

**THE MOLECULAR VALIDATION OF miRNA's AS SPECIFIC BIOMARKERS
FOR EARLY DIAGNOSIS OF OVARIAN CANCER**



**UNIVERSITY *of the*
WESTERN CAPE**

Sharneal Lottering

Thesis presented in fulfilment of the requirement for the degree of

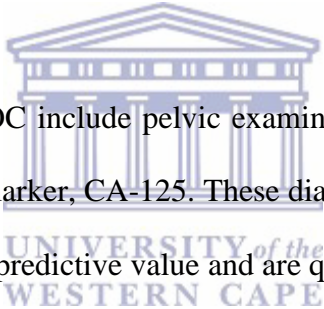
Magister Scientae

at the Faculty of Natural Science, Department of Biotechnology, University of the Western
Cape, South Africa

Supervisor: Dr Ashley Pretorius

Abstract

Ovarian Cancer (OC) is the most common reproductive and the most lethal gynaecological malignant tumour. The majority of Ovarian Cancers, comprising more than 95% of cases, emanate from the surface epithelium of the ovary, commonly referred to as Epithelial Ovarian Cancer (EOC). OC is the eighth most common form of cancer in women world-wide and in South Africa approximately 800 women die annually from the disease without diagnosis. OC is located deep within the pelvic region making early diagnosis and monitoring of the disease challenging. A minute group of cancer cells presents itself on the surface of one or both of the ovaries.



The current diagnostic tests for OC include pelvic examination, imaging studies, diagnostic imaging and a serum protein biomarker, CA-125. These diagnostic tools have low specificity, poor sensitivity, reduced positive predictive value and are quite invasive. Therefore, a method for early diagnosis is required that is less invasive and overcome the limitations regarding specificity, sensitivity and positive predictive value. Biomarkers are identified as feasible alternatives for early detection of Ovarian Cancer for example biological indicators such as DNA, RNA, proteins and microRNAs (miRNAs).

MicroRNAs are small non-coding RNAs that play a role in various cellular processes. Studies have shown that miRNAs are highly stable under extreme conditions, located in bodily fluids such as urine and blood, and they play a role in intracellular communication thus they can potentially be useful as diagnostic, prognostic and theranostic biomarkers. In this study we aimed to identify potential miRNA biomarkers to detect OC in its earliest stage and

evaluating these miRNAs in a panel of Ovarian Cancer cell lines as well as several other cell lines using qRT-PCR. The major objectives of the study were to identify miRNAs and their gene targets that play a significant role in the onset and advancement of OC using *in silico* methods.

An *in silico* approach was employed to identify miRNA target genes (also referred to as target genes) and miRNAs implicated in the initiation and development of Ovarian Cancer, in a previous study. A total of 7 miRNA target genes and 6 miRNAs were identified. Bioinformatics tools were used to functionally characterize the 7 target genes through protein-protein interactions, transcription factor analysis and pathway analysis. It was observed that the seven target genes were associated with various cancer-related processes. The prognostic/predictive value of the candidate target genes were monitored using three publically available databases namely Kaplan-Meier Plotter, SurvExpress and PROGene. Collectively, the data suggest that FARP1 and CILP are valuable candidate genes in Ovarian Cancer prognosis. The combined survival curves generated by the SurvExpress datasets showed the combined genes could predict the outcome of Ovarian Cancer patients pre- and post-treatment.

The identified miRNAs were experimentally validated using qRT-PCR to generate expression profiles for Ovarian Cancer as well as other cancers. Ovarian cell lines utilized in the study include OW28 (cystadenocarcinoma) and Coav-3 (adenocarcinoma). These cell lines were compared to a normal, KMST-6. In the study the expression profiles for six potential miRNA biomarkers for the detection of Ovarian Cancer was determined using qRT-PCR, and to distinguish OC from other cancers.

Expression profiling showed that the miRNAs identified through *in silico* methodologies were differentially expressed across all cell lines used in the study and the six miRNAs had a unique profile across all cell lines tested in the study. In addition miR1 was significantly up-regulated in the two ovarian cell lines, OAW28 and CaoV-3 compared to the normal cell line, KMST-6. The qRT-PCR analysis showed that the expression ratio in OAW28 was 8.91 and the expression ratio for CoaV-3 was 10.51. The results suggest that miR1 may be sufficient to differentiate between different cancer cell lines. The data also showed that miR2 and miR4 was under-expressed in CaoV-3 (expression ratios of -7.91 and -8.67 respectively) and HeLa but the expression levels of the two miRNAs were lower in the OC cell line. This is an important finding which could have potential prognostic implications in clinical practice by using it in combination with the current predictive biomarkers in the clinical setting. The study serves as a basis for future investigations for Ovarian Cancer diagnosis and prognosis, as well as for other cancers.



Keywords: Ovarian Cancer, miRNA, target genes, biomarker discovery, diagnostics, prognostics, functional genomics, prognostics, early detection, cancer, specificity, sensitivity, qRT-PCR, expression profiling, bioinformatics

Declaration

I declare that “**THE MOLECULAR VALIDATION OF miRNA’s AS SPECIFIC BIOMARKERS FOR EARLY DIAGNOSIS OF OVARIAN CANCER**” is my own work, that it has not been submitted for a degree or examination at any other university, and that all the resources I have used or quoted, and all the work which was the result of joint effort, have been indicated and acknowledged by complete references.

Lottering

Sharneal Lottering



03-07-2017

Date

Table of contents

ABSTRACT	i
DECLARATION	iv
TABLE OF CONTENTS	v
DEDICATION	xiii
ACKNOWLEDGEMENTS	xiv
LIST OF ABBREVIATIONS	xvi
LIST OF FIGURES	xx
LIST OF TABLES	xxiv
CHAPTER 1	4
Literature Review	4
1.1 Cancer Overview	4
1.1.1 Hallmarks of cancer	7
1.1.1.1 Self-sufficiency in growth signals	8
1.1.1.2 Insensitivity to antigrowth signals	9
1.1.1.3 Tissue invasion and metastasis	10
1.1.1.4 Limitless potential for replication	10
1.1.1.5 Sustained angiogenesis	11
1.1.1.6 Evading apoptosis	12
1.2 Ovarian Cancer: An overview	13



1.2.1	Histologic sub-types of Ovarian Cancer	15
1.2.1.1	Epithelial Ovarian Cancer (EOC)	15
1.2.1.2	Sex cord-stromal tumours	16
1.2.1.3	Germ cell tumours	17
1.2.2	Burden of disease in South Africa	17
1.2.3	Genomic alterations in Ovarian Cancer histological subtypes	18
1.2.4	Etiology of Ovarian Cancer	19
1.2.4.1	Age	20
1.2.4.2	Familial history and genetics	20
1.2.4.3	Reproductive factors and hormonal factors	20
1.2.4.4	Diet and life-style	21
1.2.4.5	Environmental factors	21
1.2.5	Staging of Ovarian Cancer	22
1.2.6	Signs and symptoms of Ovarian Cancer	23
1.2.7	Current screening and diagnostic tools for Ovarian Cancer	24
1.2.7.1	Pelvic examination	25
1.2.7.2	Transvaginal ultrasonography	26
1.2.7.3	Imaging studies	27
1.2.7.4	Image guided biopsy	28
1.3	Biomarkers	29
1.3.1.1	Serum protein biomarker: Human Epididymis Protein (HE4) and Cancer Antigen 125 (CA-125)	30
1.3.2	Classification of cancer biomarkers	32
1.3.3	Methods for Biomarker discovery in cancer	32



1.3.3.1 Genomic approach	33
1.3.3.2 Proteomic approach	34
1.3.3.3 Transcriptomics approach	35
1.3.3.4 Bioinformatics approach	35
1.4 MicroRNA's as biomarkers in cancer	36
1.4.1 Biogenesis and function of microRNA's	36
1.4.2 Circulating microRNAs	39
1.4.3 The role of microRNAs in the process of Epithelial-to-Mesenchymal Transition (EMT) in Ovarian Cancer	40
1.4.4 The specific role of microRNAs in Ovarian Cancer	41
1.5 Problem identification	43
1.6 Aims	44
1.7 Reference list	45
CHAPTER 2	57
Characterization of putative microRNA target genes	57
2.1 Introduction	57
2.1.1 Functional genomics	58
2.1.2 Previous study	60
2.1.3 Aim	61



2.1.4 Objectives	62
2.2 Methodology	63
2.2.1 Gene Ontology (GO) clustering	64
2.2.2 Protein-protein interaction analysis using STRING and GeneMANIA	64
2.2.2.1 STRING database	65
2.2.2.2 GeneMANIA	66
2.2.3 Transcription factor binding site analysis	66
2.2.3.1 GeneCard	66
2.2.3.2 TFactS®	67
2.2.4 Pathway analysis using Reactome	67
2.3 Results	69
2.3.1 Gene Ontology (GO) clustering using Gene Ontology Consortium	69
2.3.2 Co-expression analysis using STRING and GeneMANIA	71
2.3.2.1 STRING	71
2.3.2.2 GeneMANIA	71
2.3.3 Transcription factor analysis	74
2.3.4 Pathway analysis	76
2.4 Discussion	78
2.4.1 Gene Ontology	78



2.4.2 Protein-protein network generation	80
2.4.2.1 STRING	81
2.4.2.2 GeneMANIA	83
2.4.3 Transcription factor analysis	85
2.4.4 Pathway analysis	87
2.5 Conclusion	91
2.6 Reference list	92
CHAPTER 3	103
Prognostic studies of candidate microRNA target genes	103
3.1 Introduction	103
3.1.1 Prognostic markers/factors in Ovarian Cancer	104
3.1.2 Kaplan-Meier plot	108
3.2 Aims and objectives	108
3.2.1 Aims	108
3.2.2 Objectives	108
3.3 Methodology	109
3.3.1 Kaplan-Meier Plotter database	109
3.3.2 SurvExpress database	109



3.3.3 PROGgene database	110
3.4 Results	112
3.4.1 Kaplan-Meier Plotter database	112
3.4.2 SurvExpress database	114
3.4.3 PROGgene database	119
3.5 Discussion	124
3.5.1 Kaplan-Meier plotter database	124
3.5.2 SurvExpress database	125
3.5.3 PROGgene database	129
3.6 Conclusion	130
3.7 Reference list	131
CHAPTER 4	138
Molecular validation of identified miRNAs as biomarkers for early diagnosis of Ovarian Cancer (OC)	138
4.1 Introduction	138
4.1.1 Quantitative real-time PCR (qPCR)	141
4.1.2 Quantitative Strategies in qRT-PCR	144
4.2 Molecular methodologies	146



4.2.1 Cell culture	147
4.2.1.1 Thawing of cells	149
4.2.1.2 Trypsinization	149
4.2.1.3 Freezing down of cells	150
4.2.2 Total RNA extraction	150
4.2.3 Primer design	151
4.2.4 Poly (A) tailing and Reverse transcription	152
4.2.5 The validation of miRNA expression levels in Ovarian Cancer and other cell lines using qRT-PCR	154
4.3 Results	156
4.3.1 Analysis of amplification curves	156
4.3.2 Melting curve analysis	159
4.3.3 Analysis of qRT-PCR expression profiling data	162
4.4 Discussion	169
4.4.1 Analysis of Amplification curves and Melting curves	170
4.4.2 Expression profiling plot analysis	171
4.5 Conclusion	177
4.6 Reference List	179



CHAPTER 5	189
General discussion and future directions	189
5.1 General discussion	189
5.2 Future work	191
5.3 Reference List	193



DEDICATION

Dedicated to my Mom and Dad

Henriette Sharmain Lottering



Izak Lottering

ACKNOWLEDGEMENTS

I've come to believe that each of us has a personal calling that's unique as a fingerprint - and that the best way to succeed is to discover something you love and then find a way to offer it to others in the form of a service, working hard, and also allowing the energy of the universe to lead you.

