Daniel, M. J. (2016). Clinicopathological correlation of Bcl-2 oncoprotein expression in oral precancer and cancer. *J Oral Biol Craniofac Res, 6*(1), 18-23. doi:10.1016/j.jobcr.2015.12.011

Aslebagh, R., Channaveerappa, D., Arcaro, K. F., & Darie, C. C. (2018). Comparative two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) of human milk to identify dysregulated proteins in breast cancer. *Electrophoresis*. doi:10.1002/elps.201800025

Asokan, G. S., Jeelani, S., & Gnanasundaram, N. (2014). Promoter hypermethylation profile of tumour suppressor genes in oral leukoplakia and oral squamous cell carcinoma. *J Clin Diagn Res, 8*(10), ZC09-12. doi:10.7860/JCDR/2014/9251.4949

Avci, A., Tuzuner-Oncul, A. M., Gokcan, M. K., Namuslu, M., Ozturk, A., & Durak, I. (2009). Nitric oxide metabolism in cancerous and non-cancerous oral gingivomucosal tissues: possible implications of nitric oxide in cancer process. *J Oral Pathol Med, 38*(3), 304-306. doi:10.1111/j.1600-0714.2008.00719.x

Azizah, A., Hashimah, B., Nirmal, K., Siti Zubaidah, A., Puteri, N., Nabihah, A., . . . Azlina, A. (2019). Malaysia National Cancer Registry Report 2012-2016. *Putrajaya: National Cancer Registry, Malaysia Ministry of Health*.

Belczacka, I., Pejchinovski, M., Krochmal, M., Magalhaes, P., Frantzi, M., Mullen, W., . . . Jankowski, V. (2019). Urinary Glycopeptide Analysis for the Investigation of Novel Biomarkers. *Proteomics Clin Appl, 13*(3), e1800111. doi:10.1002/prca.201800111

Bellei, E., Bergamini, S., Monari, E., Fantoni, L. I., Cuoghi, A., Ozben, T., & Tomasi, A. (2011). High-abundance proteins depletion for serum proteomic analysis: concomitant removal of non-targeted proteins. *Amino Acids, 40*(1), 145-156. doi:10.1007/s00726-010-0628-x

Bertuzzi, M., Marelli, C., Bagnati, R., Colombi, A., Fanelli, R., Saieva, C., . . . Palli, D. (2015). Plasma clusterin as a candidate pre-diagnosis marker of colorectal cancer risk in the Florence cohort of the European Prospective Investigation into Cancer and Nutrition: a pilot study. *BMC Cancer, 15*, 56. doi:10.1186/s12885-015-1058-7

Betesh, L., Comunale, M. A., Wang, M., Liang, H., Hafner, J., Karabudak, A., . . . Mehta, A. (2017). Identification of fucosylated Fetuin-A as a potential biomarker for cholangiocarcinoma. *Proteomics Clin Appl, 11*(9-10). doi:10.1002/prca.201600141

Bhatia, V., Goel, M. M., Makker, A., Tewari, S., Yadu, A., Shilpi, P., . . . Goel, S. K. (2014). Promoter region hypermethylation and mRNA expression of MGMT and p16 genes in tissue and blood samples of human premalignant oral lesions and oral squamous cell carcinoma. *Biomed Res Int, 2014*, 248419. doi:10.1155/2014/248419

Bijian, K., Mlynarek, A. M., Balys, R. L., Jie, S., Xu, Y., Hier, M. P., . . . Alaoui-Jamali, M. A. (2009). Serum proteomic approach for the identification of serum

Universiti Malaya

biomarkers contributed by oral squamous cell carcinoma and host tissue microenvironment. *Journal of Proteome Research, 8*(5), 2173-2185. doi:10.1021/pr800979e

Biomarkers Definitions Working, G. (2001). Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clin Pharmacol Ther, 69*(3), 89-95. doi:10.1067/mcp.2001.113989

Bjellqvist, B., Ek, K., Righetti, P. G., Gianazza, E., Gorg, A., Westermeier, R., & Postel, W. (1982). Isoelectric focusing in immobilized pH gradients: principle, methodology and some applications. *J Biochem Biophys Methods, 6*(4), 317-339.

Boichenko, A. P., Govorukhina, N., Klip, H. G., van der Zee, A. G., Guzel, C., Luider, T. M., & Bischoff, R. (2014). A panel of regulated proteins in serum from patients with cervical intraepithelial neoplasia and cervical cancer. *Journal of Proteome Research, 13*(11), 4995-5007. doi:10.1021/pr500601w

Borgquist, S., Butt, T., Almgren, P., Shiffman, D., Stocks, T., Orho-Melander, M., . . . Melander, O. (2016). Apolipoproteins, lipids and risk of cancer. *Int J Cancer, 138*(11), 2648-2656. doi:10.1002/ijc.30013

Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., & Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin, 68*(6), 394-424. doi:10.3322/caac.21492

Brockhausen, I., & Stanley, P. (2015). O-GalNAc Glycans. In rd, A. Varki, R. D. Cummings, J. D. Esko, P. Stanley, G. W. Hart, M. Aebi, A. G. Darvill, T. Kinoshita, N. H. Packer, J. H. Prestegard, R. L. Schnaar, & P. H. Seeberger (Eds.), *Essentials of Glycobiology* (pp. 113-123). Cold Spring Harbor (NY).

Brown, J. M., Recht, L., & Strober, S. (2017). The Promise of Targeting Macrophages in Cancer Therapy. *Clin Cancer Res, 23*(13), 3241-3250. doi:10.1158/1078-0432.CCR-16-3122

Bunkenborg, J., Pilch, B. J., Podtelejnikov, A. V., & Wisniewski, J. R. (2004). Screening for N-glycosylated proteins by liquid chromatography mass spectrometry. *Proteomics, 4*(2), 454-465. doi:10.1002/pmic.200300556

Camisasca, D. R., da Ros Goncalves, L., Soares, M. R., Sandim, V., Nogueira, F. C., Garcia, C. H., . . . Lourenco, S. Q. (2017). A proteomic approach to compare saliva from individuals with and without oral leukoplakia. *J Proteomics, 151*, 43-52. doi:10.1016/j.jprot.2016.07.029

Campos, D., Freitas, D., Gomes, J., Magalhaes, A., Steentoft, C., Gomes, C., . . . Reis, C. A. (2015). Probing the O-glycoproteome of gastric cancer cell lines for biomarker discovery. *Mol Cell Proteomics, 14*(6), 1616-1629. doi:10.1074/mcp.M114.046862

Chaerkady, R., Thuluvath, P. J., Kim, M. S., Nalli, A., Vivekanandan, P., Simmers, J., . . . Pandey, A. (2008). O Labeling for a Quantitative Proteomic Analysis of

Universiti Malaya

Glycoproteins in Hepatocellular Carcinoma. *Clin Proteomics, 4*(3-4), 137-155. doi:10.1007/s12014-008-9013-0

Chang, C. C., Lee, W. T., Hsiao, J. R., Ou, C. Y., Huang, C. C., Tsai, S. T., . . . Chang, J. S. (2019). Oral hygiene and the overall survival of head and neck cancer patients. *Cancer Med, 8*(4), 1854-1864. doi:10.1002/cam4.2059

Chang, S.-C., Lin, W.-L., Chang, Y.-F., Lee, C.-T., Wu, J.-S., Hsu, P.-H., & Chang, C.-F. (2019). Glycoproteomic identification of novel plasma biomarkers for oral cancer. *Journal of Food and Drug Analysis*.

Chantaraamporn, J., Champattanachai, V., Khongmanee, A., Verathamjamras, C., Prasongsook, N., Mingkwan, K., . . . Svasti, J. (2020). Glycoproteomic Analysis Reveals Aberrant Expression of Complement C9 and Fibronectin in the Plasma of Patients with Colorectal Cancer. *Proteomes, 8*(3). doi:10.3390/proteomes8030026

Chanthammachat, P., Promwikorn, W., Pruegsanusak, K., Roytrakul, S., Srisomsap, C., Chokchaichamnankit, D., . . . Thongsuksai, P. (2013). Comparative proteomic analysis of oral squamous cell carcinoma and adjacent non-tumour tissue from Thailand. *Arch Oral Biol, 58*(11), 1677-1685. doi:10.1016/j.archoralbio.2013.08.002

Chen, C. B., Su, Y. C., Huang, T. T., Ho, H. C., Chang, Y. T., Tung, Y. T., & Lee, W. C. (2008). Differentially expressed serum haptoglobin alpha chain isoforms with potential application for diagnosis of head and neck cancer. *Clin Chim Acta, 398*(1-2), 48-52. doi:10.1016/j.cca.2008.08.001

Chen, J., Wu, W., Zhen, C., Zhou, H., Yang, R., Chen, L., & Hu, L. (2013). Expression and clinical significance of complement C3, complement C4b1 and apolipoprotein E in pancreatic cancer. *Oncol Lett, 6*(1), 43-48. doi:10.3892/ol.2013.1326

Chen, J. T., Chen, C. H., Ku, K. L., Hsiao, M., Chiang, C. P., Hsu, T. L., . . . Wong, C. H. (2015). Glycoprotein B7-H3 overexpression and aberrant glycosylation in oral cancer and immune response. *Proc Natl Acad Sci U S A, 112*(42), 13057-13062. doi:10.1073/pnas.1516991112

Chen, L., Deng, H., Cui, H., Fang, J., Zuo, Z., Deng, J., . . . Zhao, L. (2018). Inflammatory responses and inflammation-associated diseases in organs. *Oncotarget, 9*(6), 7204-7218. doi:10.18632/oncotarget.23208

Chen, Y., Azman, S. N., Kerishnan, J. P., Zain, R. B., Chen, Y. N., Wong, Y. L., & Gopinath, S. C. (2014). Identification of host-immune response protein candidates in the sera of human oral squamous cell carcinoma patients. *PLoS One, 9*(10), e109012. doi:10.1371/journal.pone.0109012

Chen, Y. T., Chen, H. W., Wu, C. F., Chu, L. J., Chiang, W. F., Wu, C. C., . . . Ou Yang, W. T. (2017). Development of a Multiplexed Liquid Chromatography Multiple-Reaction-Monitoring Mass Spectrometry (LC-MRM/MS) Method for Evaluation of Salivary Proteins as Oral Cancer Biomarkers. *Mol Cell Proteomics, 16*(5), 799-811. doi:10.1074/mcp.M116.064758

Chen, Y. T., Chong, Y. M., Cheng, C. W., Ho, C. L., Tsai, H. W., Kasten, F. H., . . . Chang, C. F. (2013). Identification of novel tumor markers for oral squamous cell carcinoma using glycoproteomic analysis. *Clin Chim Acta, 420*, 45-53. doi:10.1016/j.cca.2012.10.019

Cheong, S. C., Vatanasapt, P., Yi-Hsin, Y., Zain, R. B., Kerr, A. R., & Johnson, N. W. (2017). Oral cancer in South East Asia: current status and future directions. *Translational Research in Oral Oncology, 2*, 2057178X17702921.

Choi, S., & Myers, J. N. (2008). Molecular pathogenesis of oral squamous cell carcinoma: implications for therapy. *J Dent Res, 87*(1), 14-32. doi:10.1177/154405910808700104

Chung, L., Moore, K., Phillips, L., Boyle, F. M., Marsh, D. J., & Baxter, R. C. (2014). Novel serum protein biomarker panel revealed by mass spectrometry and its prognostic value in breast cancer. *Breast Cancer Res, 16*(3), R63. doi:10.1186/bcr3676

Ciereszko, A., Dietrich, M. A., Slowinska, M., Nynca, J., Ciborowski, M., Kisluk, J., . . . Niklinski, J. (2019). Identification of protein changes in the blood plasma of lung cancer patients subjected to chemotherapy using a 2D-DIGE approach. *PLoS One, 14*(10), e0223840. doi:10.1371/journal.pone.0223840

Clerc, F., Reiding, K. R., Jansen, B. C., Kammeijer, G. S., Bondt, A., & Wuhrer, M. (2016). Human plasma protein N-glycosylation. *Glycoconj J, 33*(3), 309-343. doi:10.1007/s10719-015-9626-2

Colley, K. J., Varki, A., & Kinoshita, T. (2015). Cellular Organization of Glycosylation. In rd, A. Varki, R. D. Cummings, J. D. Esko, P. Stanley, G. W. Hart, M. Aebi, A. G. Darvill, T. Kinoshita, N. H. Packer, J. H. Prestegard, R. L. Schnaar, & P. H. Seeberger (Eds.), *Essentials of Glycobiology* (pp. 41-49). Cold Spring Harbor (NY).

Comunale, M. A., Lowman, M., Long, R. E., Krakover, J., Philip, R., Seeholzer, S., . . . Mehta, A. S. (2006). Proteomic analysis of serum associated fucosylated glycoproteins in the development of primary hepatocellular carcinoma. *Journal of Proteome Research, 5*(2), 308-315. doi:10.1021/pr050328x

Comunale, M. A., Wang, M., Rodemich-Betesh, L., Hafner, J., Lamontagne, A., Klein, A., . . . Mehta, A. (2011). Novel changes in glycosylation of serum Apo-J in patients with hepatocellular carcinoma. *Cancer Epidemiol Biomarkers Prev, 20*(6), 1222-1229. doi:10.1158/1055-9965.EPI-10-1047

Costa, S. F. D. S., Brennan, P. A., Gomez, R. S., Fregnani, E. R., Santos-Silva, A. R., Martins, M. D., . . . Fonseca, F. P. (2018). Molecular basis of oral squamous cell carcinoma in young patients: Is it any different from older patients? *J Oral Pathol Med, 47*(6), 541-546. doi:10.1111/jop.12642

Coussens, L. M., & Werb, Z. (2002). Inflammation and cancer. *Nature, 420*(6917), 860-867. doi:10.1038/nature01322

Universiti Malaya

Coutinho-Camillo, C. M., Lourenço, S. V., Puga, R. D., Damascena, A. S., Teshima, T. H. N., Kowalski, L. P., & Soares, F. A. (2017). Profile of apoptotic proteins in oral squamous cell carcinoma: A cluster analysis of 171 cases. *Applied cancer research, 37*(1), 2.

Crutchfield, C. A., Thomas, S. N., Sokoll, L. J., & Chan, D. W. (2016). Advances in mass spectrometry-based clinical biomarker discovery. *Clin Proteomics, 13*, 1. doi:10.1186/s12014-015-9102-9

Csosz, E., Markus, B., Darula, Z., Medzihradszky, K. F., Nemes, J., Szabo, E., . . . Marton, I. (2018). Salivary proteome profiling of oral squamous cell carcinoma in a Hungarian population. *FEBS Open Bio, 8*(4), 556-569. doi:10.1002/2211-5463.12391

Cutilli, T., Leocata, P., Dolo, V., & Altobelli, E. (2016). p53 as a prognostic marker associated with the risk of mortality for oral squamous cell carcinoma. *Oncol Lett, 12*(2), 1046-1050. doi:10.3892/ol.2016.4742

Da Costa, G. G., Gomig, T. H., Kaviski, R., Santos Sousa, K., Kukolj, C., De Lima, R. S., . . . Ribeiro, E. M. (2015). Comparative Proteomics of Tumor and Paired Normal Breast Tissue Highlights Potential Biomarkers in Breast Cancer. *Cancer Genomics Proteomics, 12*(5), 251-261.

Dai, Z., Zhou, J., Qiu, S. J., Liu, Y. K., & Fan, J. (2009). Lectin-based glycoproteomics to explore and analyze hepatocellular carcinoma-related glycoprotein markers. *Electrophoresis, 30*(17), 2957-2966. doi:10.1002/elps.200900064

Dan, X. L., Liu, W. L., & Ng, T. B. (2016). Development and Applications of Lectins as Biological Tools in Biomedical Research. *Medicinal Research Reviews, 36*(2), 221-247. doi:10.1002/med.21363

Dempsey, E., & Rudd, P. M. (2012). Acute phase glycoproteins: bystanders or participants in carcinogenesis? *Ann N Y Acad Sci, 1253*, 122-132. doi:10.1111/j.1749-6632.2011.06420.x

Dionne, K. R., Warnakulasuriya, S., Zain, R. B., & Cheong, S. C. (2015). Potentially malignant disorders of the oral cavity: current practice and future directions in the clinic and laboratory. *Int J Cancer, 136*(3), 503-515. doi:10.1002/ijc.28754

Drabik, A., Bodzon-Kulakowska, A., Suder, P., Silberring, J., Kulig, J., & Sierzega, M. (2017). Glycosylation Changes in Serum Proteins Identify Patients with Pancreatic Cancer. *Journal of Proteome Research, 16*(4), 1436-1444. doi:10.1021/acs.jproteome.6b00775

Drake, P. M., Cho, W., Li, B., Prakobphol, A., Johansen, E., Anderson, N. L., . . . Fisher, S. J. (2010). Sweetening the pot: adding glycosylation to the biomarker discovery equation. *Clin Chem, 56*(2), 223-236. doi:10.1373/clinchem.2009.136333

Drake, R. R., Schwegler, E. E., Malik, G., Diaz, J., Block, T., Mehta, A., & Semmes, O. J. (2006). Lectin capture strategies combined with mass spectrometry for the discovery of serum glycoprotein biomarkers. *Mol Cell Proteomics, 5*(10), 1957-1967. doi:10.1074/mcp.M600176-MCP200

Universiti Malaya

Dube, D. H., & Bertozzi, C. R. (2005). Glycans in cancer and inflammation--potential for therapeutics and diagnostics. *Nat Rev Drug Discov, 4*(6), 477-488. doi:10.1038/nrd1751

Dudas, S. P., Yunker, C. K., Sternberg, L. R., Byrd, J. C., & Bresalier, R. S. (2002). Expression of human intestinal mucin is modulated by the beta-galactoside binding protein galectin-3 in colon cancer. *Gastroenterology, 123*(3), 817-826. doi:10.1053/gast.2002.35395

Dufour, J., Pommier, A., Alves, G., De Boussac, H., Lours-Calet, C., Volle, D. H., . . . Baron, S. (2013). Lack of liver X receptors leads to cell proliferation in a model of mouse dorsal prostate epithelial cell. *PLoS One, 8*(3), e58876. doi:10.1371/journal.pone.0058876

Echan, L. A., Tang, H. Y., Ali-Khan, N., Lee, K., & Speicher, D. W. (2005). Depletion of multiple high-abundance proteins improves protein profiling capacities of human serum and plasma. *Proteomics, 5*(13), 3292-3303. doi:10.1002/pmic.200401228

Edge, S. B., Byrd, D. R., Carducci, M. A., Compton, C. C., Fritz, A., & Greene, F. (2010). *AJCC cancer staging manual* (Vol. 649): Springer New York.

Egan, K., Crowley, D., Smyth, P., O'Toole, S., Spillane, C., Martin, C., . . . O'Leary, J. (2011). Platelet adhesion and degranulation induce pro-survival and pro-angiogenic signalling in ovarian cancer cells. *PLoS One, 6*(10), e26125. doi:10.1371/journal.pone.0026125

Elimairi, I., Sami, A., & Yousef, B. (2017). Oral Cancer and Potentially Malignant Disorders. In *Histopathology-An Update*: IntechOpen.

Ercetin, E., Richtmann, S., Delgado, B. M., Gomez-Mariano, G., Wrenger, S., Korenbaum, E., . . . Schneider, M. A. (2019). Clinical Significance of SERPINA1 Gene and Its Encoded Alpha1-antitrypsin Protein in NSCLC. *Cancers (Basel), 11*(9). doi:10.3390/cancers11091306

Eversole, L. R., & Sapp, J. P. (1993). c-myc oncoprotein expression in oral precancerous and early cancerous lesions. *Eur J Cancer B Oral Oncol, 29B*(2), 131-135.

Fabris, L., Ceder, Y., Chinnaiyan, A. M., Jenster, G. W., Sorensen, K. D., Tomlins, S., . . . Calin, G. A. (2016). The Potential of MicroRNAs as Prostate Cancer Biomarkers. *European Urology, 70*(2), 312-322. doi:10.1016/j.eururo.2015.12.054

Fan, Y., Murphy, T. B., Byrne, J. C., Brennan, L., Fitzpatrick, J. M., & Watson, R. W. (2011). Applying random forests to identify biomarker panels in serum 2D-DIGE data for the detection and staging of prostate cancer. *Journal of Proteome Research, 10*(3), 1361-1373. doi:10.1021/pr1011069

Farrokhi Yekta, R., Arefi Oskouie, A., Rezaei Tavirani, M., Mohajeri-Tehrani, M. R., & Soroush, A. R. (2018). Decreased apolipoprotein A4 and increased complement component 3 as potential markers for papillary thyroid carcinoma: A proteomic study. *Int J Biol Markers, 33*(4), 455-462. doi:10.1177/1724600818787752

Universiti Malaya

Fei, W., Chen, L., Chen, J., Shi, Q., Zhang, L., Liu, S., . . . Hu, X. (2017). RBP4 and THBS2 are serum biomarkers for diagnosis of colorectal cancer. *Oncotarget, 8*(54), 92254-92264. doi:10.18632/oncotarget.21173

Feist, P., & Hummon, A. B. (2015). Proteomic challenges: sample preparation techniques for microgram-quantity protein analysis from biological samples. *Int J Mol Sci, 16*(2), 3537-3563. doi:10.3390/ijms16023537

Ferens-Sieczkowska, M., Kratz, E. M., Kossowska, B., Passowicz-Muszynska, E., & Jankowska, R. (2013). Comparison of haptoglobin and alpha(1)-acid glycoprotein glycosylation in the sera of small cell and non-small cell lung cancer patients. *Postepy Hig Med Dosw (Online), 67*, 828-836. doi:10.5604/17322693.1061788

Ferlay, J., Colombet, M., Soerjomataram, I., Mathers, C., Parkin, D. M., Pineros, M., . . . Bray, F. (2019). Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. *Int J Cancer, 144*(8), 1941-1953. doi:10.1002/ijc.31937

Ferlay, J., Ervik, M., Lam, F., Colombet, M., Mery, L., Piñeros, M., . . . Bray, F. (2018). Global cancer observatory: cancer today. *Lyon, France: International Agency for Research on Cancer*.

Fitzpatrick, S. G., Hirsch, S. A., & Gordon, S. C. (2014). The malignant transformation of oral lichen planus and oral lichenoid lesions: a systematic review. *J Am Dent Assoc, 145*(1), 45-56. doi:10.14219/jada.2013.10

Flanagan, L., Whyte, L., Chatterjee, N., & Tenniswood, M. (2010). Effects of clusterin over-expression on metastatic progression and therapy in breast cancer. *BMC Cancer, 10*, 107. doi:10.1186/1471-2407-10-107

Fry, S. A., Sinclair, J., Timms, J. F., Leathem, A. J., & Dwek, M. V. (2013). A targeted glycoproteomic approach identifies cadherin-5 as a novel biomarker of metastatic breast cancer. *Cancer Lett, 328*(2), 335-344. doi:10.1016/j.canlet.2012.10.011

Furukawa, K., Kawamoto, K., Eguchi, H., Tanemura, M., Tanida, T., Tomimaru, Y., . . . Nagano, H. (2015). Clinicopathological Significance of Leucine-Rich alpha2-Glycoprotein-1 in Sera of Patients With Pancreatic Cancer. *Pancreas, 44*(1), 93-98. doi:10.1097/MPA.0000000000000205

Gabay, C., & Kushner, I. (1999). Acute-phase proteins and other systemic responses to inflammation. *N Engl J Med, 340*(6), 448-454. doi:10.1056/NEJM199902113400607

Gallenkamp, J., Spanier, G., Worle, E., Englbrecht, M., Kirschfink, M., Greslechner, R., . . . Pauly, D. (2018). A novel multiplex detection array revealed systemic complement activation in oral squamous cell carcinoma. *Oncotarget, 9*(3), 3001-3013. doi:10.18632/oncotarget.22963

Gan, Y., Chen, D., & Li, X. (2014). Proteomic analysis reveals novel proteins associated with progression and differentiation of colorectal carcinoma. *J Cancer Res Ther, 10*(1), 89-96. doi:10.4103/0973-1482.131396

Universiti Malaya

Garibay-Cerdenares, O. L., Hernandez-Ramirez, V. I., Osorio-Trujillo, J. C., Gallardo-Rincon, D., & Talamas-Rohana, P. (2015). Haptoglobin and CCR2 receptor expression in ovarian cancer cells that were exposed to ascitic fluid: exploring a new role of haptoglobin in the tumoral microenvironment. *Cell Adh Migr, 9*(5), 394-405. doi:10.1080/19336918.2015.1035504

Genden, E. M., Ferlito, A., Silver, C. E., Takes, R. P., Suarez, C., Owen, R. P., . . . Rinaldo, A. (2010). Contemporary management of cancer of the oral cavity. *Eur Arch Otorhinolaryngol, 267*(7), 1001-1017. doi:10.1007/s00405-010-1206-2

Georgila, K., Vyrla, D., & Drakos, E. (2019). Apolipoprotein A-I (ApoA-I), Immunity, Inflammation and Cancer. *Cancers (Basel), 11*(8). doi:10.3390/cancers11081097

Ghanghoria, S., Ghanghoria, A., & Shukla, A. (2015). p53 Expression in Oral cancer: A study of 50 cases. *Journal of Pathology of Nepal, 5*(9), 747-751.