- Oprah Winfrey



I would like to thank the Almighty Father, for being the source of my strength throughout the past two years. A special thanks to my supervisor Dr. A. Pretorius for his support and guidance during my Masters degree. It has been a privilege working under his supervision throughout this journey. I especially thank you for believing in me and my capabilities, and more often than not, making me believe in myself. I would like to thank the NRF for funding my Masters degree. In addition, I would like to thank the Bioinformatics Research Group (BRG) for their contribution and for all the teaching and support they have shown me over the past two years. I am grateful to have worked with them. A special thanks to Chimpape Lombe for her constant support and motivation. I will forever cherish the hours you spent training me. To my parents (Henriette and Izak Lottering), I am eternally grateful to you for all the love, support, motivation, patience and sacrifice. You are truly my pillar of strength. You have pushed me to realise my dreams and follow it with all my heart. You have sacrificed so much for me to accomplish my dreams; I will

cherish everything you have done for me always and forever. From the bottom of my heart thank you to my brother, Alphonse, and the rest of my family, for their prayers, love, support, sacrifices and patience during my Masters degree and my entire academic career. Each and every one of you has contributed to shaping me into becoming the best version of me.

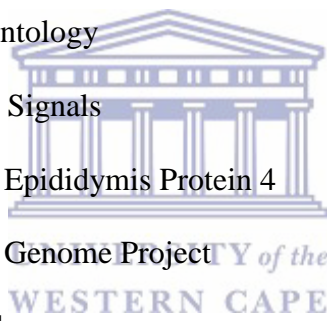
I Thank You!



LIST OF ABBREVIATIONS

%	Percent
3'-UTR	3' untranslated region
AKT	Protein Kinase B
AUROC	Area Under the Receiver Operating Characteristic Curve
BPs	Biological Processes
BRCA1	Breast Cancer type 1 Susceptibility Protein
BRCA2	Breast Cancer type 2 Susceptibility Protein
CA-125	Cancer Antigen 125
CaoV-3	Human Ovarian Cancer cell line
CC	Cellular Component
cDNA	Complementary Deoxyribonucleic Acid
CILP	Cartilage Intermediate Layer Protein
CO ₂	Carbon dioxide
CRTAP	Cartilage associated protein
C _T	Cycle threshold
CT	Computed tomography
DEU	Digital Expression Units
DMEM	Dulbecco's Modified Eagle Medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic Acid
EGFR	Epidermal Growth Factor Receptor
ELISA	Enzyme-linked immunosorbent assay
EMT	Epithelial-to-Mesenchymal Transition

EOC	Epithelial Ovarian Cancer
ERKs	Extracellular signal-regulated kinases
FAK	Focal adhesion kinases
FARP1	FERM, ARH/RhoGEF and Pleckstrin Domain Protein 1
FBS	Fetal Bovine Serum
FDA	Food and Drug Administration
FIGO	Federal Internationale de Gyneologie et di Obstetrique
FOX	Forkhead transcription factor
GCT	Germ cell tumours
GEO	Gene Expression Omnibus
GO	Gene Ontology
GS	Growth Signals
HE4	Human Epididymis Protein 4
HGP	Human Genome Project
HOX	Homeobox genes
IBS	Irritable Bowel Syndrome
KLK8	Kallikran 8
KM Plotter	Kaplan-Meier Plotter
MAP4K4	Mitogen-Activated Protein Kinase Kinase Kinase Kinase 4
MFs	Molecular Functions
miRNA	Micro Ribonucleic Acid
ml	Millilitres
mm	millimetres
mm ²	millimetres squared
MMP	Metalloproteinases



MRI	Magnetic Resonance Imaging
mRNA	Messenger ribonucleic Acid
NDEL1	NudE Neurodevelopment Protein 1 Like 1
NGS	Next Generation Sequencing
NPR3	Natriuretic Peptide Receptor 3
OC	Ovarian Cancer
OS	Overall survival
P2RX1	Purinergic Receptor R2X1
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
Penstrep	Penicillin-Streptomycin
PET	Positron emission tomography
PFS	Progression free survival
PI3K-AkT	Phosphidylinositol 3-kinase/Protein kinase B
PMS	Pre-menstrual syndrome
qRT-PCR	Quantitative Real- Time Polymerase Chain Reaction
RB	Retinoblastoma-associated Proteins
RFU	Rate of change of the fluorescence units
RISC	RNA-induced silencing complex
RNA	Ribonucleic Acid
s	seconds
SCST	Sex cord-stromal tumours
SNP	Single nucleotide polymorphism
STAB2	Stabilin 2
STRING	Search Tool for the Interacting Gene/Proteins

T	Time
TCGA	The Cancer Genome Atlas
TF	Transcription Factor
TGF β	Transcription Growth Factor beta
T _m	Melting temperature
TNF α	Tumour necrosis factor alpha
TNM	Tumour Node Metastasis
TP53	Tumour Protein P53
ug	Micro grams
ul	micro litres
UTI	Urinary tract infections
VEGF	Vascular endothelial growth factor
WHO	World Health Organisation



LIST OF FIGURES

<u>Chapter 1</u>	Page
Figure 1.1: The spread of cancer from origin	6
Figure 1.2: Acquired capabilities of cancer	7
Figure 1.3: Normal cell proliferation	9
Figure 1.4: The angiogenic switch during tumour development occurs at different stages in the tumour-progression pathway, depending on the nature of the tumour and the microenvironment	12
Figure 1.5: The female reproductive system with Ovarian Cancer in one ovary	14
Figure 1.6: Four stages of Ovarian Cancer	23
Figure 1.7: The most common symptoms of Ovarian Cancer	24
Figure 1.8: Pelvic examination by doctor	26
Figure 1.9: Transvaginal ultrasosnography	27
Figure 1.10: Ultrasound of the abdomen	29
Figure 1.11: Biopsy to detect Ovarian Cancer	28
Figure 1.12: Biogenesis of microRNAs	38

<u>Chapter 2</u>	Page
Figure 2.1: The integration of various molecular techniques of how DNA sequence is translated into complex information in a cell	59
Figure 2.2: The flow chart representing the <i>in silico</i> methodologies employed to functionally characterize the miRNA target genes during this study/chapter	63
Figure 2.3: Graphical representation of Gene Ontology (GO) clustering of the seven candidate miRNA targeted genes based on their Biological Processes (BPs) and Molecular Functions (MFs)	70
Figure 2.4: The seven candidate genes and their interacting genes produced by (STRING, 2017)	72
Figure 2.5: Co-expression analysis displaying the expression network of the putative genes and other genes linked to the network	73
Figure 2.6 Pathways identified associated with the target genes	77
<u>Chapter 3</u>	Page
Figure 3.1: Kaplan-Meier plot presenting the association between the expression of KLK8 and PFS and OS	107
Figure 3.2: Survival curves for the target genes CILP, CRTAP, FARP1, NPR3, P2RX1 and STAB2 using the Kaplan-Meier Plotter database	113

Figure 3.3:	The result summaries showing the significant target genes based on the p-values of the individual genes from the two datasets namely (a) SurvExpress dataset and (b) TCGA dataset using the SurvExpress database	115
Figure 3.4:	Kaplan-Meier analysis for the target genes for Ovarian Cancer outcomes	116
Figure 3.5:	Gene expression values across gene groups represented by box plots together with the p-value of the corresponding difference comparing risk groups	117
Figure 3.6:	Expression profile and gene ranking based on prognostic index	118
Figure 3.7:	Survival curves for the target genes CILP, CRTAP and FARP using the GSE9891 dataset comprising of 285 ovarian tumour samples	120
Figure 3.8:	Survival curves for the target genes NDEL, NPR3, P2RX1 and STAB2 using the GSE9891 dataset comprising of 285 ovarian tumour samples	121
Figure 3.9:	Survival curves for the target genes CILP, CRTAP and NDEL1 using the GSE14764 dataset	122
Figure 3.10:	Survival curves for the target genes FARP1, NPR3, P2RX1 and STAB2 using the GSE14764 dataset	123



<u>Chapter 4</u>	Page	
Figure 4.1	The graph illustrates the amplification curve produced following amplification using qPCR	143
Figure 4.2	Flow diagram representing the molecular validation of miRNA's identified through bioinformatics analysis	146
Figure 4.3	Amplification curve of Let7a in KMST-6 cell line	157
Figure 4.4	Amplification curve of miR1 a in KMST-6 cell line	158
Figure 4.5	Let7a in KMST-6 cell line	160
Figure 4.6	MiR1 in KMST-6 cell line	161
Figure 4.7	Relative expression ratio plot comparing the expression of the six candidate miRNAs in the HeLa cell line	163
Figure 4.8	Relative expression ratio plot comparing the expression of the six candidate miRNAs in the H157 cell line	164
Figure 4.9	Relative expression ratio plot comparing the expression of the six candidate miRNAs in the MCF7 cell line	165
Figure 4.10	Relative expression ratio plot comparing the expression of the six candidate miRNAs in the OAW28 cell line	166
Figure 4.11	Relative expression ratio plot comparing the expression of the six candidate miRNAs in the CoaV-3 cell line	167
Figure 4.12	Relative expression ratio plot of the six miRNAs in various cancer cell lines including Ovarian Cancer (OAW28 and CaOV3)	168

<u>Chapter 5</u>		Page
Figure 5.1	An example of a lateral flow device that can be developed for the detection of Ovarian Cancer using the discovered miRNAs	192

LIST OF TABLES

<u>Chapter 1</u>		Page
Table 1.1:	Increased and decreased risk factors associated with Ovarian Cancer	19
Table 1.2:	Epithelial-Mesenchymal-Transition related miRNAs and their target genes	41



<u>Chapter 2</u>		Page
Table 2.1:	MicroRNAs and their gene specific targets	61
Table 2.2:	Transcription Factors associated with the seven putative Ovarian Cancer genes	75

<u>Chapter 4</u>		Page
Table 4.1:	A list of cell lines used to investigate miR1-miR6	148
Table 4.2:	Poly (A) tailing and reverse transcription reagents used for cDNA synthesis	153



Chapter 1

Literature Review

1. Introduction

1.1 Cancer overview

1.1.1 Hallmarks of cancer

1.1.1.1 Self-sufficiency in growth signals

1.1.1.2 Insensitivity to antigrowth signals

1.1.1.3 Tissue invasion and metastasis

1.1.1.4 Limitless potential for replication

1.1.1.5 Sustained angiogenesis

1.1.1.6 Evading apoptosis



1.2 Ovarian Cancer: An overview

1.2.1 Histological sub-types of Ovarian Cancer

1.2.1.1 Epithelial Ovarian Cancer (EOC)

1.2.1.2 Sex cord-stromal tumours

1.2.1.3 Germ cell tumours

1.2.2 Genomic alterations in Ovarian Cancer histological sub-types

1.2.3 Etiology of Ovarian Cancer

1.2.3.1 Age

1.2.3.2 Familial history and genetics

1.2.3.4 Reproductive factors and hormonal factors

1.2.3.5 Diet and life-style

1.2.3.6 Environmental factors

1.2.4 Staging of Ovarian Cancer

1.2.5 Signs and symptoms of Ovarian Cancer

1.2.6 Statistics in South Africa

1.2.7 Current screening and diagnostic tools for Ovarian Cancer

1.2.7.1 Pelvic examination

1.2.7.2 Transvaginal ultrasonography

1.2.7.3 Imaging studies

1.2.7.4 Image guided biopsy

1.3 Biomarkers

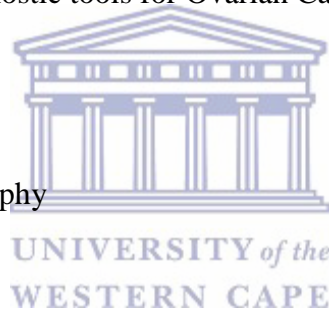
1.3.1 Serum protein biomarkers: Human Epididymis Protein 4 (HE4) and Cancer Antigen 125 (CA-125)