Ghani, W. M., Razak, I. A., Yang, Y. H., Talib, N. A., Ikeda, N., Axell, T., . . . Zain, R. B. (2011). Factors affecting commencement and cessation of betel quid chewing behaviour in Malaysian adults. *BMC Public Health, 11*, 82. doi:10.1186/1471-2458-11-82

Ghani, W. M. N., Ramanathan, A., Prime, S. S., Yang, Y. H., Razak, I. A., Abdul Rahman, Z. A., . . . Zain, R. B. (2019). Survival of Oral Cancer Patients in Different Ethnicities. *Cancer Invest*, 1-13. doi:10.1080/07357907.2019.1635614

Ghani, W. M. N., Razak, I. A., Doss, J. G., Yang, Y. H., Rahman, Z. A. A., Ismail, S. M., . . . Zain, R. B. (2019). Multi-ethnic variations in the practice of oral cancer risk habits in a developing country. *Oral Dis, 25*(2), 447-455. doi:10.1111/odi.12995

Ghantous, Y., Schussel, J. L., & Brait, M. (2018). Tobacco and alcohol-induced epigenetic changes in oral carcinoma. *Curr Opin Oncol, 30*(3), 152-158. doi:10.1097/CCO.0000000000000444

Gkouveris, I., Nikitakis, N., Karanikou, M., Rassidakis, G., & Sklavounou, A. (2014). Erk1/2 activation and modulation of STAT3 signaling in oral cancer. *Oncol Rep, 32*(5), 2175-2182. doi:10.3892/or.2014.3440

Gong, J., Jaiswal, R., Dalla, P., Luk, F., & Bebawy, M. (2015). Microparticles in cancer: A review of recent developments and the potential for clinical application. *Semin Cell Dev Biol, 40*, 35-40. doi:10.1016/j.semcdb.2015.03.009

Gorg, A., Postel, W., & Gunther, S. (1988). The current state of two-dimensional electrophoresis with immobilized pH gradients. *Electrophoresis, 9*(9), 531-546. doi:10.1002/elps.1150090913

Gourin, C. G., Zhi, W., & Adam, B. L. (2009). Proteomic identification of serum biomarkers for head and neck cancer surveillance. *Laryngoscope, 119*(7), 1291-1302. doi:10.1002/lary.20279

Granger, J., Siddiqui, J., Copeland, S., & Remick, D. (2005). Albumin depletion of human plasma also removes low abundance proteins including the cytokines. *Proteomics, 5*(18), 4713-4718. doi:10.1002/pmic.200401331

Universiti Malaya

Graves, P. R., & Haystead, T. A. (2002). Molecular biologist's guide to proteomics. *Microbiol Mol Biol Rev, 66*(1), 39-63; table of contents. doi:10.1128/mmbr.66.1.39-63.2002

Greulich, T., Nell, C., Herr, C., Vogelmeier, C., Kotke, V., Wiedmann, S., . . . Koczulla, A. R. (2016). Results from a large targeted screening program for alpha-1-antitrypsin deficiency: 2003 - 2015. *Orphanet J Rare Dis, 11*(1), 75. doi:10.1186/s13023-016-0453-8

Guidry, J. T., Birdwell, C. E., & Scott, R. S. (2018). Epstein-Barr virus in the pathogenesis of oral cancers. *Oral Dis, 24*(4), 497-508. doi:10.1111/odi.12656

Gundry, R. L., Fu, Q., Jelinek, C. A., Van Eyk, J. E., & Cotter, R. J. (2007). Investigation of an albumin-enriched fraction of human serum and its albuminome. *Proteomics Clin Appl, 1*(1), 73-88. doi:10.1002/prca.200600276

Guo, X., Hao, Y., Kamilijiang, M., Hasimu, A., Yuan, J., Wu, G., . . . Abudula, A. (2015). Potential predictive plasma biomarkers for cervical cancer by 2D-DIGE proteomics and Ingenuity Pathway Analysis. *Tumour Biol, 36*(3), 1711-1720. doi:10.1007/s13277-014-2772-5

Gupta, A. (2020). Emerging applications of lectins in cancer detection and biomedicine. *Materials Today: Proceedings, 31*, 651-661.

Gupta, B., & Johnson, N. W. (2014). Systematic review and meta-analysis of association of smokeless tobacco and of betel quid without tobacco with incidence of oral cancer in South Asia and the Pacific. *PLoS One, 9*(11), e113385. doi:10.1371/journal.pone.0113385

Gupta, N., Gupta, R., Acharya, A. K., Patthi, B., Goud, V., Reddy, S., . . . Singla, A. (2016). Changing Trends in oral cancer - a global scenario. *Nepal J Epidemiol, 6*(4), 613-619. doi:10.3126/nje.v6i4.17255

Gupta, R., & Brunak, S. (2002). Prediction of glycosylation across the human proteome and the correlation to protein function. *Pac Symp Biocomput*, 310-322.

Guttman, O., Baranovski, B. M., Schuster, R., Kaner, Z., Freixo-Lima, G. S., Bahar, N., . . . Lewis, E. C. (2015). Acute-phase protein alpha1-anti-trypsin: diverting injurious innate and adaptive immune responses from non-authentic threats. *Clin Exp Immunol, 179*(2), 161-172. doi:10.1111/cei.12476

Guu, S. Y., Lin, T. H., Chang, S. C., Wang, R. J., Hung, L. Y., Fang, P. J., . . . Chang, C. F. (2017). Serum N-glycome characterization and anti-carbohydrate antibody profiling in oral squamous cell carcinoma patients. *PLoS One, 12*(6), e0178927. doi:10.1371/journal.pone.0178927

Haj-Ahmad, T. A., Abdalla, M. A., & Haj-Ahmad, Y. (2014). Potential Urinary Protein Biomarker Candidates for the Accurate Detection of Prostate Cancer among Benign Prostatic Hyperplasia Patients. *J Cancer, 5*(2), 103-114. doi:10.7150/jca.6890

Universiti Malaya

Hakomori, S. (2002). Glycosylation defining cancer malignancy: new wine in an old bottle. *Proc Natl Acad Sci U S A, 99*(16), 10231-10233. doi:10.1073/pnas.172380699

Hanahan, D., & Weinberg, R. A. (2000). The hallmarks of cancer. *Cell, 100*(1), 57-70.

Hanahan, D., & Weinberg, R. A. (2011). Hallmarks of cancer: the next generation. *Cell, 144*(5), 646-674. doi:10.1016/j.cell.2011.02.013

Hashim, O. H., Jayapalan, J. J., & Lee, C. S. (2017). Lectins: an effective tool for screening of potential cancer biomarkers. *PeerJ, 5*, e3784. doi:10.7717/peerj.3784

He, K. F., Zhang, L., Huang, C. F., Ma, S. R., Wang, Y. F., Wang, W. M., . . . Sun, Z. J. (2014). CD163+ tumor-associated macrophages correlated with poor prognosis and cancer stem cells in oral squamous cell carcinoma. *Biomed Res Int, 2014*, 838632. doi:10.1155/2014/838632

He, Y., Chen, Q., & Li, B. (2008). ATM in oral carcinogenesis: association with clinicopathological features. *J Cancer Res Clin Oncol, 134*(9), 1013-1020. doi:10.1007/s00432-008-0365-7

Helen-Ng, L. C., Razak, I. A., Ghani, W. M., Marhazlinda, J., Norain, A. T., Raja Jallaludin, R. L., . . . Zain, R. B. (2012). Dietary pattern and oral cancer risk--a factor analysis study. *Community Dent Oral Epidemiol, 40*(6), 560-566. doi:10.1111/j.1600-0528.2012.00704.x

His, M., Zelek, L., Deschasaux, M., Pouchieu, C., Kesse-Guyot, E., Hercberg, S., . . . Touvier, M. (2014). Prospective associations between serum biomarkers of lipid metabolism and overall, breast and prostate cancer risk. *Eur J Epidemiol, 29*(2), 119-132. doi:10.1007/s10654-014-9884-5

Hsieh, Y. T., Chen, Y. F., Lin, S. C., Chang, K. W., & Li, W. C. (2019). Targeting Cellular Metabolism Modulates Head and Neck Oncogenesis. *Int J Mol Sci, 20*(16). doi:10.3390/ijms20163960

Hu, S., Arellano, M., Boontheung, P., Wang, J., Zhou, H., Jiang, J., . . . Wong, D. T. (2008). Salivary proteomics for oral cancer biomarker discovery. *Clin Cancer Res, 14*(19), 6246-6252. doi:10.1158/1078-0432.CCR-07-5037

Huang da, W., Sherman, B. T., & Lempicki, R. A. (2009a). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc, 4*(1), 44-57. doi:10.1038/nprot.2008.211

Huang da, W., Sherman, B. T., & Lempicki, R. A. (2009b). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res, 37*(1), 1-13. doi:10.1093/nar/gkn923

Huang, S. F., Cheng, S. D., Chien, H. T., Liao, C. T., Chen, I. H., Wang, H. M., . . . Hsieh, L. L. (2012). Relationship between epidermal growth factor receptor gene copy number and protein expression in oral cavity squamous cell carcinoma. *Oral Oncol, 48*(1), 67-72. doi:10.1016/j.oraloncology.2011.06.511

Universiti Malaya

Huang, S. H., & O'Sullivan, B. (2013). Oral cancer: Current role of radiotherapy and chemotherapy. *Med Oral Patol Oral Cir Bucal, 18*(2), e233-240.

Huang, Y. F., Wu, Y. H., Cheng, W. F., Peng, S. L., Shen, W. L., & Chou, C. Y. (2018). Vitamin D-Binding Protein Enhances Epithelial Ovarian Cancer Progression by Regulating the Insulin-like Growth Factor-1/Akt Pathway and Vitamin D Receptor Transcription. *Clin Cancer Res, 24*(13), 3217-3228. doi:10.1158/1078-0432.CCR-17-2943

Hudler, P., Kocevar, N., & Komel, R. (2014). Proteomic approaches in biomarker discovery: new perspectives in cancer diagnostics. *ScientificWorldJournal, 2014*, 260348. doi:10.1155/2014/260348

Hussein, A. A., Helder, M. N., de Visscher, J. G., Leemans, C. R., Braakhuis, B. J., de Vet, H. C. W., & Forouzanfar, T. (2017). Global incidence of oral and oropharynx cancer in patients younger than 45 years versus older patients: A systematic review. *Eur J Cancer, 82*, 115-127. doi:10.1016/j.ejca.2017.05.026

IARC. (2004a). Tobacco smoke and involuntary smoking. *IARC Monogr Eval Carcinog Risks Hum, 83*, 1-1438.

IARC. (2004b). Betel-quid and areca-nut chewing and some areca-nut derived nitrosamines. *IARC Monogr Eval Carcinog Risks Hum, 85*, 1-334.

Idris, S. F., Ahmad, S. S., Scott, M. A., Vassiliou, G. S., & Hadfield, J. (2013). The role of high-throughput technologies in clinical cancer genomics. *Expert Rev Mol Diagn, 13*(2), 167-181. doi:10.1586/erm.13.1

Inchingolo, F., Santacroce, L., Ballini, A., Topi, S., Dipalma, G., Haxhirexha, K., . . . Charitos, I. A. (2020). Oral Cancer: A Historical Review. *Int J Environ Res Public Health, 17*(9). doi:10.3390/ijerph17093168

Islam, T., Rahman, R., Gov, E., Turanli, B., Gulfidan, G., Haque, A., . . . Haque Mollah, N. (2018). Drug Targeting and Biomarkers in Head and Neck Cancers: Insights from Systems Biology Analyses. *OMICS, 22*(6), 422-436. doi:10.1089/omi.2018.0048

Jain, M. (2016). Assesment of Correlation of Herpes Simplex Virus-1 with Oral Cancer and Precancer- A Comparative Study. *J Clin Diagn Res, 10*(8), ZC14-17. doi:10.7860/JCDR/2016/18593.8229

Jamal, B., Sengupta, P. K., Gao, Z. N., Nita-Lazar, M., Amin, B., Jalisi, S., . . . Kukuruzinska, M. A. (2012). Aberrant amplification of the crosstalk between canonical Wnt signaling and N-glycosylation gene DPAGT1 promotes oral cancer. *Oral Oncol, 48*(6), 523-529. doi:10.1016/j.oraloncology.2012.01.010

Jassim, M. M. A., Mahmood, M. M., & Musa, S. Q. (2020). Herpetic Oncogenic Virus Co-Infections Impinge Cell Cycle Regulatory Gene Expressions in Oral Cavity Malignancies. *Systematic Reviews in Pharmacy, 11*(3). doi:10.31838/srp.2020.3.106

Universiti Malaya

Jessie, K., Jayapalan, J. J., Ong, K. C., Abdul Rahim, Z. H., Zain, R. M., Wong, K. T., & Hashim, O. H. (2013). Aberrant proteins in the saliva of patients with oral squamous cell carcinoma. *Electrophoresis, 34*(17), 2495-2502. doi:10.1002/elps.201300107

Jiang, S. Y., Ma, Z., & Ramachandran, S. (2010). Evolutionary history and stress regulation of the lectin superfamily in higher plants. *BMC Evol Biol, 10*, 79. doi:10.1186/1471-2148-10-79

Jiang, X., Wu, J., Wang, J., & Huang, R. (2019). Tobacco and oral squamous cell carcinoma: A review of carcinogenic pathways. *Tob Induc Dis, 17*, 29. doi:10.18332/tid/105844

Jiao, C., Cui, L., Ma, A., Li, N., & Si, H. (2016). Elevated Serum Levels of Retinol-Binding Protein 4 Are Associated with Breast Cancer Risk: A Case-Control Study. *PLoS One, 11*(12), e0167498. doi:10.1371/journal.pone.0167498

Jin, J., Kim, J. M., Hur, Y. S., Cho, W. P., Lee, K. Y., Ahn, S. I., . . . Park, I. S. (2012). Clinical significance of clusterin expression in pancreatic adenocarcinoma. *World J Surg Oncol, 10*, 146. doi:10.1186/1477-7819-10-146

Jin, R., Chen, X., Han, D., Luo, X., & Li, H. (2017). Clusterin modulates transdifferentiation of non-small-cell lung cancer. *BMC Cancer, 17*(1), 661. doi:10.1186/s12885-017-3649-y

Jin, Y. T., & Lin, L. M. (1989). Lectin binding patterns in squamous epithelium in experimentally induced hamster buccal pouch carcinoma. *J Oral Pathol Med, 18*(8), 446-450. doi:10.1111/j.1600-0714.1989.tb01341.x

Johnson, D. E., O'Keefe, R. A., & Grandis, J. R. (2018). Targeting the IL-6/JAK/STAT3 signalling axis in cancer. *Nat Rev Clin Oncol, 15*(4), 234-248. doi:10.1038/nrclinonc.2018.8

Johnson, N. W., Jayasekara, P., & Amarasinghe, A. A. (2011). Squamous cell carcinoma and precursor lesions of the oral cavity: epidemiology and aetiology. *Periodontol 2000, 57*(1), 19-37. doi:10.1111/j.1600-0757.2011.00401.x

Jou, Y. J., Lin, C. D., Lai, C. H., Chen, C. H., Kao, J. Y., Chen, S. Y., . . . Lin, C. W. (2010). Proteomic identification of salivary transferrin as a biomarker for early detection of oral cancer. *Analytica Chimica Acta, 681*(1-2), 41-48. doi:10.1016/j.aca.2010.09.030

Kadam, R., & Teni, T. (2016). Clusterin in Cancer: Dual role as a Tumor Suppressor Gene and an Oncogene. *Biomedical Research Journal, 3*(2), 130.