1.3.2 Methods for biomarker discovery in cancer

1.3.2.1 Genomic approach

1.3.2.2 Proteomic approach

1.3.2.3 Transcriptomics approach



1.3.2.4 Bioinformatics approach

1.3.3 Classification of biomarkers

1.4 MicroRNAs as biomarkers in cancer

1.4.1 Biogenesis and function of microRNAs

1.4.2 Circulating microRNAs

1.4.3 The role of microRNAs in the process of Epithelial-to-Mesenchymal Transition (EMT)
in Ovarian Cancer

1.4.4 The specific role of microRNAs in Ovarian Cancer

1.5 Problem identification

1.6 Aims

1.7 Reference List

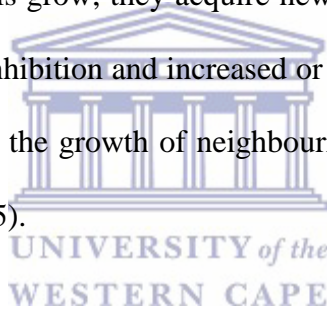


Chapter 1

Literature Review

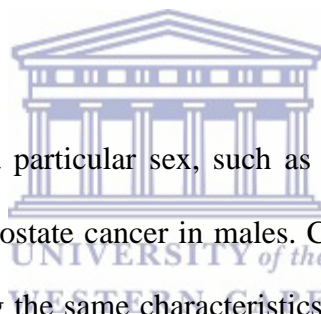
1.1 Cancer Overview

Cancer is generally described as a cellular growth disorder (neoplasia), affecting different organs and tissues in the body (Hejmadi, 2010). It is a group of diseases in which abnormal cells grow in an uncontrollable manner, disregarding the rules for normal cell division. In normal cells, specific signals determine when a cell should divide or undergo apoptosis (Hejmadi, 2010). Cancer cells are autonomous to these signals, resulting in uncontrolled proliferation. As these cancer cells grow, they acquire new characteristics including changes in cell structure, loss of contact inhibition and increased or decreased protein expression. The changes allow the cells to inhibit the growth of neighbouring cells and it allows the cells to invade other tissues (Movva, 2015).



Cancer is caused by genetic and epigenetic changes that fundamentally alter the properties of a normal cell (DeBerardinis *et al.*, 2008; Sadikovic *et al.*, 2008). One mutation will not render the cell cancerous but as the mutations accumulate, cancer can occur. Frequently, genes that make proteins to repair DNA damage are themselves mutated. Subsequently, mutations will increase in the cell, causing further abnormalities (DeBerardinis *et al.*, 2008). All the cells produced by division of the ancestral cell will display inappropriate proliferation. Cancer has the ability to spread through the lymphatic system and bloodstream, and grow into another tumour away from the site of origin. This is referred to as secondary cancer (metastasis). Once cancer has spread beyond the point of origin it becomes difficult to treat and the mortality rate of the disease increases (Evans and Roett, 2009).

Cancer is a disease that affects people of all ages, sex, races and socio-economic backgrounds. Globally, cancer is responsible for 13 % of all deaths. According to the World Health Organisation (WHO), there were approximately 12 million new cases of cancer and 7 million deaths in 2011 as a result of cancer (World Health Organisation, 2011). By 2012, the number of new cases rose to 14 million and the death toll rose to 8.2 million. Worldwide a 75 % increase is expected resulting in a total incidence of 22.2 million by 2030, with 17 million cancer-related deaths. It is suggested that South Africa could see an increase of 78 % in the number of cancer cases by 2030 (Economic Intelligent Unit Limited, 2009). The prevalence of cancer in developing countries is lower than developed countries, but this is quickly changing because of the westernised lifestyle being adopted (Stefan, 2015).



Several cancers are specific to a particular sex, such as ovarian cancer, breast cancer and cervical cancer in females and prostate cancer in males. Cancer comprises of approximately 200 different diseases, all sharing the same characteristics. Cancer is classified based on the type of cell, organ or tissue it originates from and are divided into six major categories including carcinomas, sarcomas, myelomas, leukaemias, lymphomas and mixed types (Dorak and Karpuzoglu, 2012). Carcinomas are the most common, accounting for 80-90 % of all cancers. The high prevalence of carcinomas is attributed to the fact that human epithelial cells are the site of cell proliferation and they are constantly exposed to carcinogenic factors. Sarcomas, myelomas, leukaemias and lymphomas originate from connective tissues, plasma cells of bone marrow, bone marrow and the lymphatic system respectively (Ceusters *et al.*, 2005). The majority of cancers occur as a result of multiple mutations that results in uncontrollable cellular growth. The cancer cells will continue to propagate and produce new cells until they crowd out all the normal cells. The lethality of cancer is due to the ability of cancer cells to spread to distant sites in the body. A primary tumour is relatively easy to

remove through surgery but once cancer has spread to many locations, surgery alone becomes impossible (Valastyan and Weinerg, 2011). For that reason, metastasis and invasion of normal tissue by cancer cells are the hallmarks of cancer as described by Hanahan and Weinerg in 2000. Tumours that are metastatic have the ability to move through the circulation system to other organs. As a consequence of this aggressive behaviour, secondary tumours far away from the primary tumour will develop that are more aggressive and difficult to eliminate. This process is described by the Invasion-Metastasis cascade illustrated in Figure 1.1 (Valastyan and Weinberg, 2011).

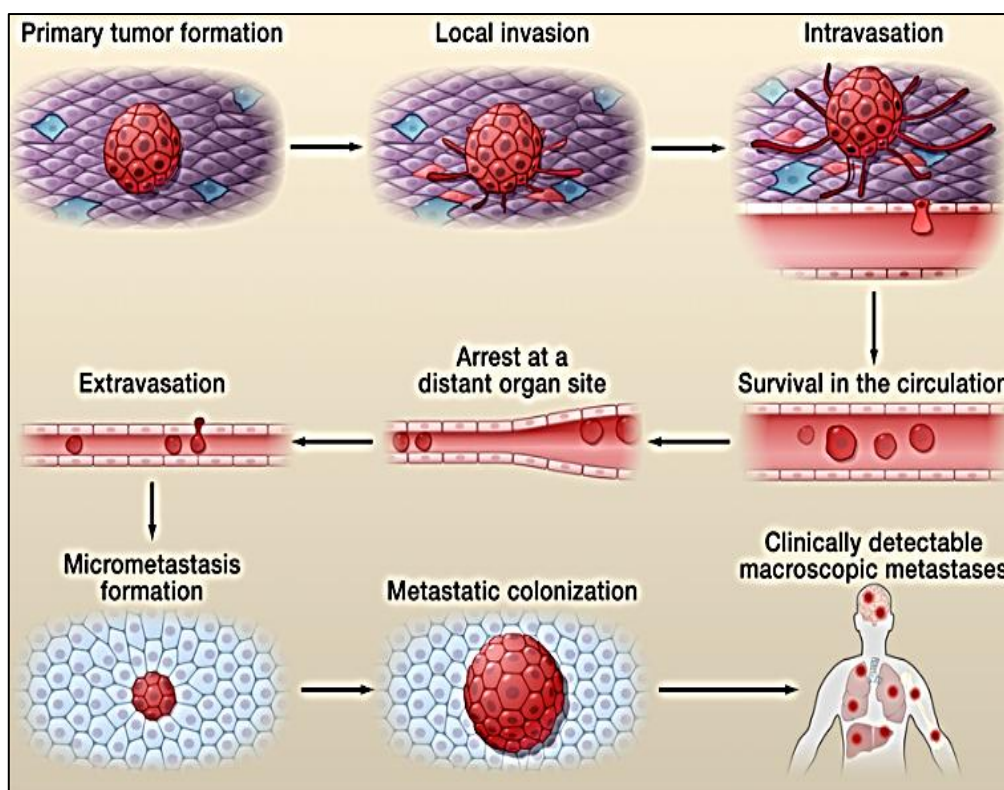


Figure 1.1: The spread of cancer from origin. The invasion-metastatic process consists of sequential steps. Cancer cells proliferate; break through basal membranes to migrate to capillaries and other suitable tissue for invasion and proliferation (Valastayn and Weinberg, 2011).

1.1.1 Hallmarks of cancer

There are six hallmarks of cancer including (i) self-sufficiency in growth signals, (ii) insensitivity to antigrowth signals, (iii) tissue invasion and metastasis, (iv) limitless potential for replication, (v) sustained angiogenesis and (vi) the ability to evade apoptosis shown in Figure 1.2.

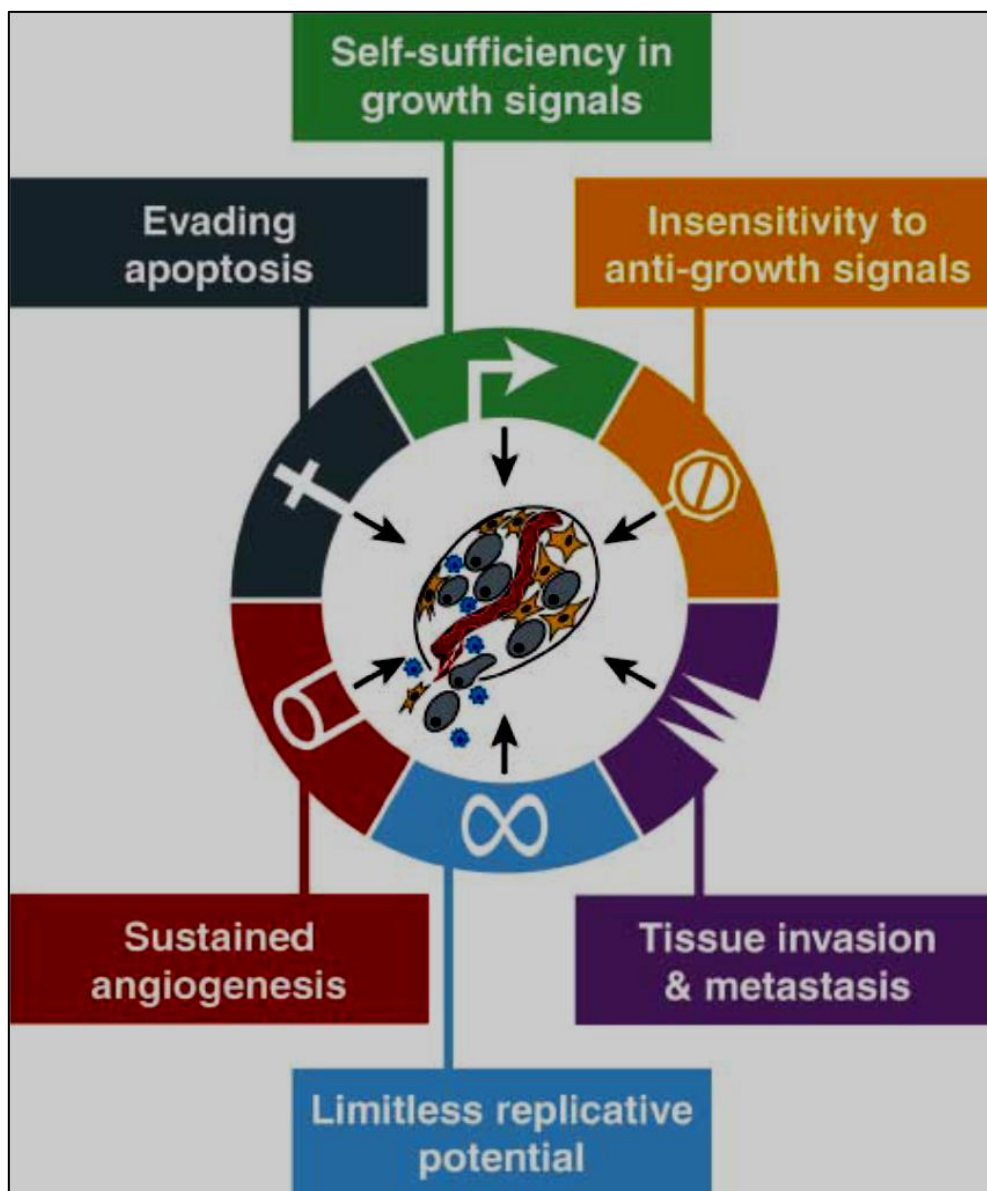


Figure 1.2: Acquired capabilities of cancer. The figure illustrates the six hallmarks of cancer originally proposed in 2000 (Hanhan and Weinberg, 2011).