Kailemia, M. J., Park, D., & Lebrilla, C. B. (2017). Glycans and glycoproteins as specific biomarkers for cancer. *Anal Bioanal Chem, 409*(2), 395-410. doi:10.1007/s00216-016-9880-6

Kannan, S., Balaram, P., Chandran, G. J., Pillai, M. R., Mathew, B., & Nair, M. K. (1993). Expression of lectin-specific cellular glycoconjugates during oral carcinogenesis. *J Cancer Res Clin Oncol, 119*(11), 689-694. doi:10.1007/bf01215989

Universiti Malaya

Kapila, S. N., Natarajan, S., & Boaz, K. (2017). A Comparison of Clinicopathological Differences in Oral Squamous Cell Carcinoma in Patients Below and Above 40 Years of Age. *J Clin Diagn Res, 11*(9), ZC46-ZC50. doi:10.7860/JCDR/2017/27828.10600

Karagiannis, G. S., Pavlou, M. P., & Diamandis, E. P. (2010). Cancer secretomics reveal pathophysiological pathways in cancer molecular oncology. *Mol Oncol, 4*(6), 496-510. doi:10.1016/j.molonc.2010.09.001

Karas, M., Ingendoh, A., Bahr, U., & Hillenkamp, F. (1989). Ultraviolet-Laser Desorption Ionization Mass-Spectrometry of Femtomolar Amounts of Large Proteins. *Biomedical and Environmental Mass Spectrometry, 18*(9), 841-843. doi:DOI 10.1002/bms.1200180931

Kaskas, N. M., Moore-Medlin, T., McClure, G. B., Ekshyyan, O., Vanchiere, J. A., & Nathan, C. A. (2014). Serum biomarkers in head and neck squamous cell cancer. *JAMA Otolaryngol Head Neck Surg, 140*(1), 5-11. doi:10.1001/jamaoto.2013.5688

Kataria, Y., Deaton, R. J., Enk, E., Jin, M., Petrauskaite, M., Dong, L., . . . Gann, P. H. (2016). Retinoid and carotenoid status in serum and liver among patients at high-risk for liver cancer. *BMC Gastroenterol, 16*, 30. doi:10.1186/s12876-016-0432-5

Kawahara, R., Bollinger, J. G., Rivera, C., Ribeiro, A. C., Brandao, T. B., Paes Leme, A. F., & MacCoss, M. J. (2016). A targeted proteomic strategy for the measurement of oral cancer candidate biomarkers in human saliva. *Proteomics, 16*(1), 159-173. doi:10.1002/pmic.201500224

Keam, S. P., Gulati, T., Gamell, C., Caramia, F., Huang, C., Schittenhelm, R. B., . . . Williams, S. G. (2018). Exploring the oncoproteomic response of human prostate cancer to therapeutic radiation using data-independent acquisition (DIA) mass spectrometry. *Prostate, 78*(8), 563-575. doi:10.1002/pros.23500

Kelemen, O., Pla, I., Sanchez, A., Rezeli, M., Szasz, A. M., Malm, J., . . . Marko-Varga, G. (2020). Proteomic analysis enables distinction of early- versus advanced-stage lung adenocarcinomas. *Clin Transl Med, 10*(2), e106. doi:10.1002/ctm2.106

Kesarwala, A. H., Krishna, M. C., & Mitchell, J. B. (2016). Oxidative stress in oral diseases. *Oral Dis, 22*(1), 9-18. doi:10.1111/odi.12300

Khanna, R., Vidhyarthi, A. K., Khanna, S., Singh, U., & Singh, U. C. (2012). Expression of p53 protein in leukoplakia and oral squamous cell carcinoma. *World Journal of Surgical Medical and Radiation Oncology, 1*(16).

Kim, Y. W., Bae, S. M., Lim, H., Kim, Y. J., & Ahn, W. S. (2012). Development of multiplexed bead-based immunoassays for the detection of early stage ovarian cancer using a combination of serum biomarkers. *PLoS One, 7*(9), e44960. doi:10.1371/journal.pone.0044960

Universiti Malaya

Kirwan, A., Utratna, M., O'Dwyer, M. E., Joshi, L., & Kilcoyne, M. (2015). Glycosylation-Based Serum Biomarkers for Cancer Diagnostics and Prognostics. *Biomed Res Int, 2015*, 490531. doi:10.1155/2015/490531

Kopp, H. G., Placke, T., & Salih, H. R. (2009). Platelet-derived transforming growth factor-beta down-regulates NKG2D thereby inhibiting natural killer cell antitumor reactivity. *Cancer Res, 69*(19), 7775-7783. doi:10.1158/0008-5472.CAN-09-2123

Korde, S. D., Basak, A., Chaudhary, M., Goyal, M., & Vagga, A. (2011). Enhanced nitrosative and oxidative stress with decreased total antioxidant capacity in patients with oral precancer and oral squamous cell carcinoma. *Oncology, 80*(5-6), 382-389. doi:10.1159/000329811

Kornfeld, R., & Kornfeld, S. (1985). Assembly of asparagine-linked oligosaccharides. *Annu Rev Biochem, 54*, 631-664. doi:10.1146/annurev.bi.54.070185.003215

Kozakiewicz, P., & Grzybowska-Szatkowska, L. (2018). Application of molecular targeted therapies in the treatment of head and neck squamous cell carcinoma. *Oncol Lett, 15*(5), 7497-7505. doi:10.3892/ol.2018.8300

Kratzer, T. B., Weinstein, S. J., Albanes, D., & Mondul, A. M. (2019). Vitamin D binding protein and risk of renal cell carcinoma in the prostate, lung, colorectal and ovarian cancer screening trial. *Int J Cancer*. doi:10.1002/ijc.32758

Krishna, A., Singh, S., Kumar, V., & Pal, U. S. (2015). Molecular concept in human oral cancer. *Natl J Maxillofac Surg, 6*(1), 9-15. doi:10.4103/0975-5950.168235

Kulasingam, V., & Diamandis, E. P. (2008). Strategies for discovering novel cancer biomarkers through utilization of emerging technologies. *Nat Clin Pract Oncol, 5*(10), 588-599. doi:10.1038/ncponc1187

Kurihara-Shimomura, M., Sasahira, T., Shimomura, H., & Kirita, T. (2020). Peroxidan Plays a Tumor-Promoting Role in Oral Squamous Cell Carcinoma. *Int J Mol Sci, 21*(15). doi:10.3390/ijms21155416

Kwak, J. W., Laskowski, J., Li, H. Y., McSharry, M. V., Sippel, T. R., Bullock, B. L., . . . Nemenoff, R. A. (2018). Complement Activation via a C3a Receptor Pathway Alters CD4(+) T Lymphocytes and Mediates Lung Cancer Progression. *Cancer Res, 78*(1), 143-156. doi:10.1158/0008-5472.CAN-17-0240

Lastovickova, M., Strouhalova, D., & Bobalova, J. (2020). Use of Lectin-based Affinity Techniques in Breast Cancer Glycoproteomics: A Review. *Journal of Proteome Research, 19*(5), 1885-1899. doi:10.1021/acs.jproteome.9b00818

Lee, D. S., Ramirez, R. J., Lee, J. J., Valenzuela, C. V., Zevallos, J. P., Mazul, A. L., . . . Jackson, R. S. (2020). Survival of Young Versus Old Patients With Oral Cavity Squamous Cell Carcinoma: A Meta-Analysis. *Laryngoscope*. doi:10.1002/lary.29260

Universiti Malaya

Lee, G., Liu, S., & Huang, C.-Y. (2014). Human serum proteins recognized by CA215 and cancerous immunoglobulins and implications in cancer immunology. *J Cancer Clin Oncol, 3*, 51-69.

Lei, T., Zhao, X., Jin, S., Meng, Q., Zhou, H., & Zhang, M. (2013). Discovery of potential bladder cancer biomarkers by comparative urine proteomics and analysis. *Clin Genitourin Cancer, 11*(1), 56-62. doi:10.1016/j.clgc.2012.06.003

Levy, A. P., Asleh, R., Blum, S., Levy, N. S., Miller-Lotan, R., Kalet-Litman, S., . . . Goldenstein, H. (2010). Haptoglobin: basic and clinical aspects. *Antioxid Redox Signal, 12*(2), 293-304. doi:10.1089/ars.2009.2793

Li, C., Li, H., Zhang, T., Li, J., Liu, L., & Chang, J. (2014). Discovery of Apo-A1 as a potential bladder cancer biomarker by urine proteomics and analysis. *Biochem Biophys Res Commun, 446*(4), 1047-1052. doi:10.1016/j.bbrc.2014.03.053

Li, D., Zhou, L., Ma, C., Chen, W., Zhang, Y., Yu, S., . . . Qiu, L. (2019). Comparative analysis of the serum proteome profiles of thyroid cancer: An initial focus on the lipid profile. *Oncol Lett, 18*(3), 3349-3357. doi:10.3892/ol.2019.10655

Li, M. H., Fu, S. B., & Xiao, H. S. (2015). Genome-wide analysis of microRNA and mRNA expression signatures in cancer. *Acta Pharmacol Sin, 36*(10), 1200-1211. doi:10.1038/aps.2015.67

Li, Q. K., Shah, P., Tian, Y., Hu, Y., Roden, R. B. S., Zhang, H., & Chan, D. W. (2017). An integrated proteomic and glycoproteomic approach uncovers differences in glycosylation occupancy from benign and malignant epithelial ovarian tumors. *Clin Proteomics, 14*, 16. doi:10.1186/s12014-017-9152-2

Li, Y., Tao, S. C., Bova, G. S., Liu, A. Y., Chan, D. W., Zhu, H., & Zhang, H. (2011). Detection and Verification of Glycosylation Patterns of Glycoproteins from Clinical Specimens Using Lectin Microarrays and Lectin-Based Immunosorbent Assays. *Analytical Chemistry, 83*(22), 8509-8516. doi:10.1021/ac201452f

Liang, Y., Ma, T., Thakur, A., Yu, H., Gao, L., Shi, P., . . . Chen, M. (2015). Differentially expressed glycosylated patterns of alpha-1-antitrypsin as serum biomarkers for the diagnosis of lung cancer. *Glycobiology, 25*(3), 331-340. doi:10.1093/glycob/cwu115

Liao, K. A., Tsay, Y. G., Huang, L. C., Huang, H. Y., Li, C. F., & Wu, T. F. (2011). Search for the tumor-associated proteins of oral squamous cell carcinoma collected in Taiwan using proteomics strategy. *Journal of Proteome Research, 10*(5), 2347-2358. doi:10.1021/pr101146w

Lihong, H., Linlin, G., Yiping, G., Yang, S., Xiaoyu, Q., Zhuzhu, G., . . . Shujuan, S. (2014). Proteomics approaches for identification of tumor relevant protein targets in pulmonary squamous cell carcinoma by 2D-DIGE-MS. *PLoS One, 9*(4), e95121. doi:10.1371/journal.pone.0095121

Lim, L. C., Looi, M. L., Zakaria, S. Z., Sagap, I., Rose, I. M., Chin, S. F., & Jamal, R. (2016). Identification of Differentially Expressed Proteins in the Serum of

Universiti Malaya

Colorectal Cancer Patients Using 2D-DIGE Proteomics Analysis. *Pathol Oncol Res, 22*(1), 169-177. doi:10.1007/s12253-015-9991-y

Lin, K., He, S., He, L., Chen, J., Cheng, X., Zhang, G., & Zhu, B. (2014). Complement component 3 is a prognostic factor of nonsmall cell lung cancer. *Mol Med Rep, 10*(2), 811-817. doi:10.3892/mmr.2014.2230

Lin, L. H., Xu, Y. W., Huang, L. S., Hong, C. Q., Zhai, T. T., Liao, L. D., . . . Peng, Y. H. (2017). Serum proteomic-based analysis identifying autoantibodies against PRDX2 and PRDX3 as potential diagnostic biomarkers in nasopharyngeal carcinoma. *Clinical Proteomics, 14*. doi:10.1186/s12014-017-9141-5

Lin, M. C., Huang, M. J., Liu, C. H., Yang, T. L., & Huang, M. C. (2014). GALNT2 enhances migration and invasion of oral squamous cell carcinoma by regulating EGFR glycosylation and activity. *Oral Oncol, 50*(5), 478-484. doi:10.1016/j.oraloncology.2014.02.003

Lin, W. J., Jiang, R. S., Wu, S. H., Chen, F. J., & Liu, S. A. (2011). Smoking, alcohol, and betel quid and oral cancer: a prospective cohort study. *J Oncol, 2011*, 525976. doi:10.1155/2011/525976

Lin, W. L., Lin, Y. S., Shi, G. Y., Chang, C. F., & Wu, H. L. (2015). Lewisy promotes migration of oral cancer cells by glycosylation of epidermal growth factor receptor. *PLoS One, 10*(3), e0120162. doi:10.1371/journal.pone.0120162

Lin, X., Hong, S., Huang, J., Chen, Y., Chen, Y., & Wu, Z. (2017). Plasma apolipoprotein A1 levels at diagnosis are independent prognostic factors in invasive ductal breast cancer. *Discov Med, 23*(127), 247-258.

Lingen, M. W., & Kumar, V. (2005). Head and neck. *Robbins and cotran pathologic basis of disease. 7th ed. Philadelphia: Elsevier, Saunders*, 773-795.

Linxweiler, J., Kollipara, L., Zahedi, R. P., Lampel, P., Zimmermann, R., & Greiner, M. (2014). Proteomic insights into non-small cell lung cancer: New ideas for cancer diagnosis and therapy from a functional viewpoint. *EuPA Open Proteomics, 4*, 25-39.

Lis, H., & Sharon, N. (1998). Lectins: Carbohydrate-Specific Proteins That Mediate Cellular Recognition. *Chem Rev, 98*(2), 637-674.

Liu, B., Han, M. T., Zhang, J., Lu, P., Li, J., Song, N., . . . Zhang, W. (2013). Downregulation of clusterin expression in human testicular seminoma. *Cell Physiol Biochem, 32*(4), 1117-1123. doi:10.1159/000354511

Liu, G., Sengupta, P. K., Jamal, B., Yang, H. Y., Bouchie, M. P., Lindner, V., . . . Kukuruzinska, M. A. (2013). N-glycosylation induces the CTHRC1 protein and drives oral cancer cell migration. *J Biol Chem, 288*(28), 20217-20227. doi:10.1074/jbc.M113.473785

Liu, L., So, A. Y.-L., & Fan, J.-B. (2015). Analysis of cancer genomes through microarrays and next-generation sequencing. *Translational Cancer Research, 4*(3), 212-218.

Universiti Malaya

Liu, T., Liu, D., Liu, R., Jiang, H., Yan, G., Li, W., . . . Guo, K. (2017). Discovering potential serological biomarker for chronic Hepatitis B Virus-related hepatocellular carcinoma in Chinese population by MAL-associated serum glycoproteomics analysis. *Sci Rep, 7*, 38918. doi:10.1038/srep38918

Liu, W., Liu, B., Xin, L., Zhang, Y., Chen, X., Zhu, Z., & Lin, Y. (2007). Down-regulated expression of complement factor I: a potential suppressive protein for gastric cancer identified by serum proteome analysis. *Clin Chim Acta, 377*(1-2), 119-126. doi:10.1016/j.cca.2006.09.005

Liu, Y., Chen, H., Sun, Z., & Chen, X. (2015). Molecular mechanisms of ethanol-associated oro-esophageal squamous cell carcinoma. *Cancer Lett, 361*(2), 164-173. doi:10.1016/j.canlet.2015.03.006

Lo, W. Y., Lai, C. C., Hua, C. H., Tsai, M. H., Huang, S. Y., Tsai, C. H., & Tsai, F. J. (2007). S100A8 is identified as a biomarker of HPV18-infected oral squamous cell carcinomas by suppression subtraction hybridization, clinical proteomics analysis, and immunohistochemistry staining. *Journal of Proteome Research, 6*(6), 2143-2151. doi:10.1021/pr060551+

Lorkova, L., Pospisilova, J., Lacheta, J., Leahomschi, S., Zivny, J., Cibula, D., . . . Petrak, J. (2012). Decreased concentrations of retinol-binding protein 4 in sera of epithelial ovarian cancer patients: a potential biomarker identified by proteomics. *Oncol Rep, 27*(2), 318-324. doi:10.3892/or.2011.1513

Lu, J., Wang, Y., Yan, M., Feng, P., Yuan, L., Cai, Y., . . . Li, L. (2016). High serum haptoglobin level is associated with tumor progression and predicts poor prognosis in non-small cell lung cancer. *Oncotarget, 7*(27), 41758-41766. doi:10.18632/oncotarget.9676

Maan, M., Peters, J. M., Dutta, M., & Patterson, A. D. (2018). Lipid metabolism and lipophagy in cancer. *Biochem Biophys Res Commun, 504*(3), 582-589. doi:10.1016/j.bbrc.2018.02.097

Mabert, K., Cojoc, M., Peitzsch, C., Kurth, I., Souchelnytskyi, S., & Dubrovska, A. (2014). Cancer biomarker discovery: current status and future perspectives. *Int J Radiat Biol, 90*(8), 659-677. doi:10.3109/09553002.2014.892229

Makridakis, M., Roubelakis, M. G., Bitsika, V., Dimuccio, V., Samiotaki, M., Kossida, S., . . . Vlahou, A. (2010). Analysis of secreted proteins for the study of bladder cancer cell aggressiveness. *Journal of Proteome Research, 9*(6), 3243-3259. doi:10.1021/pr100189d

Mangaraj, M., Nanda, R., & Panda, S. (2016). Apolipoprotein A-I: A Molecule of Diverse Function. *Indian J Clin Biochem, 31*(3), 253-259. doi:10.1007/s12291-015-0513-1

Manoharan, S., Padmanabhan, M., Kolanjiappan, K., Ramachandran, C. R., & Suresh, K. (2004). Analysis of glycoconjugates in patients with oral squamous cell carcinoma. *Clin Chim Acta, 339*(1-2), 91-96. doi:10.1016/j.cccn.2003.09.006

Universiti Malaya

Markiewski, M. M., DeAngelis, R. A., Benencia, F., Ricklin-Lichtsteiner, S. K., Koutoulaki, A., Gerard, C., . . . Lambris, J. D. (2008). Modulation of the antitumor immune response by complement. *Nat Immunol, 9*(11), 1225-1235. doi:10.1038/ni.1655

Markopoulos, A. K. (2012). Current aspects on oral squamous cell carcinoma. *Open Dent J, 6*, 126-130. doi:10.2174/1874210601206010126

Martin, L. J., Melnichouk, O., Huszti, E., Connelly, P. W., Greenberg, C. V., Minkin, S., & Boyd, N. F. (2015). Serum lipids, lipoproteins, and risk of breast cancer: a nested case-control study using multiple time points. *J Natl Cancer Inst, 107*(5). doi:10.1093/jnci/djv032

Marusyk, A., & Polyak, K. (2010). Tumor heterogeneity: causes and consequences. *Biochim Biophys Acta, 1805*(1), 105-117. doi:10.1016/j.bbcan.2009.11.002

Mathur, R., Singhavi, H. R., Malik, A., Nair, S., & Chaturvedi, P. (2019). Role of Poor Oral Hygiene in Causation of Oral Cancer-a Review of Literature. *Indian J Surg Oncol, 10*(1), 184-195. doi:10.1007/s13193-018-0836-5

Mazumdar, S., SenGupta, S. K., Param, R., & Sinha, S. N. (1993). Binding pattern of eight different lectins in healthy subjects and patients with dysplastic and malignant lesions of the oral cavity. *Int J Oral Maxillofac Surg, 22*(5), 301-305. doi:10.1016/s0901-5027(05)80521-1

Meany, D. L., & Chan, D. W. (2011). Aberrant glycosylation associated with enzymes as cancer biomarkers. *Clin Proteomics, 8*(1), 7. doi:10.1186/1559-0275-8-7

Medema, R. H., & Macurek, L. (2012). Checkpoint control and cancer. *Oncogene, 31*(21), 2601-2613. doi:10.1038/onc.2011.451

Medzihradszky, K. F., Campbell, J. M., Baldwin, M. A., Falick, A. M., Juhasz, P., Vestal, M. L., & Burlingame, A. L. (2000). The characteristics of peptide collision-induced dissociation using a high-performance MALDI-TOF/TOF tandem mass spectrometer. *Anal Chem, 72*(3), 552-558.

Milioli, H. H., Santos Sousa, K., Kaviski, R., Dos Santos Oliveira, N. C., De Andrade Urban, C., De Lima, R. S., . . . De Souza Fonseca Ribeiro, E. M. (2015). Comparative proteomics of primary breast carcinomas and lymph node metastases outlining markers of tumor invasion. *Cancer Genomics Proteomics, 12*(2), 89-101.

Miranda-Filho, A., & Bray, F. (2020). Global patterns and trends in cancers of the lip, tongue and mouth. *Oral Oncol, 102*, 104551. doi:10.1016/j.oraloncology.2019.104551

Miyamoto, R., Uzawa, N., Nagaoka, S., Hirata, Y., & Amagasa, T. (2003). Prognostic significance of cyclin D1 amplification and overexpression in oral squamous cell carcinomas. *Oral Oncol, 39*(6), 610-618.

Miyamoto, S., Ruhaak, L. R., Stroble, C., Salemi, M. R., Phinney, B., Lebrilla, C. B., & Leiserowitz, G. S. (2016). Glycoproteomic Analysis of Malignant Ovarian Cancer

Universiti Malaya

Ascites Fluid Identifies Unusual Glycopeptides. *Journal of Proteome Research, 15*(9), 3358-3376. doi:10.1021/acs.jproteome.6b00548

Montero, P. H., & Patel, S. G. (2015). Cancer of the oral cavity. *Surg Oncol Clin N Am, 24*(3), 491-508. doi:10.1016/j.soc.2015.03.006

Mori, K., Emoto, M., & Inaba, M. (2011). Fetuin-A: a multifunctional protein. *Recent Pat Endocr Metab Immune Drug Discov, 5*(2), 124-146.

Mu, A. K., Lim, B. K., Hashim, O. H., & Shuib, A. S. (2013). Identification of O-glycosylated proteins that are aberrantly excreted in the urine of patients with early stage ovarian cancer. *Int J Mol Sci, 14*(4), 7923-7931. doi:10.3390/ijms14047923

Munkley, J., & Elliott, D. J. (2016). Hallmarks of glycosylation in cancer. *Oncotarget, 7*(23), 35478-35489. doi:10.18632/oncotarget.8155

Murugan, A. K., Munirajan, A. K., & Tsuchida, N. (2012). Ras oncogenes in oral cancer: the past 20 years. *Oral Oncol, 48*(5), 383-392. doi:10.1016/j.oraloncology.2011.12.006

Nakahara, Y., Shintani, S., Mihara, M., Kiyota, A., Ueyama, Y., & Matsumura, T. (2000). Alterations of Rb, p16(INK4A) and cyclin D1 in the tumorigenesis of oral squamous cell carcinomas. *Cancer Lett, 160*(1), 3-8.

Naqvi, S. U., Khan, S., Ahmed, A., Lail, A., Gul, S., & Ahmed, S. (2020). Prevalence of EBV, CMV, and HPV in oral squamous cell carcinoma patients in the Pakistani population. *J Med Virol*. doi:10.1002/jmv.25791

Nedjadi, T., Benabdelkamal, H., Albarakati, N., Masood, A., Al-Sayyad, A., Alfadda, A. A., . . . Al-Maghrabi, J. (2020). Circulating proteomic signature for detection of biomarkers in bladder cancer patients. *Sci Rep, 10*(1), 10999. doi:10.1038/s41598-020-67929-z

Neville, B. W., & Day, T. A. (2002). Oral cancer and precancerous lesions. *CA Cancer J Clin, 52*(4), 195-215. doi:10.3322/canjclin.52.4.195

O'Farrell, P. H. (1975). High resolution two-dimensional electrophoresis of proteins. *J Biol Chem, 250*(10), 4007-4021.

Ola, M. S., Nawaz, M., & Ahsan, H. (2011). Role of Bcl-2 family proteins and caspases in the regulation of apoptosis. *Mol Cell Biochem, 351*(1-2), 41-58. doi:10.1007/s11010-010-0709-x

Omar, E. A. (2013). The outline of prognosis and new advances in diagnosis of oral squamous cell carcinoma (OSCC): review of the literature. *Journal of Oral Oncology, 2013*.

Omar, Z. A., & Tamin, N. S. I. (2011). National Cancer Registry Report: Malaysia Cancer Statistics–Data and Figure 2007. *Putrajaya: National Cancer Registry, Malaysia Ministry of Health*.

Universiti Malaya

Omura, K. (2014). Current status of oral cancer treatment strategies: surgical treatments for oral squamous cell carcinoma. *Int J Clin Oncol, 19*(3), 423-430. doi:10.1007/s10147-014-0689-z

Padhi, S. S., Roy, S., Kar, M., Saha, A., Roy, S., Adhya, A., . . . Banerjee, B. (2017). Role of CDKN2A/p16 expression in the prognostication of oral squamous cell carcinoma. *Oral Oncol, 73*, 27-35. doi:10.1016/j.oraloncology.2017.07.030

Pagel, O., Loroch, S., Sickmann, A., & Zahedi, R. P. (2015). Current strategies and findings in clinically relevant post-translational modification-specific proteomics. *Expert Rev Proteomics, 12*(3), 235-253. doi:10.1586/14789450.2015.1042867

Pai, R., Pai, S., Lalitha, R., Kumaraswamy, S., Lalitha, N., Johnston, R., & Bhargava, M. (2009). Over-expression of c-Myc oncoprotein in oral squamous cell carcinoma in the South Indian population. *Ecancermedicalscience, 3*, 128. doi:10.3332/ecancer.2008.128

Palasz, P., Adamski, L., Gorska-Chrzastek, M., Starzynska, A., & Studniarek, M. (2017). Contemporary Diagnostic Imaging of Oral Squamous Cell Carcinoma - A Review of Literature. *Pol J Radiol, 82*, 193-202. doi:10.12659/PJR.900892

Pan, S., Brentnall, T. A., & Chen, R. (2016). Glycoproteins and glycoproteomics in pancreatic cancer. *World J Gastroenterol, 22*(42), 9288-9299. doi:10.3748/wjg.v22.i42.9288

Pan, S., Chen, R., Aebersold, R., & Brentnall, T. A. (2011). Mass spectrometry based glycoproteomics--from a proteomics perspective. *Mol Cell Proteomics, 10*(1), R110 003251. doi:10.1074/mcp.R110.003251

Pande, P., Mathur, M., Shukla, N. K., & Ralhan, R. (1998). pRb and p16 protein alterations in human oral tumorigenesis. *Oral Oncol, 34*(5), 396-403.

Pang, W. W., Abdul-Rahman, P. S., Wan-Ibrahim, W. I., & Hashim, O. H. (2010). Can the acute-phase reactant proteins be used as cancer biomarkers? *Int J Biol Markers, 25*(1), 1-11.

Parkin, D. M., Bray, F., Ferlay, J., & Pisani, P. (2005). Global cancer statistics, 2002. *CA Cancer J Clin, 55*(2), 74-108.