1.1.1.1 Self-sufficiency in growth signals

According to Hanahan and Weinberg (2011), sustaining proliferative signals is arguably the most important trait in cancer cells. Normal cells require certain growth signals (GS) before proliferation can take place. Normal tissue cautiously control the production and transmission of these growth signals, by the transmembrane receptors. In order for cell proliferation to occur in a normal cell, consecutive steps, illustrated in Figure 1.3, is carefully followed. Based on current studies, no normal cell can evade these steps and proliferate in the absence of these growth stimulating signals (Witsch, Sela and Yarden, 2010). Cancer cells or tumour cells have the ability to generate their own growth signals thus reducing their dependence on stimulation from the normal tissue microenvironment. The generation of their own growth stimulating signals disrupts essential homeostatic mechanisms that ensures appropriate behaviour of various cells (Hanahan and Weinberg, 2000; Hanahan and Weinberg, 2011). Strategies for achieving growth signal autonomy include alteration of extracellular growth signals, alteration of transcellular transducers of those signals and alteration of intracellular circuits that translate those signals into action. Cell surface receptors that are responsible for the transduction of growth-stimulatory signals are also targets for deregulation during pathogenesis. Growth factor receptors are often over expressed in many cancers. Cancer cells favour the extracellular matrix receptors that transmit progrowth signals by altering the normal integrins and are thus liberated from dependence on exogeneous growth factor signals (Hanahan and Weinberg, 2000; Hanahan and Weinberg, 2011; Groten, Borner and Mertelsmann, 2016).

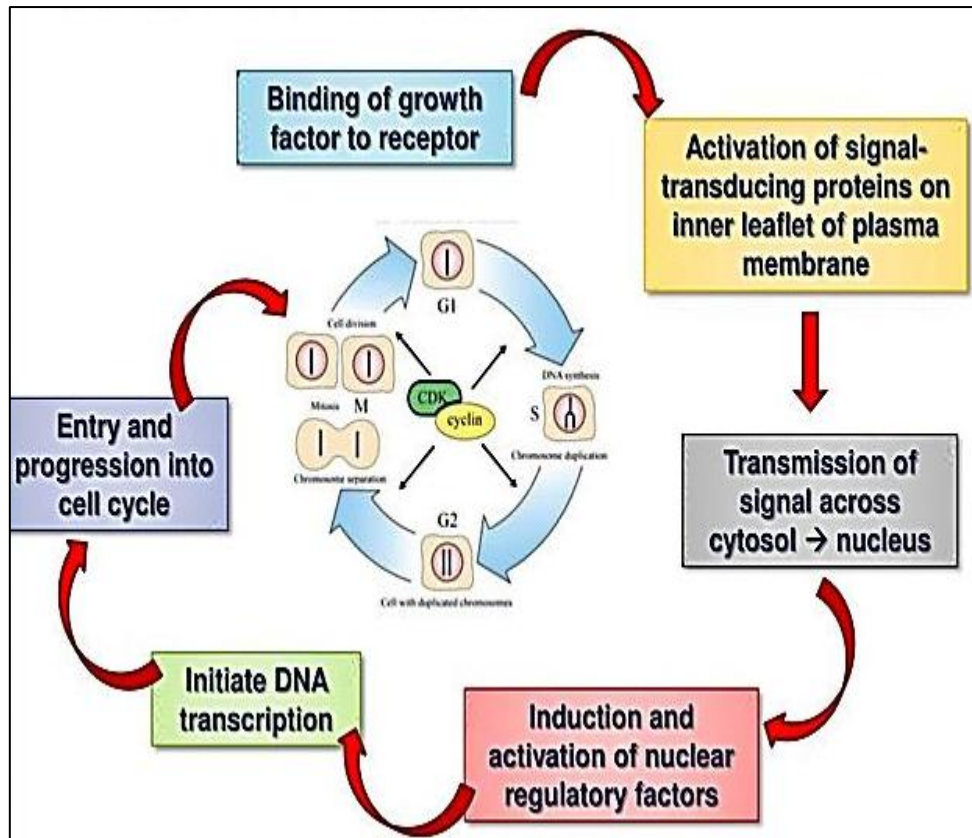


Figure 1.3: Normal cell proliferation (Hanahan and Weinberg, 2011).

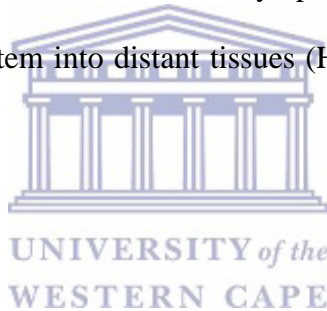
1.1.1.2 Insensitivity to antigrowth signals

Multiple antiproliferative signals such as soluble growth inhibitors and immobilized inhibitors maintain cellular quiescence and homeostasis in normal cells (Hanahan and Weinberg, 2011). These signals are controlled by tumour suppressors and are responsible for blocking cell proliferation. Cancer cells have the ability to bypass these antiproliferative signals in order to proliferate. The two quintessential tumour suppressors include TP53 (tumour protein P53) and RB (retinoblastoma-associated) proteins (Hanahan and Weinberg, 2011; Dai *et al.*, 2016). RB proteins govern whether or not a cell should proceed through the cell cycle; and TP53 halts proliferation or promotes apoptosis. Cancer cells lack these important factors of cell cycle control which might cause proliferation of the cells to cease

and enter the resting (G_0) phase of the cell cycle (Hanahan and Weinberg, 2000; Hanahan and Weinberg, 2011; Chow, 2010; Casimiro *et al.*, 2012).

1.1.1.3 Tissue invasion and metastasis

Cancer cells have the ability to migrate from the site of origin resulting in an invasive and metastatic phenotype. The invasion-metastasis cascade is a multistep process consisting of a distinct sequence of events illustrated in Figure 1.1. The cascade begins at a cellular level with reduced expression of surface adhesion molecules. This is followed by local invasion and intravasation of the cancer cells in the blood and lymphatic system. The cancer cells then move through the circulatory system into distant tissues (Hanahan and Weinberg, 2011; Dai *et al.*, 2016).



1.1.1.4 Limitless potential for replication

Cancer cells require the ability to replicate infinitely in order to produce solid mass tumours (Movva, 2015). Normal cells only pass through a limited number of successive cell growth and division cycles. Normal cells have a limited amount of division before they go into senescence or they experience cell death as a result of “crisis”. Evidence suggests that the telomeres, protecting the end of chromosomes, are centrally involved in the capability of unlimited replication (Shay and Wright, 2000). Telomeric DNA shortens with every successive cell division, until the telomeres become eroded and consequently lose their protective function. In cancer cells specialized DNA polymerase that adds telomere repeats to the ends of telomeric DNA is expressed at significant levels. Cancer cells bypass this barrier

by manipulating telomerase, specialized DNA polymerase, thus they can divide indefinitely, without initiating senescence (Kelland, 2007; Hanahan and Weinberg, 2011).

1.1.1.5 Sustained angiogenesis

Angiogenesis is a process whereby new blood vessels are formed, once tissues are formed. All cells in the tissue require a nearby blood supply, capillary blood vessel, oxygen and nutrients; and the ability to remove carbon dioxide (CO₂) and metabolic waste (Hanahan and Weinberg, 2011). Similarly, cancer cells require access to oxygen, nutrients and waste disposal. Cancer cells acquire the ability to initiate the production of new vasculature through the activation of the ‘angiogenic switch’ (Figure 1.4) (Hanahan and Weinberg, 2011, Hanahan and Weinberg, 2000). Cancer cells can counterbalance positive and negative signals that respectively encourage or block angiogenesis. The primary angiogenic growth factor is Vascular Endothelial Growth Factor (VEGF), VEGF promotes the survival, migration and replication of endothelial cells. Other angiogenic factors include endostatin, angiostatin and thrombospondin-1. The formation of new blood vessels is not only important in supporting tumour growth but also in metastasis. The newly formed capillaries are easily penetrated by tumour cells, providing the cancer cells the opportunity to enter the circulatory system and enter metastasis (Hanahan and Weinberg, 2011; Weis and Cheresh, 2011; Pezzella *et al.*, 2015).

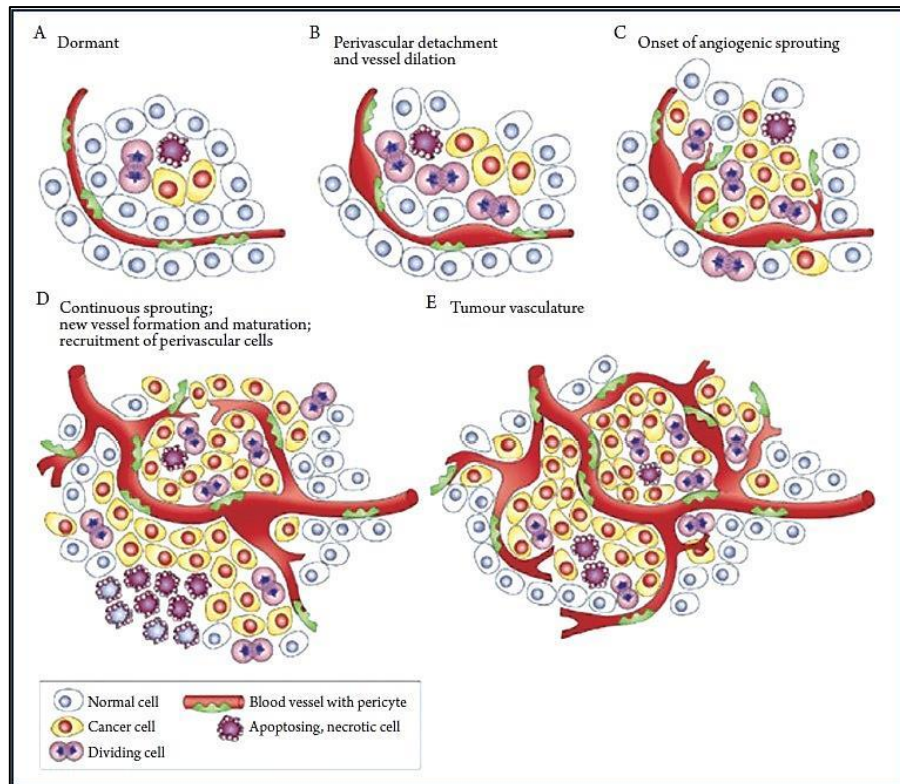


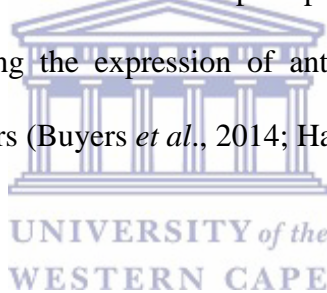
Figure 1.4: The angiogenic switch during tumour development occurs at different stages in the tumour-progression pathway, depending on the nature of the tumour and the microenvironment (Baker, Cherry-Bohannon and Francis, 2012).

1.1.1.6 Evading apoptosis

Apoptosis is the opposite of cell growth; it is a form of programmed cell death also known as cellular suicide (Orrenius and Zhivotovsky, 2010). In order for a cancer cell to divide and grow uncontrollably, a cancer cell has to hijack and bypass normal cellular growth pathways. Apoptosis is required for maintaining tissues in the body and it's initiated when cells are damaged or infected (Hanahan and Weinberg, 2011; Fernald and Kurokawa, 2013). Evidence, principally from mouse models, cultured cells and from descriptive analysis of biopsied stages in human carcinogenesis, shows that all types of cancer cells acquired various

means of resisting apoptosis (Fernald and Kurokawa, 2013; Orrenius and Zhivotovsky, 2010).