Patel, A., Pathak, Y., Patel, J., & Sutariya, V. (2018). Role of nutritional factors in pathogenesis of cancer. *Food Quality and Safety, 2*(1), 27-36. doi:10.1093/fqsafe/fyx033

Pavlou, M. P., Diamandis, E. P., & Blasutig, I. M. (2013). The long journey of cancer biomarkers from the bench to the clinic. *Clin Chem, 59*(1), 147-157. doi:10.1373/clinchem.2012.184614

Pawlik, T. M., Hawke, D. H., Liu, Y., Krishnamurthy, S., Fritsche, H., Hunt, K. K., & Kuerer, H. M. (2006). Proteomic analysis of nipple aspirate fluid from women with early-stage breast cancer using isotope-coded affinity tags and tandem mass spectrometry reveals differential expression of vitamin D binding protein. *BMC Cancer, 6*, 68. doi:10.1186/1471-2407-6-68

Universiti Malaya

Peng, H., Yan, Z., Zeng, X., Zhang, S., Jiang, H., Huang, H., & Zhuo, H. (2019). Serum and tissue proteomic signatures of patients with hepatocellular carcinoma using 2D gel electrophoresis. *Mol Med Rep, 20*(2), 1025-1038. doi:10.3892/mmr.2019.10311

Peng, L., Cantor, D. I., Huang, C., Wang, K., Baker, M. S., & Nice, E. C. (2018). Tissue and plasma proteomics for early stage cancer detection. *Mol Omics, 14*(6), 405-423. doi:10.1039/c8mo00126j

Perera, M., Al-Hebshi, N. N., Perera, I., Ipe, D., Ulett, G. C., Speicher, D. J., . . . Johnson, N. W. (2018). Inflammatory Bacteriome and Oral Squamous Cell Carcinoma. *J Dent Res, 97*(6), 725-732. doi:10.1177/0022034518767118

Perez-Holanda, S., Blanco, I., Menendez, M., & Rodrigo, L. (2014). Serum concentration of alpha-1 antitrypsin is significantly higher in colorectal cancer patients than in healthy controls. *BMC Cancer, 14*, 355. doi:10.1186/1471-2407-14-355

Perez-Sayans, M., Suarez-Penaranda, J. M., Pilar, G. D., Barros-Angueira, F., Gandara-Rey, J. M., & Garcia-Garcia, A. (2011). What real influence does the proto-oncogene c-myc have in OSCC behavior? *Oral Oncol, 47*(8), 688-692. doi:10.1016/j.oraloncology.2011.05.016

Petti, S. (2003). Pooled estimate of world leukoplakia prevalence: a systematic review. *Oral Oncol, 39*(8), 770-780.

Petti, S., Masood, M., & Scully, C. (2013). The magnitude of tobacco smoking-betel quid chewing-alcohol drinking interaction effect on oral cancer in South-East Asia. A meta-analysis of observational studies. *PLoS One, 8*(11), e78999. doi:10.1371/journal.pone.0078999

Pindborg, J. J., Reichart, P. A., Smith, C. J., & Van der Waal, I. (1997). *Histological Typing of Cancer and Precancer of the Oral Mucosa: In Collaboration with LH Sobin and Pathologists in 9 Countries*: Springer Science & Business Media.

Pinho, S. S., & Reis, C. A. (2015). Glycosylation in cancer: mechanisms and clinical implications. *Nat Rev Cancer, 15*(9), 540-555. doi:10.1038/nrc3982

Poddar, N. K., Maurya, S. K., & Saxena, V. (2017). Role of serine proteases and inhibitors in cancer. In *Proteases in Physiology and Pathology* (pp. 257-287): Springer.

Polaskova, V., Kapur, A., Khan, A., Molloy, M. P., & Baker, M. S. (2010). High-abundance protein depletion: comparison of methods for human plasma biomarker discovery. *Electrophoresis, 31*(3), 471-482. doi:10.1002/elps.200900286

Pomastowski, P., & Buszewski, B. (2014). Two-dimensional gel electrophoresis in the light of new developments. *Trac-Trends in Analytical Chemistry, 53*, 167-177. doi:10.1016/j.trac.2013.09.010

Qi, Y. J., Ward, D. G., Pang, C., Wang, Q. M., Wei, W., Ma, J., . . . Johnson, P. J. (2014). Proteomic profiling of N-linked glycoproteins identifies ConA-binding

Universiti Malaya

procathepsin D as a novel serum biomarker for hepatocellular carcinoma. *Proteomics, 14*(2-3), 186-195. doi:10.1002/pmic.201300226

Rao, S. V. K., Mejia, G., Roberts-Thomson, K., & Logan, R. (2013). Epidemiology of oral cancer in Asia in the past decade--an update (2000-2012). *Asian Pac J Cancer Prev, 14*(10), 5567-5577.

Redondo, M., Rodrigo, I., Alcaide, J., Tellez, T., Roldan, M. J., Funez, R., . . . Jimenez, E. (2010). Clusterin expression is associated with decreased disease-free survival of patients with colorectal carcinomas. *Histopathology, 56*(7), 932-936. doi:10.1111/j.1365-2559.2010.03565.x

Rekha, K. P., & Angadi, P. V. (2010). Verrucous carcinoma of the oral cavity: a clinico-pathologic appraisal of 133 cases in Indians. *Oral Maxillofac Surg, 14*(4), 211-218. doi:10.1007/s10006-010-0222-0

Remani, P., Bhattathiri, V. N., Bindu, L., Chandralekha, B., Vijayakumar, T., & Nair, M. K. (1997). Correlation of lectin binding with lymph node metastasis in oral cancers. *Oral Oncol, 33*(1), 19-22. doi:10.1016/s0964-1955(96)00044-9

Ren, L., Yi, J., Li, W., Zheng, X., Liu, J., Wang, J., & Du, G. (2019). Apolipoproteins and cancer. *Cancer Med, 8*(16), 7032-7043. doi:10.1002/cam4.2587

Reuter, S., Gupta, S. C., Chaturvedi, M. M., & Aggarwal, B. B. (2010). Oxidative stress, inflammation, and cancer: how are they linked? *Free Radic Biol Med, 49*(11), 1603-1616. doi:10.1016/j.freeradbiomed.2010.09.006

Rho, J. H., Roehrl, M. H., & Wang, J. Y. (2009). Glycoproteomic analysis of human lung adenocarcinomas using glycoarrays and tandem mass spectrometry: differential expression and glycosylation patterns of vimentin and fetuin A isoforms. *Protein J, 28*(3-4), 148-160. doi:10.1007/s10930-009-9177-0

Riihila, P., Nissinen, L., Farshchian, M., Kallajoki, M., Kivisaari, A., Meri, S., . . . Kahari, V. M. (2017). Complement Component C3 and Complement Factor B Promote Growth of Cutaneous Squamous Cell Carcinoma. *Am J Pathol, 187*(5), 1186-1197. doi:10.1016/j.ajpath.2017.01.006

Rodriguez-Pineiro, A. M., Ayude, D., Rodriguez-Berrocal, F. J., & de la Cadena, M. P. (2004). Concanavalin A chromatography coupled to two-dimensional gel electrophoresis improves protein expression studies of the serum proteome. *J Chromatogr B Analyt Technol Biomed Life Sci, 803*(2), 337-343. doi:10.1016/j.jchromb.2004.01.019

Roman, E., Lunde, M. L., Miron, T., Warnakulasauriya, S., Johannessen, A. C., Vasstrand, E. N., & Ibrahim, S. O. (2013). Analysis of protein expression profile of oral squamous cell carcinoma by MALDI-TOF-MS. *Anticancer Res, 33*(3), 837-845.

Roth, Z., Yehezkel, G., & Khalaila, I. (2012). Identification and quantification of protein glycosylation. *International Journal of Carbohydrate Chemistry, 2012*. doi:10.1155/2012/640923

Universiti Malaya

Rowley, H., Sherrington, P., Helliwell, T. R., Kinsella, A., & Jones, A. S. (1998). p53 expression and p53 gene mutation in oral cancer and dysplasia. *Otolaryngol Head Neck Surg, 118*(1), 115-123. doi:10.1016/S0194-5998(98)70387-0

Saawarn, S., Astekar, M., Saawarn, N., Dhakar, N., & Gomateshwar Sagari, S. (2012). Cyclin d1 expression and its correlation with histopathological differentiation in oral squamous cell carcinoma. *ScientificWorldJournal, 2012*, 978327. doi:10.1100/2012/978327

Said, I. T., Shamsuddin, A. M., Sherief, M. A., Taleb, S. G., Aref, W. F., & Kumar, D. (1999). Comparison of different techniques for detection of Gal-GalNAc, an early marker of colonic neoplasia. *Histol Histopathol, 14*(2), 351-357. doi:10.14670/HH-14.351

Sakwe, A. M., Koumangoye, R., Goodwin, S. J., & Ochieng, J. (2010). Fetuin-A ({alpha}2HS-glycoprotein) is a major serum adhesive protein that mediates growth signaling in breast tumor cells. *J Biol Chem, 285*(53), 41827-41835. doi:10.1074/jbc.M110.128926

Saleem, M., Raza, S. K., & S, G. M. (2019). A comparative protein analysis of lung cancer, along with three controls using a multidimensional proteomic approach. *Exp Biol Med (Maywood), 244*(1), 36-41. doi:10.1177/1535370219826525

Sallam, R. M. (2015). Proteomics in cancer biomarkers discovery: challenges and applications. *Dis Markers, 2015*, 321370. doi:10.1155/2015/321370

Sanchez, J. C., Appel, R. D., Golaz, O., Pasquali, C., Ravier, F., Bairoch, A., & Hochstrasser, D. F. (1995). Inside SWISS-2DPAGE database. *Electrophoresis, 16*(7), 1131-1151.

Sand, L., & Jalouli, J. (2014). Viruses and oral cancer. Is there a link? *Microbes Infect, 16*(5), 371-378. doi:10.1016/j.micinf.2014.02.009

Sankaranarayanan, R., Ramadas, K., Amarasinghe, H., Subramanian, S., & Johnson, N. (2015). Oral Cancer: Prevention, Early Detection, and Treatment. In H. Gelband, P. Jha, R. Sankaranarayanan, & S. Horton (Eds.), *Cancer: Disease Control Priorities, Third Edition (Volume 3)*. Washington (DC).

Santos, C. R., & Schulze, A. (2012). Lipid metabolism in cancer. *FEBS J, 279*(15), 2610-2623. doi:10.1111/j.1742-4658.2012.08644.x

Saraswat, M., Joenvaara, S., Seppanen, H., Mustonen, H., Haglund, C., & Renkonen, R. (2017). Comparative proteomic profiling of the serum differentiates pancreatic cancer from chronic pancreatitis. *Cancer Med, 6*(7), 1738-1751. doi:10.1002/cam4.1107

Saraswat, M., Makitie, A., Tohmola, T., Dickinson, A., Saraswat, S., Joenvaara, S., & Renkonen, S. (2018). Tongue Cancer Patients Can be Distinguished from Healthy Controls by Specific N-Glycopeptides Found in Serum. *Proteomics Clin Appl, 12*(6), e1800061. doi:10.1002/prca.201800061

Universiti Malaya

Saravani, S., Kadeh, H., Miri-Moghaddam, E., Zekri, A., Sanadgol, N., & Gholami, A. (2015). Human Cytomegalovirus in Oral Squamous Cell Carcinoma in Southeast of Iran. *Jundishapur J Microbiol, 8*(8), e21838. doi:10.5812/jjm.21838

Sarkis, S. A., Abdullah, B. H., Majeed, B. A. A., & Talabani, N. G. (2010). Immunohistochemical expression of epidermal growth factor receptor (EGFR) in oral squamous cell carcinoma in relation to proliferation, apoptosis, angiogenesis and lymphangiogenesis. *Head Neck Oncol, 2*, 13. doi:10.1186/1758-3284-2-13

Sarode, G., Sarode, S. C., Tupkari, J., & Patil, S. (2017). Is oral squamous cell carcinoma unique in terms of intra-and inter-tumoral heterogeneity? *Translational Research in Oral Oncology, 2*, 2057178X17703578.

Sarrats, A., Saldova, R., Pla, E., Fort, E., Harvey, D. J., Struwe, W. B., . . . Peracaula, R. (2010). Glycosylation of liver acute-phase proteins in pancreatic cancer and chronic pancreatitis. *Proteomics Clin Appl, 4*(4), 432-448. doi:10.1002/prca.200900150

Satiman, E. A. F. E. N., Ahmad, H., Ramzi, A. B., Abdul Wahab, R., Kaderi, M. A., Wan Harun, W. H. A., . . . Arzmi, M. H. (2020). The role of Candida albicans candidalysin ECE1 gene in oral carcinogenesis. *J Oral Pathol Med, 49*(9), 835-841. doi:10.1111/jop.13014

Scaltriti, M., Brausi, M., Amorosi, A., Caporali, A., D'Arca, D., Astancolle, S., . . . Bettuzzi, S. (2004). Clusterin (SGP-2, ApoJ) expression is downregulated in low- and high-grade human prostate cancer. *Int J Cancer, 108*(1), 23-30. doi:10.1002/ijc.11496

Schmitd, L. B., Tjioe, K. C., Assao, A., & Oliveira, D. T. (2015). Oral Squamous Cell Carcinoma in Young Population—Risk Factors, Clinical Presentation, and Prognosis. In *Contemporary Issues in Head and Neck Cancer Management*: IntechOpen.