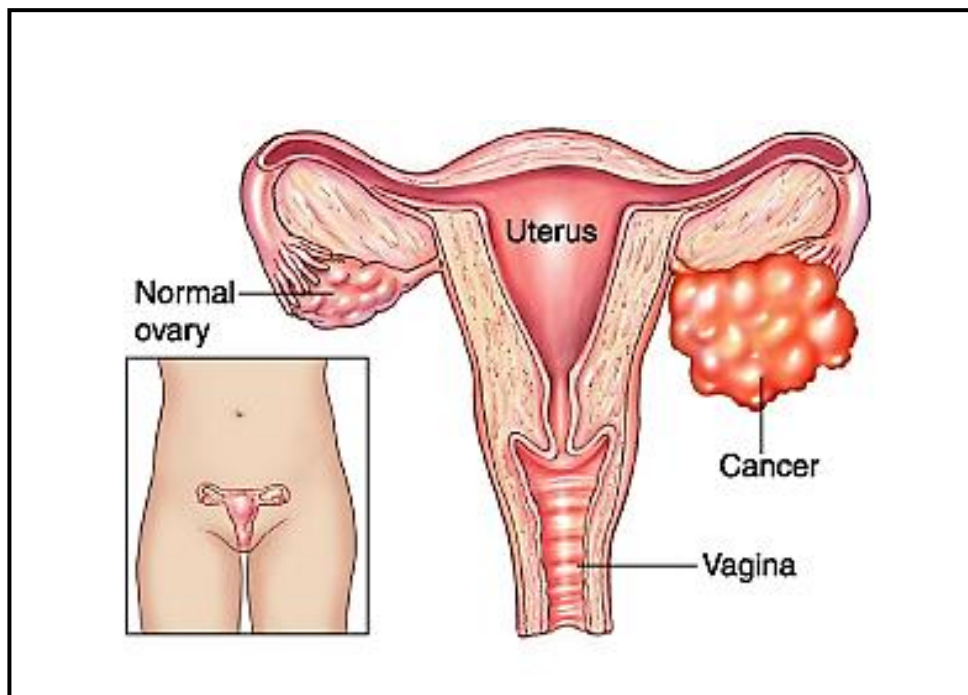
Apoptosis either occurs through intrinsic apoptotic or extrinsic apoptotic pathways. During the intrinsic apoptotic pathways, the signals from within the cell activate the process of apoptosis whereas during the extrinsic apoptotic pathway, the death signals from outside the cell are received and processed by the cell to activate apoptosis (Allocati *et al.*, 2012). It is suggested that intrinsic apoptotic pathways are more important in cancer prevention compared to extrinsic apoptotic pathways. The primary regulators of apoptosis belong to the Bcl-2 family of proteins. These proteins are either pro-apoptotic or anti-apoptotic. Tumours also evade apoptosis by elevating the expression of anti-apoptotic regulators and down-regulating the pro-apoptotic factors (Buyers *et al.*, 2014; Hanahan and Weinberg, 2011).



1.2 Ovarian Cancer: An overview

Women have two ovaries, on either side of the uterus in the female pelvic cavity. Each ovary is approximately the size of an almond, producing ova as well as the hormones, oestrogen and progesterone (MD Guidelines, 2010). It is the most important organ of the female reproductive system. The importance of these tiny glands is derived from their role in producing female sex hormones and female gametes that are fertilized for the production of embryos (Agarwal *et al.*, 2011). The development of the ovaries is part of the prenatal development of the reproductive system. The ovaries are held together by fibrous tissue that extends from the upper uterus to the lower region of the ovary (Agarwal *et al.*, 2011). Just

like any other organ, the ovary can become cancerous illustrated in Figure 1.5 (MD Guidelines, 2010).



UNIVERSITY of the
WESTERN CAPE

Figure 1.5: The female reproductive system with Ovarian Cancer in one ovary. The illustration shows cancerous tissue in the right ovary and normal ovary on the right (MD Guidelines, 2010).

Ovarian Cancer is a disease that affects women only. It transpires when cells in the ovary become abnormal and multiplies uncontrollably, resulting in the formation of a tumour. Ovarian Cancer is caused by insertions, deletions and genetic recombination in the DNA of a normal ovarian cell (Kurman and Shih, 2010). Approximately 10-15 % of Ovarian Cancers are linked to inherited genetic mutations, commonly referred to as germline mutations. But, most cases of OC are sporadic. These cases are associated with genetic mutations known as somatic mutations, which are acquired during an individuals' life-time. These sporadic

genetic mutations are only present in certain cells. In principle, these sporadic and germline mutations; result in the formation of anomalous growth in ovarian cells and tissues (Al Bakir and Gabra, 2014). There are different types of OC namely Epithelial Ovarian Cancer (EOC), borderline tumours, germ cell tumours and sex-cord stromal cell tumours. EOC is the most common, forming 90 % of all OC cases (Konishi, Koshiyama and Matsumra, 2014).

1.2.1 Histologic sub-types of Ovarian Cancer

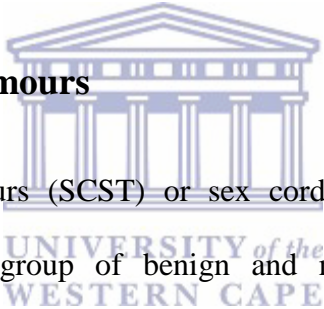
The difficulty in understanding Ovarian Cancer is due to the heterogeneity of the disease. It is composed of different types of tumours or subsets of the disease. All these tumours differ in clinicopathologic, behaviour, histogenetic principles with regard to their derivation from surface epithelium, germ cells and mesenchyme (the stroma and the sex-cord). Sub-classification of OC is biologically and therapeutically important. Ovarian Cancer is divided into two types of tumours namely Type I and Type II (Konishi, Koshiyama and Matsumra, 2014; Kurman and Shih, 2010). Type I ovarian tumours is clinically idle and it usually detected at an early stage. Upon diagnosis, these tumours are usually confined to the ovary and it includes low-grade serous, low-grade endometrioid, low-grade clear cell and low-grade mucinous carcinomas. Type II ovarian tumours are more aggressive, it's genetically unstable and it's usually diagnosed at an advanced stage (Kurman and Shih, 2010).

1.2.1.1 Epithelial Ovarian Cancer (EOC)

Epithelial cells cover the ovary and 80-95 % of Ovarian Cancers emanate from these cells. EOC is further divided into malignant and benign tumours. EOC is further differentiated under the microscope as follows: serous, mucinous, endometrioid, clear cell, transitional cell

tumours (Brenner tumours), carcinosarcoma, mixed epithelial tumour and undifferentiated tumours. Serous, mucinous, endometrioid and clear cell is the most common types of EOC (Konishi, Koshiyama and Matsumra, 2014). Clear cell and endometrioid carcinomas are highly associated with endometriosis. In stage distribution (stages where tumours has spread from primary site), serous carcinoma is found predominantly at an advanced stage. In contrast, clear cell and endometrioid carcinomas tend to remain confined to the ovary. Clear cell and endometrioid carcinomas may be unique histological types compared with serous carcinomas with respect to stage distribution and association with endometriosis (Al Bakir and Gabra, 2014).

1.2.1.2 Sex cord-stromal tumours



Ovarian sex cord-stromal tumours (SCST) or sex cord-gonadal stromal tumours are a morphologically heterogeneous group of benign and malignant tumours arising from granulosa, theca, sertoli and leydig cells. This histopathologic ovarian tumour group is infrequent, presenting approximately 7 % out of all primary ovarian tumours (Al Bakir and Gabra, 2014; Cunha and Horta, 2015). In contrast to EOC, malignant SCST are rare, comprising 1.2 % of all primary Ovarian Cancers. Although sex cord-stromal tumours are present in a broad age group, most of these tumours are present in younger patients. Furthermore, most patients are diagnosed in the primary stage and majority of these tumours are low-grade (Cunha and Horta, 2015).

1.2.1.3 Germ cell tumours

An Ovarian Germ Cell tumour (GCT) is a female specific tumour often referred to as ovarian teratoma. GCT account for 20-25 % of all ovarian tumours and constitute the second largest group of ovarian neoplasms; but it accounts for only 5 % of all malignant ovarian neoplasms. GCT derives from primitive germ cells of the embryonic gonad, principally affecting teenagers and young women. Approximately one-third of germ cell tumours are malignant. Germ cell tumours tend to affect only one ovary, and most are curable even if diagnosed during advanced stages (Isaacs, 2013).

1.2.2 Burden of disease in South Africa

Every year, approximately 8.2 million people die of cancer world-wide. Scientists suggest the number will double by 2050 which can be attributed to the changes in life-style. In most countries cancer falls in the top three causes of death, but according to South African statistics cancer don't even fall in the top ten. The westernized life-style being adopted by many South Africans is rapidly changing these statistics (Stefan, 2015). The statistics in South Africa is not currently up to date even though recording all cases of cancer became obligatory in 2011. The incompleteness is caused by a lack of funding and many people are reluctant to disclose information because it could be regarded as a breach of privacy. Also, various other diseases such as HIV/AIDS and TB occur more frequently in South Africa (Stefan, 2015). In South Africa, the latest statistics according to the National Cancer Registry was done in 2009. The National Cancer Registry estimated that 425 women were diagnosed with Ovarian Cancer. The highest percentage of individuals diagnosed between 50-60 years of age. The lowest percentage of women diagnosed was between 0-20 and above 80 years of age. This is evident that OC is a post-menopausal disease (National Cancer Registry, 2009).

The Burden of Disease Estimates for Comparative Risk Factor Assessment estimated that OC was the 15th cause of cancer-related death in 2000, in both men and women. Ovarian Cancer is ranked 9th when comparing all cancer-related deaths in women (Bradshaw *et al.*, 2006).


1.2.3 Genomic alterations in Ovarian Cancer histological subtypes

Molecularly the alterations in gene expression profiles in OC histological types associate with their counterparts in normal tissue. High Grade Serous ovarian tumours are generally categorized by loss of heterozygosity or mutations in p53. Alternatively, Low Grade Serous ovarian tumours are characterized by mutations in K-ras, BRAF and HER-2 genes. Mutations occurring in CTNNB1 (38-50 % of cases), PTEN (20 % of cases) and microsatellite instability (19 % of cases) are commonly seen in Endometrioid carcinomas (Croce and Di Leva, 2013; Birrer *et al.*, 2013). In mucinous carcinomas, K-ras is commonly mutated and it is associated to early events in tumourigenesis. Over-expression of HER-2 is linked to 15-20 % of mucinous tumours. Clear-cell carcinomas are characterized by a high frequency of mutations in PI3KCA. Mutation in the PI3KCA gene promotes the activation of the PI3K/AKT pathway, resulting in improved cell survival and invasion. Various histological subtypes have been associated with alterations in homeobox genes (HOX). The production of HOX genes are stimulated in the production of gynecological organs. Homeobox 9 (HOX9), homeobox 10 (HOX10) and homeobox 11 (HOX11) is highly expressed in serous, endometrioid and mucinous correspondingly (Croce and Di Leva, 2013; Dong, Lu and Lu, 2016).

1.2.4 Etiology of Ovarian Cancer

All women have some risk of getting Ovarian Cancer but some women have a higher risk than others. It is suggested that approximately 1 in every 72 women will be diagnosed with Ovarian Cancer in their lifetime (American Cancer Society, 2014). There are multiple factors that is associated with increased and decreased risk of getting Ovarian Cancer, illustrated in Table 1.1. Some of the risk factors or suggested causes of Ovarian Cancer include age, familial history and genetics, reproductive and hormonal factors, diet and life-style and environmental factors.

Table 1.1: Increased and decreased risk factors associated with Ovarian Cancer
(American Cancer Society, 2014).