Schulman, I. G. (2017). Liver X receptors link lipid metabolism and inflammation. *FEBS Lett, 591*(19), 2978-2991. doi:10.1002/1873-3468.12702

Sekikawa, S., Onda, T., Miura, N., Nomura, T., Takano, N., Shibahara, T., & Honda, K. (2018). Underexpression of alpha-1-microglobulin/bikunin precursor predicts a poor prognosis in oral squamous cell carcinoma. *Int J Oncol, 53*(6), 2605-2614. doi:10.3892/ijo.2018.4581

Semaan, S. M., Wang, X., Marshall, A. G., & Sang, Q. X. (2012). Identification of Potential Glycoprotein Biomarkers in Estrogen Receptor Positive (ER+) and Negative (ER-) Human Breast Cancer Tissues by LC-LTQ/FT-ICR Mass Spectrometry. *J Cancer, 3*, 269-284. doi:10.7150/jca.4592

Shah, M., Telang, S., Raval, G., Shah, P., & Patel, P. S. (2008). Serum fucosylation changes in oral cancer and oral precancerous conditions: alpha-L-fucosidase as a marker. *Cancer, 113*(2), 336-346. doi:10.1002/cncr.23556

Shah, P., Wang, X., Yang, W., Toghi Eshghi, S., Sun, S., Hoti, N., . . . Zhang, H. (2015). Integrated Proteomic and Glycoproteomic Analyses of Prostate Cancer Cells

Universiti Malaya

Reveal Glycoprotein Alteration in Protein Abundance and Glycosylation. *Mol Cell Proteomics, 14*(10), 2753-2763. doi:10.1074/mcp.M115.047928

Shan, J., Sun, Z., Yang, J., Xu, J., Shi, W., Wu, Y., . . . Li, H. (2019). Discovery and preclinical validation of proteomic biomarkers in saliva for early detection of oral squamous cell carcinomas. *Oral Dis, 25*(1), 97-107. doi:10.1111/odi.12971

Shannan, B., Seifert, M., Leskov, K., Willis, J., Boothman, D., Tilgen, W., & Reichrath, J. (2006). Challenge and promise: roles for clusterin in pathogenesis, progression and therapy of cancer. *Cell Death Differ, 13*(1), 12-19. doi:10.1038/sj.cdd.4401779

Sharon, N., & Lis, H. (2004). History of lectins: from hemagglutinins to biological recognition molecules. *Glycobiology, 14*(11), 53R-62R. doi:10.1093/glycob/cwh122

Shevchenko, A., Wilm, M., Vorm, O., & Mann, M. (1996). Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Anal Chem, 68*(5), 850-858.

Shi, F., Wu, H., Qu, K., Sun, Q., Li, F., Shi, C., . . . She, J. (2018). Identification of serum proteins AHSG, FGA and APOA-I as diagnostic biomarkers for gastric cancer. *Clin Proteomics, 15*, 18. doi:10.1186/s12014-018-9194-0

Shield-Artin, K. L., Bailey, M. J., Oliva, K., Liovic, A. K., Barker, G., Dellios, N. L., . . . Rice, G. E. (2012). Identification of ovarian cancer-associated proteins in symptomatic women: A novel method for semi-quantitative plasma proteomics. *Proteomics Clin Appl, 6*(3-4), 170-181. doi:10.1002/prca.201100008

Shukla, K., Vun, I., Lov, I., Laparidis, G., McCamley, C., & Ariyawardana, A. (2019). Role of Candida infection in the malignant transformation of oral leukoplakia: A systematic review of observational studies. *Translational Research in Oral Oncology, 4*, 2057178X19828229. doi:10.1177/2057178X19828229

Solier, C., & Langen, H. (2014). Antibody-based proteomics and biomarker research - current status and limitations. *Proteomics, 14*(6), 774-783. doi:10.1002/pmic.201300334

Soni, S., Kaur, J., Kumar, A., Chakravarti, N., Mathur, M., Bahadur, S., . . . Ralhan, R. (2005). Alterations of rb pathway components are frequent events in patients with oral epithelial dysplasia and predict clinical outcome in patients with squamous cell carcinoma. *Oncology, 68*(4-6), 314-325. doi:10.1159/000086970

Speeckaert, M. M., Speeckaert, R., van Geel, N., & Delanghe, J. R. (2014). Vitamin D binding protein: a multifunctional protein of clinical importance. *Adv Clin Chem, 63*, 1-57. doi:10.1016/b978-0-12-800094-6.00001-7

Speight, P. M., Khurram, S. A., & Kujan, O. (2018). Oral potentially malignant disorders: risk of progression to malignancy. *Oral Surg Oral Med Oral Pathol Oral Radiol, 125*(6), 612-627. doi:10.1016/j.oooo.2017.12.011

Universiti Malaya

Srinivasan, M., & Jewell, S. D. (2001). Evaluation of TGF-alpha and EGFR expression in oral leukoplakia and oral submucous fibrosis by quantitative immunohistochemistry. *Oncology, 61*(4), 284-292. doi:10.1159/000055335

Steen, P. V. D., Rudd, P. M., Dwek, R. A., & Opdenakker, G. (1998). Concepts and principles of O-linked glycosylation. *Crit Rev Biochem Mol Biol, 33*(3), 151-208. doi:10.1080/10409239891204198

Steentoft, C., Vakhrushev, S. Y., Joshi, H. J., Kong, Y., Vester-Christensen, M. B., Schjoldager, K. T., . . . Clausen, H. (2013). Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. *EMBO J, 32*(10), 1478-1488. doi:10.1038/emboj.2013.79

Stempfer, R., Kubicek, M., Lang, I. M., Christa, N., & Gerner, C. (2008). Quantitative assessment of human serum high-abundance protein depletion. *Electrophoresis, 29*(21), 4316-4323. doi:10.1002/elps.200800211

Stowell, S. R., Ju, T., & Cummings, R. D. (2015). Protein glycosylation in cancer. *Annu Rev Pathol, 10*, 473-510. doi:10.1146/annurev-pathol-012414-040438

Strimbu, K., & Tavel, J. A. (2010). What are biomarkers? *Curr Opin HIV AIDS, 5*(6), 463-466. doi:10.1097/COH.0b013e32833ed177

Su, C.-C., Sheng, K.-H., Chen, S.-Y., Tsay, Y.-G., & Wu, T.-F. (2016). Searching urinary tumor-associated proteins for bladder transitional cell carcinoma in southwestern Taiwan using gel-based proteomics. *Urological Science, 27*(4), 250-257.

Sun, L., Hu, S., Yu, L., Guo, C., Sun, L., Yang, Z., . . . Ran, Y. (2016). Serum haptoglobin as a novel molecular biomarker predicting colorectal cancer hepatic metastasis. *Int J Cancer, 138*(11), 2724-2731. doi:10.1002/ijc.29993

Swiatly, A., Horala, A., Matysiak, J., Hajduk, J., Nowak-Markwitz, E., & Kokot, Z. J. (2018). Understanding Ovarian Cancer: iTRAQ-Based Proteomics for Biomarker Discovery. *Int J Mol Sci, 19*(8). doi:10.3390/ijms19082240

Tabassum, U., Reddy, O., & Mukherjee, G. (2012). Elevated serum haptoglobin is associated with clinical outcome in triple-negative breast cancer patients. *Asian Pac J Cancer Prev, 13*(9), 4541-4544. doi:10.7314/apjcp.2012.13.9.4541

Tai, C. S., Lin, Y. R., Teng, T. H., Lin, P. Y., Tu, S. J., Chou, C. H., . . . Chen, W. L. (2017). Haptoglobin expression correlates with tumor differentiation and five-year overall survival rate in hepatocellular carcinoma. *PLoS One, 12*(2), e0171269. doi:10.1371/journal.pone.0171269

Tan, A. A., Phang, W. M., Gopinath, S. C., Hashim, O. H., Kiew, L. V., & Chen, Y. (2015). Revealing Glycoproteins in the Secretome of MCF-7 Human Breast Cancer Cells. *Biomed Res Int, 2015*, 453289. doi:10.1155/2015/453289

Tanaka, T., & Ishigamori, R. (2011). Understanding carcinogenesis for fighting oral cancer. *J Oncol, 2011*, 603740. doi:10.1155/2011/603740

Universiti Malaya

Tang, D., Tao, D., Fang, Y., Deng, C., Xu, Q., & Zhou, J. (2017). TNF-Alpha Promotes Invasion and Metastasis via NF-Kappa B Pathway in Oral Squamous Cell Carcinoma. *Med Sci Monit Basic Res, 23*, 141-149. doi:10.12659/msmbr.903910

Thiel, U. J., Feltens, R., Adryan, B., Gieringer, R., Brochhausen, C., Schuon, R., . . . Brieger, J. (2011). Analysis of differentially expressed proteins in oral squamous cell carcinoma by MALDI-TOF MS. *J Oral Pathol Med, 40*(5), 369-379. doi:10.1111/j.1600-0714.2010.00982.x

Thompson, P. D., Sakwe, A., Koumangoye, R., Yarbrough, W. G., Ochieng, J., & Marshall, D. R. (2014). Alpha-2 Heremans Schmid Glycoprotein (AHSG) modulates signaling pathways in head and neck squamous cell carcinoma cell line SQ20B. *Exp Cell Res, 321*(2), 123-132. doi:10.1016/j.yexcr.2013.12.003

Tian, W. D., Li, J. Z., Hu, S. W., Peng, X. W., Li, G., Liu, X., . . . Li, X. P. (2015). Proteomic identification of alpha-2-HS-glycoprotein as a plasma biomarker of hypopharyngeal squamous cell carcinoma. *Int J Clin Exp Pathol, 8*(8), 9021-9031.

Torres-Luquis, O., Madden, K., N'Dri N, M., Berg, R., Olopade, O. F., Ngwa, W., . . . Mohammed, S. I. (2019). LXR/RXR pathway signaling associated with triple-negative breast cancer in African American women. *Breast Cancer (Dove Med Press), 11*, 1-12. doi:10.2147/BCTT.S185960

Tran, T., Nguyen, T., Nguyen, B., & Phan, V. (2008). Changes of serum glycoproteins in lung cancer patients. *J Proteomics Bioinform, 1*(1), 011-016.

Tu, C., Rudnick, P. A., Martinez, M. Y., Cheek, K. L., Stein, S. E., Slebos, R. J., & Liebler, D. C. (2010). Depletion of abundant plasma proteins and limitations of plasma proteomics. *Journal of Proteome Research, 9*(10), 4982-4991. doi:10.1021/pr100646w

Tung, C. L., Lin, S. T., Chou, H. C., Chen, Y. W., Lin, H. C., Tung, C. L., . . . Chan, H. L. (2013). Proteomics-based identification of plasma biomarkers in oral squamous cell carcinoma. *J Pharm Biomed Anal, 75*, 7-17. doi:10.1016/j.jpba.2012.11.017

Turhani, D., Krapfenbauer, K., Thurnher, D., Langen, H., & Fountoulakis, M. (2006). Identification of differentially expressed, tumor-associated proteins in oral squamous cell carcinoma by proteomic analysis. *Electrophoresis, 27*(7), 1417-1423. doi:10.1002/elps.200500510

Uen, Y. H., Lin, K. Y., Sun, D. P., Liao, C. C., Hsieh, M. S., Huang, Y. K., . . . Lin, C. Y. (2013). Comparative proteomics, network analysis and post-translational modification identification reveal differential profiles of plasma Con A-bound glycoprotein biomarkers in gastric cancer. *J Proteomics, 83*, 197-213. doi:10.1016/j.jprot.2013.03.007

Uribarri, M., Hormaeche, I., Zalacain, R., Lopez-Vivanco, G., Martinez, A., Nagore, D., & Ruiz-Arguello, M. B. (2014). A new biomarker panel in bronchoalveolar lavage for an improved lung cancer diagnosis. *J Thorac Oncol, 9*(10), 1504-1512. doi:10.1097/JTO.0000000000000282

Universiti Malaya

Vadrevu, S. K., Chintala, N. K., Sharma, S. K., Sharma, P., Cleveland, C., Riediger, L., . . . Markiewski, M. M. (2014). Complement c5a receptor facilitates cancer metastasis by altering T-cell responses in the metastatic niche. *Cancer Res, 74*(13), 3454-3465. doi:10.1158/0008-5472.CAN-14-0157

Vijayan, K. K., Remani, P., Beevi, V. M., Ankathil, R., Vijayakumar, T., Rajendran, R., . . . Vasudevan, D. M. (1987). Tissue binding patterns of lectins in premalignant and malignant lesions of the oral cavity. *J Exp Pathol, 3*(3), 295-304.

Villa, A., Villa, C., & Abati, S. (2011). Oral cancer and oral erythroplakia: an update and implication for clinicians. *Aust Dent J, 56*(3), 253-256. doi:10.1111/j.1834-7819.2011.01337.x

Walsh, C. T., Garneau-Tsodikova, S., & Gatto, G. J., Jr. (2005). Protein posttranslational modifications: the chemistry of proteome diversifications. *Angew Chem Int Ed Engl, 44*(45), 7342-7372. doi:10.1002/anie.200501023

Wang, D. B., Lu, X. K., Zhang, X., Li, Z. G., & Li, C. X. (2016). Carbonic anhydrase 1 is a promising biomarker for early detection of non-small cell lung cancer. *Tumour Biol, 37*(1), 553-559. doi:10.1007/s13277-015-3834-z

Wang, H., Peng, R., Chen, X., Jia, R., Huang, C., Huang, Y., . . . Guo, G. (2018). Effect of HK2, PKM2 and LDHA on Cetuximab efficacy in metastatic colorectal cancer. *Oncol Lett, 15*(4), 5553-5560. doi:10.3892/ol.2018.8005

Wang, X. P., Li, X. H., Zhang, L., Lin, J. H., Huang, H., Kang, T., . . . Zheng, X. (2016). High level of serum apolipoprotein A-I is a favorable prognostic factor for overall survival in esophageal squamous cell carcinoma. *BMC Cancer, 16*, 516. doi:10.1186/s12885-016-2502-z

Wang, Y., Kuramitsu, Y., Yoshino, S., Takashima, M., Zhang, X., Ueno, T., . . . Nakamura, K. (2011). Screening for serological biomarkers of pancreatic cancer by two-dimensional electrophoresis and liquid chromatography-tandem mass spectrometry. *Oncol Rep, 26*(1), 287-292. doi:10.3892/or.2011.1278

Wang, Y., Liu, Y. H., Mai, S. J., He, L. J., Liao, Y. J., Deng, H. X., . . . Xie, D. (2010). Evaluation of serum clusterin as a surveillance tool for human hepatocellular carcinoma with hepatitis B virus related cirrhosis. *J Gastroenterol Hepatol, 25*(6), 1123-1128. doi:10.1111/j.1440-1746.2009.06205.x

Wang, Y. P., Chen, C., Hua, Q. Q., Wang, L., Li, F., Li, M., . . . Tao, Z. Z. (2017). Downregulation of leucine-rich-alpha-2-glycoprotein 1 expression is associated with the tumorigenesis of head and neck squamous cell carcinoma. *Oncology Reports, 37*(3), 1503-1510. doi:10.3892/or.2017.5377

Wang, Z., Feng, X., Liu, X., Jiang, L., Zeng, X., Ji, N., . . . Chen, Q. (2009). Involvement of potential pathways in malignant transformation from oral leukoplakia to oral squamous cell carcinoma revealed by proteomic analysis. *BMC Genomics, 10*, 383. doi:10.1186/1471-2164-10-383

Wang, Z., Wang, C., Huang, X., Shen, Y., Shen, J., & Ying, K. (2012). Differential proteome profiling of pleural effusions from lung cancer and benign inflammatory

Universiti Malaya

disease patients. *Biochim Biophys Acta, 1824*(4), 692-700. doi:10.1016/j.bbapap.2012.01.016

Warnakulasuriya, S. (2009). Global epidemiology of oral and oropharyngeal cancer. *Oral Oncol, 45*(4-5), 309-316. doi:10.1016/j.oraloncology.2008.06.002

Warnakulasuriya, S. (2014). Prognostic and predictive markers for oral squamous cell carcinoma: the importance of clinical, pathological and molecular markers. *Saudi Journal of Medicine and Medical Sciences, 2*(1), 12.

Warnakulasuriya, S., Johnson, N. W., & van der Waal, I. (2007). Nomenclature and classification of potentially malignant disorders of the oral mucosa. *J Oral Pathol Med, 36*(10), 575-580. doi:10.1111/j.1600-0714.2007.00582.x

Warnakulasuriya, S., Reibel, J., Bouquot, J., & Dabelsteen, E. (2008). Oral epithelial dysplasia classification systems: predictive value, utility, weaknesses and scope for improvement. *J Oral Pathol Med, 37*(3), 127-133. doi:10.1111/j.1600-0714.2007.00584.x

Whitehouse, C. M., Dreyer, R. N., Yamashita, M., & Fenn, J. B. (1985). Electrospray interface for liquid chromatographs and mass spectrometers. *Anal Chem, 57*(3), 675-679.

Wilkins, M. R., Sanchez, J. C., Gooley, A. A., Appel, R. D., Humphery-Smith, I., Hochstrasser, D. F., & Williams, K. L. (1996). Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it. *Biotechnol Genet Eng Rev, 13*, 19-50.

Wu, C. C., Chu, H. W., Hsu, C. W., Chang, K. P., & Liu, H. P. (2015). Saliva proteome profiling reveals potential salivary biomarkers for detection of oral cavity squamous cell carcinoma. *Proteomics, 15*(19), 3394-3404. doi:10.1002/pmic.201500157

Wu, J., Xie, X., Nie, S., Buckanovich, R. J., & Lubman, D. M. (2013). Altered expression of sialylated glycoproteins in ovarian cancer sera using lectin-based ELISA assay and quantitative glycoproteomics analysis. *Journal of Proteome Research, 12*(7), 3342-3352. doi:10.1021/pr400169n

Yang, E. C., Tan, M. T., Schwarz, R. A., Richards-Kortum, R. R., Gillenwater, A. M., & Vigneswaran, N. (2018). Noninvasive diagnostic adjuncts for the evaluation of potentially premalignant oral epithelial lesions: current limitations and future directions. *Oral Surg Oral Med Oral Pathol Oral Radiol, 125*(6), 670-681. doi:10.1016/j.oooo.2018.02.020

Yang, N., Feng, S., Shedden, K., Xie, X., Liu, Y., Rosser, C. J., . . . Goodison, S. (2011). Urinary glycoprotein biomarker discovery for bladder cancer detection using LC/MS-MS and label-free quantification. *Clin Cancer Res, 17*(10), 3349-3359. doi:10.1158/1078-0432.CCR-10-3121

Yang, Y., Franc, V., & Heck, A. J. R. (2017). Glycoproteomics: A Balance between High-Throughput and In-Depth Analysis. *Trends Biotechnol, 35*(7), 598-609. doi:10.1016/j.tibtech.2017.04.010

Universiti Malaya

Yates, J. R., Ruse, C. I., & Nakorchevsky, A. (2009). Proteomics by mass spectrometry: approaches, advances, and applications. *Annu Rev Biomed Eng, 11*, 49-79. doi:10.1146/annurev-bioeng-061008-124934

Yekta, R. F., Oskouie, A. A., Tavirani, M. R., Mohajeri-Tehrani, M. R., & Soroush, A. R. (2018). Decreased apolipoprotein A4 and increased complement component 3 as potential markers for papillary thyroid carcinoma: A proteomic study. *International Journal of Biological Markers, 33*(4), 455-462. doi:10.1177/1724600818787752

Yu, C. J., Chang, K. P., Chang, Y. J., Hsu, C. W., Liang, Y., Yu, J. S., . . . Wu, C. C. (2011). Identification of Guanylate-Binding Protein 1 as a Potential Oral Cancer Marker Involved in Cell Invasion Using Omics-Based Analysis. *Journal of Proteome Research, 10*(8), 3778-3788. doi:10.1021/pr2004133

Yu, J. S., Chen, Y. T., Chiang, W. F., Hsiao, Y. C., Chu, L. J., See, L. C., . . . Hartwell, L. H. (2016). Saliva protein biomarkers to detect oral squamous cell carcinoma in a high-risk population in Taiwan. *Proc Natl Acad Sci U S A, 113*(41), 11549-11554. doi:10.1073/pnas.1612368113

Yu, T., Wu, Y., Helman, J. I., Wen, Y., Wang, C., & Li, L. (2011). CXCR4 promotes oral squamous cell carcinoma migration and invasion through inducing expression of MMP-9 and MMP-13 via the ERK signaling pathway. *Mol Cancer Res, 9*(2), 161-172. doi:10.1158/1541-7786.MCR-10-0386

Yue, T. T., Goldstein, I. J., Hollingsworth, M. A., Kaul, K., Brand, R. E., & Haab, B. B. (2009). The Prevalence and Nature of Glycan Alterations on Specific Proteins in Pancreatic Cancer Patients Revealed Using Antibody-Lectin Sandwich Arrays. *Molecular & Cellular Proteomics, 8*(7), 1697-1707. doi:10.1074/mcp.M900135-MCP200

Zamanian-Daryoush, M., & DiDonato, J. A. (2015). Apolipoprotein A-I and Cancer. *Front Pharmacol, 6*, 265. doi:10.3389/fphar.2015.00265

Zeng, Z., Hincapie, M., Pitteri, S. J., Hanash, S., Schalkwijk, J., Hogan, J. M., . . . Hancock, W. S. (2011). A proteomics platform combining depletion, multi-lectin affinity chromatography (M-LAC), and isoelectric focusing to study the breast cancer proteome. *Anal Chem, 83*(12), 4845-4854. doi:10.1021/ac2002802

Zhang, C., Deng, X., Qiu, L., Peng, F., Geng, S., Shen, L., & Luo, Z. (2018). Knockdown of C1GalT1 inhibits radioresistance of human esophageal cancer cells through modifying β1-integrin glycosylation. *Journal of Cancer, 9*(15), 2666.

Zhang, G., Annan, R. S., Carr, S. A., & Neubert, T. A. (2014). Overview of peptide and protein analysis by mass spectrometry. *Curr Protoc Mol Biol, 108*, 10 21 11-30. doi:10.1002/0471142727.mb1021s108

Zhang, H., Liu, A. Y., Loriaux, P., Wollscheid, B., Zhou, Y., Watts, J. D., & Aebersold, R. (2007). Mass spectrometric detection of tissue proteins in plasma. *Mol Cell Proteomics, 6*(1), 64-71. doi:10.1074/mcp.M600160-MCP200

Universiti Malaya

Zhang, L. Y., Ying, W. T., Mao, Y. S., He, H. Z., Liu, Y., Wang, H. X., . . . Zhao, X. H. (2003). Loss of clusterin both in serum and tissue correlates with the tumorigenesis of esophageal squamous cell carcinoma via proteomics approaches. *World J Gastroenterol, 9*(4), 650-654. doi:10.3748/wjg.v9.i4.650

Zhang, Q., Huang, R., Tang, Q., Yu, Y., Huang, Q., Chen, Y., . . . Wang, X. (2018). Leucine-rich alpha-2-glycoprotein-1 is up-regulated in colorectal cancer and is a tumor promoter. *Onco Targets Ther, 11*, 2745-2752. doi:10.2147/OTT.S153375

Zhang, S., Shang, S., Li, W., Qin, X., & Liu, Y. (2016). Insights on N-glycosylation of human haptoglobin and its association with cancers. *Glycobiology, 26*(7), 684-692. doi:10.1093/glycob/cww016

Zhang, Z., Bast, R. C., Jr., Yu, Y., Li, J., Sokoll, L. J., Rai, A. J., . . . Chan, D. W. (2004). Three biomarkers identified from serum proteomic analysis for the detection of early stage ovarian cancer. *Cancer Res, 64*(16), 5882-5890. doi:10.1158/0008-5472.CAN-04-0746

Zhao, C., Annamalai, L., Guo, C., Kothandaraman, N., Koh, S. C., Zhang, H., . . . Choolani, M. (2007). Circulating haptoglobin is an independent prognostic factor in the sera of patients with epithelial ovarian cancer. *Neoplasia, 9*(1), 1-7. doi:10.1593/neo.06619

Zhao, J., Fan, Y. X., Yang, Y., Liu, D. L., Wu, K., Wen, F. B., . . . Zhao, S. (2015). Identification of potential plasma biomarkers for esophageal squamous cell carcinoma by a proteomic method. *Int J Clin Exp Pathol, 8*(2), 1535-1544.

Zhao, J., Qiu, W., Simeone, D. M., & Lubman, D. M. (2007). N-linked glycosylation profiling of pancreatic cancer serum using capillary liquid phase separation coupled with mass spectrometric analysis. *Journal of Proteome Research, 6*(3), 1126-1138. doi:10.1021/pr0604458

Zhao, J., Ye, W., Wu, J., Liu, L., Yang, L., Gao, L., . . . Li, Y. (2015). Sp1-CD147 positive feedback loop promotes the invasion ability of ovarian cancer. *Oncol Rep, 34*(1), 67-76. doi:10.3892/or.2015.3999

Zhao, Y., Yu, D., Li, H., Nie, P., Zhu, Y., Liu, S., . . . Fang, B. (2014). Cyclin D1 overexpression is associated with poor clinicopathological outcome and survival in oral squamous cell carcinoma in Asian populations: insights from a meta-analysis. *PLoS One, 9*(3), e93210. doi:10.1371/journal.pone.0093210

Zhao, Z., Ma, J., Mao, Y., Dong, L., Li, S., & Zhang, Y. (2018). Silence of alpha1-Antitrypsin Inhibits Migration and Proliferation of Triple Negative Breast Cancer Cells. *Med Sci Monit, 24*, 6851-6860. doi:10.12659/MSM.910665

Zhou, J., Trock, B., Tsangaris, T. N., Friedman, N. B., Shapiro, D., Brotzman, M., . . . Li, J. (2010). A unique proteolytic fragment of alpha1-antitrypsin is elevated in ductal fluid of breast cancer patient. *Breast Cancer Res Treat, 123*(1), 73-86. doi:10.1007/s10549-009-0625-5

Universiti Malaya

LIST OF PUBLICATIONS AND PAPERS PRESENTED

List of Publications:

1. YL Wong, R Anand, KM Yuen, WMW Mustafa, MT Abraham, KK Tay ZA Abdul Rahman, Y Chen. Comparative sera proteomics analysis of differentially expressed proteins in oral squamous cell carcinoma. (Submitted)
2. YL Wong, R Anand, KM Yuen, WMW Mustafa, MT Abraham, KK Tay, ZA Abdul Rahman, Y Chen. Identification of potential glycoprotein biomarkers in oral squamous cell carcinoma using sweet strategies. (Accepted)

List of Papers Presented:

1. YL Wong, R Anand, ZAA Rahman, Y Chen. Identification of serum proteomic signatures associated with oral squamous cell carcinoma. 3rd International Conference on Oral Microbiology & Oral Immunology. Kuala Lumpur, Malaysia (2020).
2. YL Wong, ZA Abdul Rahman, Y Chen. Glycoproteomics analysis of oral squamous cell carcinoma. The 4th RSU National and International Research Conference on Science and Technology, Social Science, and Humanities 2019 (RSUSSH 2019). Rangsit University, Pathum Thani, Thailand (2019).
3. YL Wong, ZA Abdul Rahman, Y Chen. Protein profiles of oral squamous cell carcinoma: a preliminary study. Dental Congregation 2016. The Royale Chulan Damansara, Petaling Jaya, Selangor, Malaysia (2016).
4. YL Wong. Identification of glycosylated proteins expression patterns in oral squamous cell carcinoma. Presented: Three Minutes Thesis 2016 Competition, University of Malaya, Kuala Lumpur, Malaysia (2016).

Universiti Malaya