Increased risk factors	Decreased Risk factors
Delayed child bearing/Low parity	Breastfeeding for more than 18 months
Early menstruation/ Late menopause	Early menopause
Endometriosis	Multiparity pregnancy
Eostrogen replacement therapy	Hysterectomy/ Tubal ligation
Genetic predisposition	Late menstruation
Genetic syndromes	Oral contraceptives

1.2.4.1 Age

Ovarian Cancer can occur in women at any age but studies have shown that the risk increases with age. Approximately 2 in every 3 women diagnosed with Ovarian Cancer are between the ages 50-69 years old. The mean age of women diagnosed is 63 years of age (Correia *et al.*, 2002).

1.2.4.2 Familial history and genetics

Approximately 5-10 % of women diagnosed with OC have an inherited increased risk of developing the disease, therefore in many cases Ovarian Cancer runs in the family. This is indicated by an increased incidence of OC among women with a familial history and by the observation of some families where multiple family members have/had OC (Mutch and Prat, 2014). Estimates of 10-15 % of all women diagnosed with OC are due to genetic factors. An inherited genetic mutation in one or both of the breast cancer genes (BRCA1 and BRCA 2) is one of the major genetic traits that increase the risk of women developing OC. Women that inherit genetic mutations in one or both of these genes, risk increases by 25-54 % (Mutch and Prat, 2014). Other mutated genes that can be inherited include PMS1 Homolog 1, mismatch repair system component (PMS-1), PMS1 Homolog 2, Mismatch Repair System Component (PMS-2) and MutL Homolog 1 (MLH-1) (Correia *et al.*, 2002).

1.2.4.4 Reproductive factors and hormonal factors

The connotation between OC and reproductive and hormonal factors differ by histological types. The most consistently described OC risk factors related to reproduction and hormones

include low parity and no oral contraceptives. There are multiple hypotheses postulated namely the ‘incessant ovulation theory’, the ‘pituitary gonadotropin hypothesis’, the ‘inflammation hypothesis’ and ‘the ovarian stromal hypothesis’ (Berchuck *et al.*, 2008). The ‘incessant ovulation theory’ states that the risk of OC is increased through the number of times a women ovulates in her life-time. Ovulation increases the rate of cellular division associated with repair of the epithelial cells therefore increasing the risk of spontaneous mutations. The ‘pituitary gonadotropin hypothesis’ postulates that an increased level in progesterone and oestrogen is linked to an increased proliferation and malignant transformation of ovarian cells. The ‘inflammation hypothesis’ protists that inflammatory conditions such as endometriosis stimulates OC formation. The ‘ovarian stromal hypothesis’ suggests that following ovulation there may be failure of apoptosis thereby stimulating the formation of OC (Berchuck *et al.*, 2008).



1.2.4.5 Diet and life-style

Diet and life-style are also major risk factors in the development of Ovarian Cancer. These factors are modifiable unlike inherited factors. Recent studies shows that smoking, obesity and an unhealthy diet increase the risk of women getting the disease (Sidaway, 2015).

1.2.4.6 Environmental factor

Exposure to talcum powder may increase the risk of OC but the evidence is controversial. A study done by Harvard Medical School showed that constant exposure to talcum powder doubled the risk of an individual getting OC, but other studies showed no relation between OC and talcum powder (Huncharek and Muscat, 2008; Sidaway, 2015). Talcum powder

contains particles of asbestos, a known carcinogen of OC, therefore researches believe it may be carcinogenic. Women that are exposed to asbestos on a daily bases have an increased chance of getting the disease than normal (Sidaway, 2015).

1.2.5 Staging of Ovarian Cancer

Ovarian Cancer is staged according to the extent the disease has spread from its origin. Doctors assign the stage of Ovarian Cancer to the Tumour Node Metastasis (TNM) or the Federation Internationale de Gynecologie et d'Obstetrique (FIGO) classification system. OC is divided into four stages illustrated in Figure 1.6 (Nordqvist, 2015). These systems are used to determine how far the cancer has spread; from the primary site and it's extremely important because it aids in providing the patient with the appropriate treatment required. In stage I, the cancer cells can either be present in one or both of the ovaries or it can be present on the surface layers of one or both of the ovaries. In stage II, the cancer cells have spread to the lymph nodes and into the pelvic region. In stage III, the cancer cells have spread to the abdominal cavity. In stage IV, also referred to as the advanced stage, the cancer cells has metastazied. The cancer cells have spread to distant organs and tissues (KK Women's and Children's Hospital, 2014).

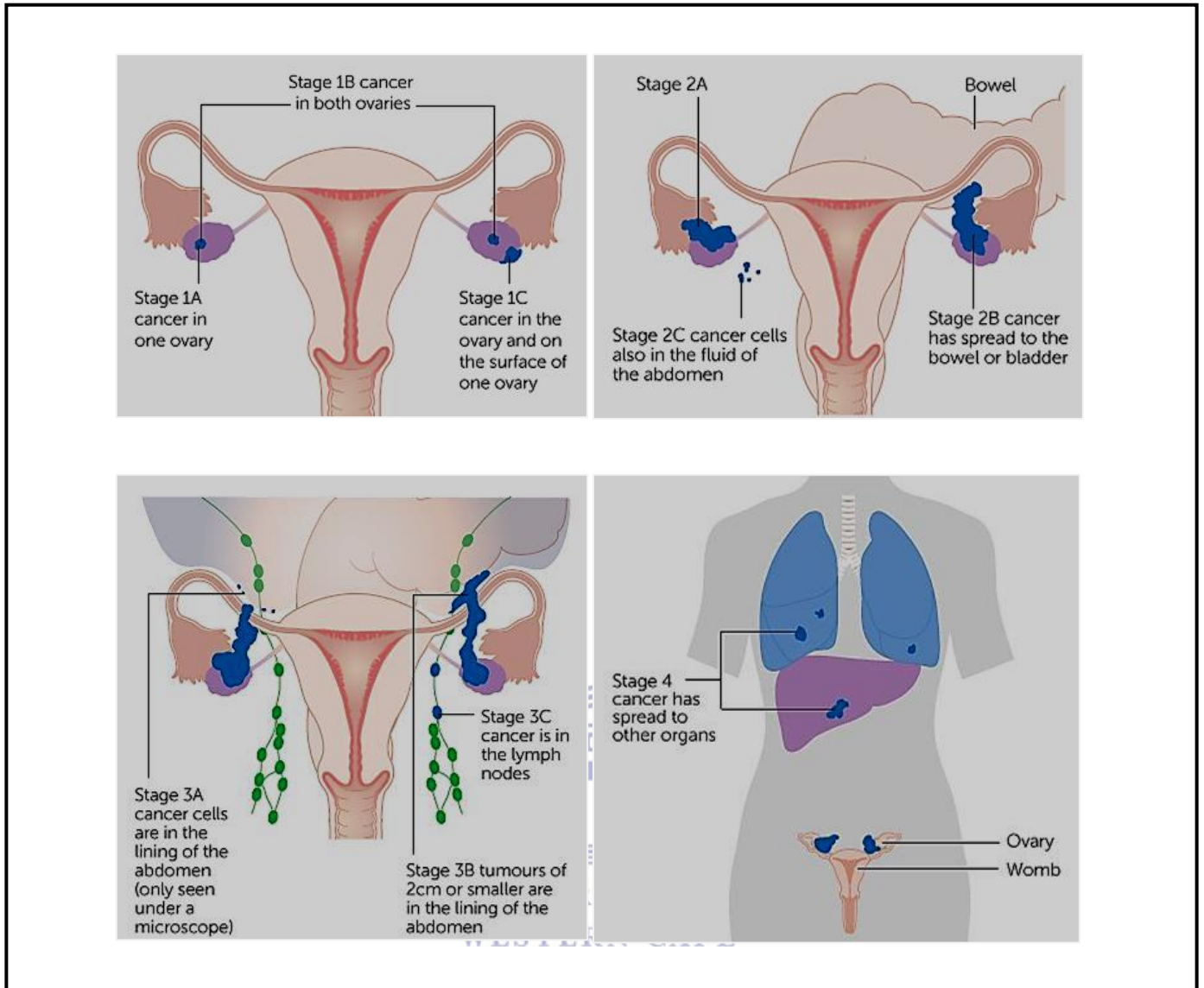


Figure 1.6: Four stages of Ovarian Cancer. From left to right. Stage I: Cancer cells located inside or on the surface of the ovary. Stage II: Cancer has spread to other tissues in the ovary. Stage III: Cancer has spread to tissues outside the pelvis and regional lymph nodes. Stage IV: Cancer has spread to tissues outside the abdomen and pelvis (Cancer Research UK, 2016).

1.2.6 Signs and symptoms of Ovarian Cancer

OC is challenging to diagnose because the symptoms are commonly experienced by women from time to time. It often resembles that of benign conditions such as irritable bowel

syndrome (IBS) or pre-menstrual syndrome (PMS). The most common symptoms include persistent abdominal bloating as the tumour creates pressure on the bladder and rectum; and fluid begins to form, early satiety (20 % of cases) and urinary frequency (Figure 1.9) (Cancer Research UK, 2016; Cho and Shih, 2010). Other symptoms include fatigue, upset stomach or heartburn (16 % of cases), back pain, pain during sex, menstrual changes, nausea (13 % of cases) and abnormal vaginal bleeding (Cho and Shih, 2010).



Figure 1.7: The most common symptoms of Ovarian Cancer (Cancer Research UK, 2016).



1.2.7 Current screening and diagnostic tools for Ovarian Cancer

The current diagnostic tools for Ovarian Cancer have very little success in early detection. There isn't one specific test that gives a definitive result as to whether or not an individual has OC. Diagnostic accuracy is important if it's being used for further tests and treatment of a disease. The two most important measures of diagnostic accuracy include specificity and sensitivity. Test sensitivity is defined as the ability to positively diagnose an individual with the disease, whereas test specificity is defined as the ability of a test to accurately identify those patient's without the disease (Simundic, 2009). An efficient diagnostic tool requires a

specificity of at least 99.6 %, a sensitivity of 75 % and a positive predictive value of at least 10 %. Ideally, it should be non-invasive and cost effective. The current diagnostic tools for Ovarian Cancer include pelvic examination, transvaginal ultrasonography, imaging studies and serum protein biomarker, CA-125; none of which is performed on its own (Cardenas-Goicoechea *et al.*, 2013).

1.2.7.1 Pelvic examination

A pelvic examination can be useful for the detection of some of the reproductive cancers but in most cases, OC is not detected during a routine pelvic exam because ovarian tumours are difficult or sometimes impossible to feel, even for a skilled practitioner. A pelvic exam is a physical exam done by a doctor to check for abnormalities such as enlarged ovaries. The doctor will exert pressure on a women's abdomen and insert his/her fingers into the vagina to feel the ovaries illustrated in Figure 1.8. If any masses are detected a needle will be passed through the skin to extract a fluid sample. This process is called paracentesis (Gajjar *et al.*, 2012). The major disadvantage of this diagnostic tool is that it's invasive. Many women feel uncomfortable during this procedure. Additionally, the sensitivity and specificity is below the requirement for an effective test for early diagnosis. It has been shown to be 40 % sensitive and 90 % specific for this procedure (Cardenas-Goicoechea *et al.*, 2013).

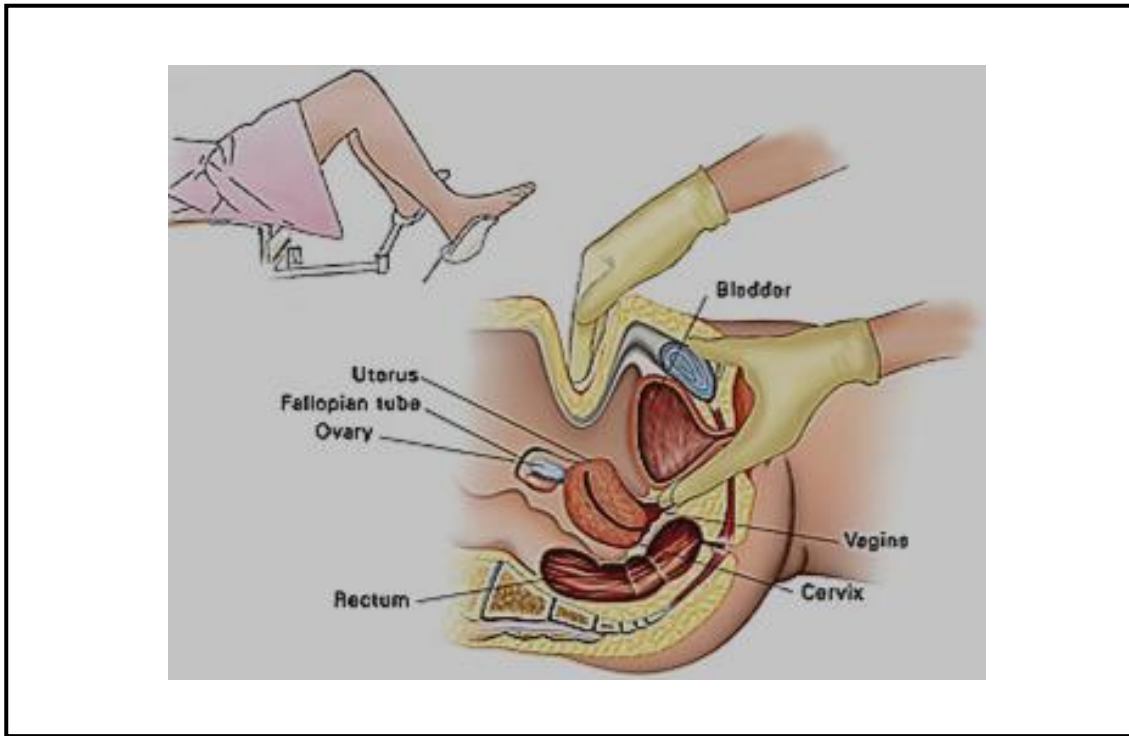
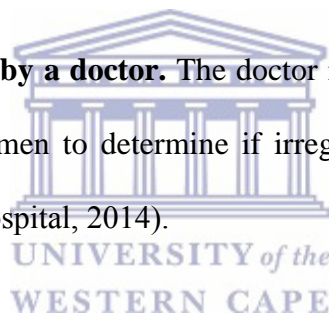


Figure 1.8: Pelvic examination by a doctor. The doctor inserts two fingers into the vagina and exerts pressure on the abdomen to determine if irregularities exist within the ovaries. (KK Women's and Children's Hospital, 2014).



1.2.7.2 Transvaginal ultrasonography

A transvaginal ultrasound uses sound waves to examine the uterus, fallopian tubes, and ovaries (Hoff and van Nagell Jr., 2013). Different sound waves are emitted by healthy tissue, fluid-filled cyst and solid mass tumours. During this test, a long thin transvaginal transducer, covered with a latex sheath and lubricant will be inserted into the vagina illustrated in Figure 1.9. The transducer will be angled to bring the areas of study into focus. Images of organs and structures will be displayed on a computer screen called a sonogram (Hoff and van Nagell Jr., 2013). The disadvantages of this diagnostic tool include low specificity and sensitivity for a conclusive diagnosis. Most women find the procedure very uncomfortable but not painful.

Also, it is suggested that the procedure should be done annually as a screening mechanism but it's not cost effective. The procedure is between R4000-R17 600, which is not feasible for most South Africans (Lockwood and Ritzert, 2013).

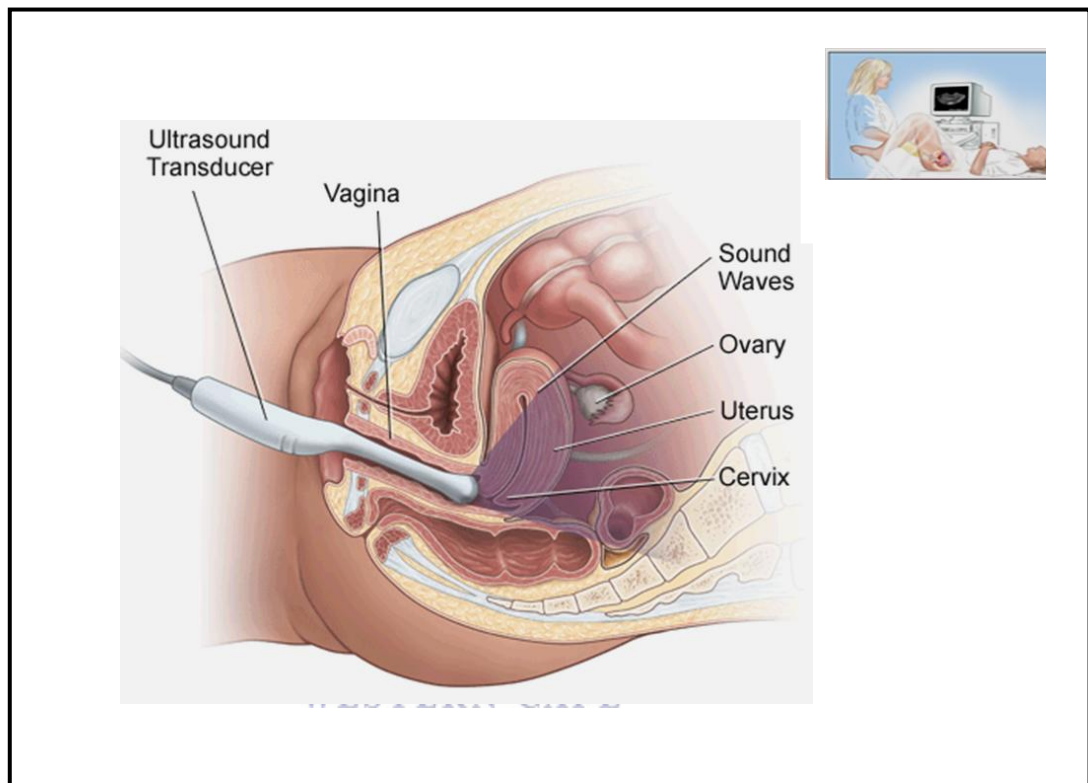


Figure 1.9: Transvaginal ultrasosnography The medical technician inserts a transducer into a women's vagina. The transducer emits sound waves generating an image of the pelvic organs (as shown in the top right hand corner) (KK Women's and Children's Hospital, 2014).

1.2.7.3 Imaging studies

Imaging studies include ultrasound (Figure 1.10), computed tomography (CT) scans, magnetic resonance imaging (MRI) scans and a positron emission tomography (PET) scan. These tests are conducted to determine whether or not a pelvic mass is present and if the

cancer has spread to distant tissues and organs. The disadvantage of these tests is that it will not confirm that the solid mass is a tumour. For definitive diagnosis, further studies are required (Ohmichi and Tanaka., 2012).

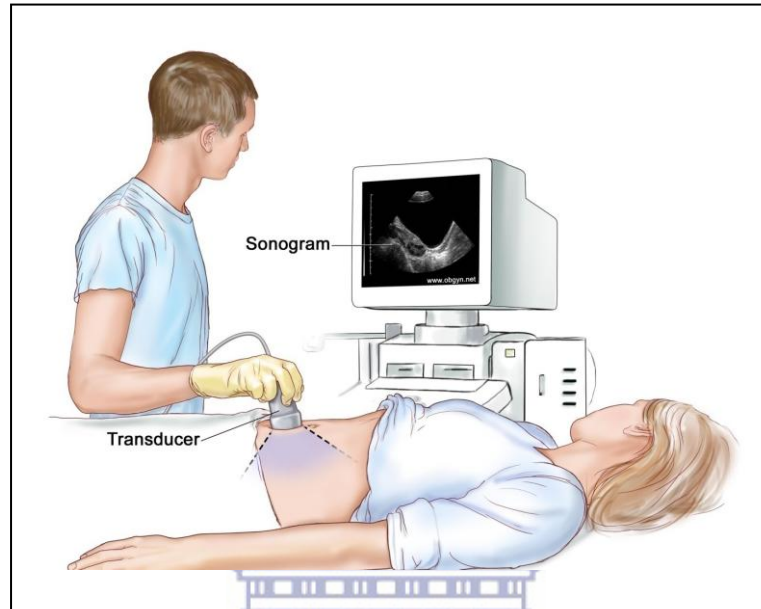


Figure 1.10: Ultrasound of the abdomen A small probe is placed on the surface of a women's abdomen which releases sound waves. This translates into a picture (National Cancer Institute, 2016).

1.2.7.4 Image guided Biopsy

A biopsy (Figure 1.11) is commonly done to remove the tumour in the ovary. In rare cases, a patient suspected of OC may be biopsied by taking tissue samples from the ovaries or from a sheet of fatty tissue inside the abdomen since OC may spread there (Cancer Research UK, 2016). During this procedure an ultrasound or a CT scan is used to guide the biopsy procedure. The procedure takes approximately 10-20 minutes under local anaesthetic. Multiple samples may be taken from different locations thus there may be several needle

punctures. The tissue samples extracted will be examined in the laboratory to check for OC. Following the procedure, the patient pulse and blood pressure will be monitored whilst resting in bed. The patient may be hospitalized overnight for further monitoring (National Cancer Institute, 2016; Ohmichi and Tanaka., 2012). As a result of the limitations of the current diagnostics tools biomarkers were investigated.

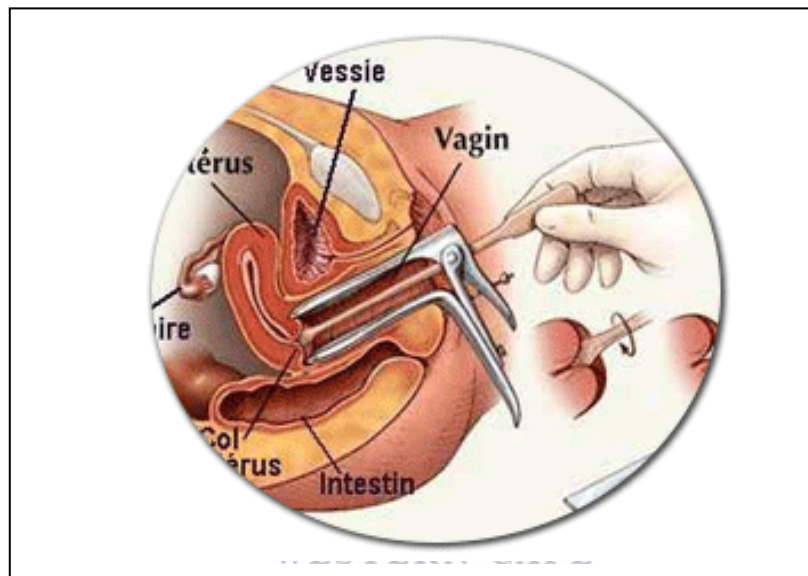
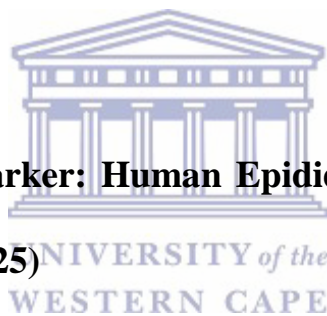


Figure 1.11: Biopsy to detect Ovarian Cancer. The doctor inserts a special needle to extract cells and tissues from the ovary to be viewed under a microscope by a pathologist to detect for signs of cancer (National Cancer Institute, 2016).

1.3 Biomarkers

According to the National Cancer Institute (2016), a biomarker is a biological marker found in blood, other bodily fluids, or tissues that is a sign of a normal or abnormal process, or a condition or disease, such as cancer. It is used as an index of the intensity of a disease or

other physiological states in an organism. A biomarker has an important role in medical research and practice providing insight into the mechanism and course of a disease (Hayes and Henry, 2012). Biomarkers are used for various purposes such as diagnostics, screening and prognostics. Regardless of the function of the biomarkers it's clinical significance is dependent on their sensitivity, specificity, predictive value, precision, reliability, reproducibility and the possibility of easy use and wide spread application. It is important for all biomarkers, according to their purpose, to possess certain characteristics, meet certain requirements and answer certain questions such as 'is it an optimal drug for the disease', 'what is the optimal dosage' and 'what is the likelihood of developing the disease' (Hayes and Henry, 2012 and Singh, 2011).



1.3.1 Serum protein biomarker: Human Epididymis Protein 4 (HE4) and Cancer Antigen 125 (CA-125)

HE4, a member of the *wey acidic protein* gene family, is expressed in the reproductive tract and its over expressed in ovarian cancer cells, especially in serous and endometrioid ovarian carcinoma (Amant *et al.*, 2011). It has been suggested to be a serological marker of OC (Amant *et al.*, 2011). A study done by Moore *et al.*, (2008, 2009) evaluated nine potential biomarkers, of which HE4 was most effective in detecting OC. They observed a sensitivity of 72.9 % and a specificity of 95 %. The average level of serum HE4 in malignant lesions were 248.7 Pm whereas in normal control and benign lesions a lower concentration was observed (Moore *et al.*, 2009). Moore and associates (2008, 2009) went further by combining HE4 and CA-125 which yielded a higher sensitivity of 76.4 % and a specificity of 95 % (Area Under the Receiver Operating Characteristic curve (AUROC) of 0.91); suggesting that using these biomarkers in combination predicts OC more accurately than either alone. Regardless of

these positive results, HE4 as a diagnostic biomarker is still controversial. Conversely, a study done by Kristjansdottir *et al.* (2013) suggested a reduction in the sensitivity (48.3 %) and specificity (75 %) with a AUROC of 0.73 thus the diagnostic accuracy of HE4 is contradicting and require further assessment (Gupta, Rastogi and Sachan, 2016; Moore *et al.*, 2008; Moore *et al.*, 2009; Kristjansdottir *et al.*, 2013).

Cancer Antigen-125 (CA-125) is a glycoprotein, tumour biomarker commonly used for ovarian tumours. It is expressed by OC cells and it's expressed in tissue derived from Müllerian and coelomic epithelia. It circulates in the blood therefore it can be measured with a blood test (Goy *et al.*, 2012). According to Bray-Ward *et al* (2005), a single protein cannot be used to distinguish an individual that has cancer from a healthy control. There are many challenges related to this biomarker but so far no other serum biomarker has outperformed, CA-125. Some other benign conditions of the womb such as endometriosis, pregnancy and pelvic inflammatory disease also produce high levels of CA-125 therefore the test gives false positive result (Agarwal *et al.*, 2011). The specificity, sensitivity and positive predictive value is lower than required for an effective diagnostic tool for early diagnosis. The sensitivity is between 50-60 %, specificity 90 % and positive predictive value is 7 %. The CA-125 biomarker has been used in combination with other protein biomarkers to improve its specificity and sensitivity (Goy *et al.*, 2012).

The OVA1 blood test, approved by the Food and Drug Administration (FDA), is designed to be a highly sensitive tool. It tests for all types of OC and determines if the tumour is low risk OC or high risk OC. The OVA1 blood test measures the level of CA-125 in the blood, in combination with four other proteins namely beta-2 microglobulin, transferrin, apolipoprotein

A1 and transthyretin (prealbumin). The OVA1 blood test has an approximate sensitivity of 96 %, specificity of 35 % and a positive predictive value of 40 %. Regardless of the increased specificity, sensitivity and positive predictive value, the test is still not performed on its own. It is used in addition to, not in place of, other diagnostic tools (Bristow *et al.*, 2014).

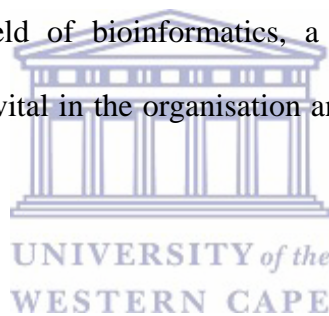
1.3.2 Classification of cancer biomarkers

The use of biomarkers were first described in 1980 by Isaakson. However, all biomarkers don't share all the same characteristics and are classified based on their characteristics and their application. There are multiple types of biomarkers, which fulfills various roles (Corella and Ordovas, 2015). Biomarkers are generally defined into 6 categories, namely (i) biomarkers of risk prediction; whether an individual is predisposed to developing any disorder, (ii) screening/detection biomarkers are real time indicators of the presence of a disease (iii) diagnostic or trait biomarkers; are measurable characteristics reflecting the presence of the disease state, (iv) state or acuity biomarkers reflect the severity of the disease and the probability of recurrence (v) predictive biomarkers predict a patients response to treatment, determine the optimum dosage or the efficiency of the drug and (vi) monitoring biomarkers predict and monitor a patients recurrence of the disease following treatment. Collectively, these biomarkers provide information about a patient at virtually every stage of the disease (Berk, 2015; Mishra and Verma, 2010).

1.3.3 Methods for Biomarker discovery in cancer

Biomarkers are the cornerstone of medical care. The discovery of novel biomarkers is imperative because it provides a dynamic and powerful approach to understanding the

prediction, cause, diagnosis, progression, regression or the outcome following treatment of a disease (Mayeux, 2004). In essence it aids in reflecting the entire continuum of a disease from it's earliest stage to the terminal stage (Mayeux, 2004). Biomarkers can be detected in all parts of the body including bodily fluids. These biomarkers include mRNA, DNA, protein, metabolites and miRNA biomarkers. The evolution of 'omics' technologies have been dedicated to identify large numbers of candidate biomarkers. This explosion in 'high-throughput' technologies led to an increased interest in the discovery and validation of molecular biomarkers (Diamandis and Kulasingam, 2008; Ghosh and Poisson, 2008). These methods of discovery include genomics, proteomics and transcriptomics; and prompted a large amount of data that led to the need for computerised databases to store, organise and analyse generated data. The field of bioinformatics, a discipline of computational and biological sciences, has become vital in the organisation and analyses of this vast amount of biological data (Lewis, 2008).



1.3.3.1 Genomic approach

Genomics is a discipline in genetics that concerns function and structure of an organisms genome. The Human Genome Project (HGP) has a major impact on molecular diagnostics and the discovery of biomarkers. The widely used genomic technologies include microarray technologies and single nucleotide polymorphism (SNP) array and next generation sequencing (NGS) technologies (Chen *et al.*, 2007). These technologies are used to discover and identify biomarkers related to genome alterations caused by cancer, including chromosomal rearrangements, epigenetic modification and copy number aberrations. Several of these technologies produce a massive amount of data per experiment. As a result elaborate computational tools are required for the analysis of these huge datasets. Nevertheless, there

has been tremendous growth in genomic applications (Chen *et al.*, 2007; Pollack, 2007; Quackenbush, 2006). This growth has contributed to the discovery of a highly promising diagnostic biomarkers for example cell migration-inducing and hyaluronan-binding protein (CEMIP) gene was identified through oligonucleotide microarray analysis as a biomarker for colorectal cancer (Sung and Wu, 2013).

1.3.3.2 Proteomic approach

For decades proteins have been an integral part of the field of clinical chemistry. Recent advances in the technology and the completion of the HGP gave rise to new opportunities for analysing proteins for clinical diagnostic purposes. Proteomics aims to interrogate extremely complex protein mixtures in blood and tissues. Blood contains approximately 100 000 different proteins. Several proteomics approaches have been use to identify novel biomarkers for example mesothelin (Diamandis and Kulasingam, 2008).

Mesothelin is a cell surface protein overexpressed in OC. A study done by Huang *et al* (2009) used preoperative mesothelin levels from patients with benign ovarian tumours and patients with ovarian carcinomas. A direct Enzyme-linked immunosorbent assay (ELISA) was used to determine the measurement of mesothelin in the serum. Higher levels of mesothelin were found in OC patients compared to benign ovarian tumours. Additionally, mesothelin was significantly increased from early to advanced stages and revealed a poorer overall survival for OC patients. Therefore, mesothelin was presented as a novel tumour marker for diagnosis as well as prognosis for Ovarian Cancer (Huang *et al.*, 2009; Tinalli *et al.*, 2007).

Regardless of the advances in the technologies, proteomics has limited ability to identify and quantify proteins in complex mixtures. Also, only 10 % of the proteins in human serum can be detected with currently available approaches. There is still potential for further discovery of novel biomarkers but only a selected few available applications are suitable for analysis of biomarker discovery (Diamandis and Kulasingam, 2008; Sallam, 2015; Srivastava *et al.*, 2005).

1.3.3.3 Transcriptomics approach

Transcriptomics is a technique used to determine the differential expression of RNA transcripts over-time and/or between cells and diseases (Chen *et al.*, 2007). It is one of the oldest and widely used high throughput technologies in biomarker discovery. The main aim of this technique is to describe and quantify RNAs; and to determine their variations in response to stimuli or disease. Expression profiling by microarray has been very successful in biomarker discovery however direct sequencing offers a greater potential for the detection of more transcripts and their variants (Chen *et al.*, 2007).

1.3.3.4 Bioinformatics approach

Bioinformatics is a discipline in computational and biological sciences, also referred to as computational biology. It is important in the sorting and analysis of the enormous amount of biological information generated by other ‘omics’ technologies (Ghosh and Poisson, 2013). The focal purpose of bioinformatics is to identify significant biological information within a pool of raw data eventually leading to the discovery of novel biomarkers. Various pipelines can be used to simulate complex disease physiologies and generate a list of valuable targets

such as miRNAs, genes or proteins that can be validated as potential biomarkers for diseases (Calvert-Joshua, 2013 and Ngcoza, 2013).

1.3.4 Classification of cancer biomarkers

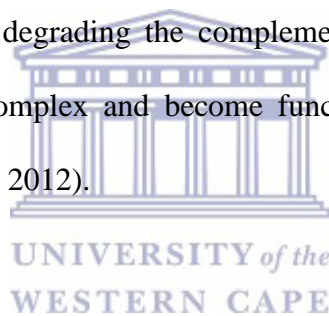
The use of biomarkers were first described in 1980 by Isaakson. However, all biomarkers don't share all the same characteristics and are classified based on their characteristics and their application. There are multiple types of biomarkers, which fulfills various roles (Corella and Ordovas, 2015). Biomarkers are generally defined into 6 categories, namely (i) biomarkers of risk prediction; whether an individual is predisposed to developing any disorder, (ii) screening/detection biomarkers are real time indicators of the presence of a disease (iii) diagnostic or trait biomarkers; are measurable characteristics reflecting the presence of the disease state, (iv) state or acuity biomarkers reflect the severity of the disease and the probability of recurrence (v) predictive biomarkers predict a patients response to treatment, determine the optimum dosage or the efficiency of the drug and (vi) monitoring biomarkers predict and monitor a patients recurrence of the disease following treatment. Collectively, these biomarkers provide information about a patient at virtually every stage of the disease (Berk, 2015; Mishra and Verma, 2010).

1.4 MicroRNAs as biomarkers in cancer

1.4.1 Biogenesis and function of microRNAs

MicroRNAs were first discovered in the nematode, *Caenorhabditis elegans* by Lee, Feinbaum and Ambros in 1993 (Chauhan *et al.*, 2012). They are a highly conserved class of

small (~ 22 nucleotides) non-coding RNA molecules. The mature microRNA molecule is produced through a series of steps (Chen *et al.*, 2014). Mature microRNAs are derived from stem-loop precursors, pri-miRNA, transcribed by RNA polymerase II. The pri-miRNA has a 3' poly-A-tail and a 7-methylguanosine cap at the 5' end. It is cleaved by Drosha and its co-factor producing pre-miRNA. The pre-miRNA is transported into the cytoplasm from the nucleus by the nuclear transport receptor, Exportin-5 and the nuclear protein, Ran-GTP. The exported pre-miRNA is further processed, producing a small duplex miRNA molecule, ~22 nucleotides, by the enzyme Dicer. The duplex miRNA further assembles into a ribonucleoprotein complex known as RNA-induced silencing complex (RISC). The RISC complex is responsible for the induction of unwinding the double-stranded molecule into a single-stranded miRNA thereby degrading the complementary strand. The single-stranded miRNA remains in the RISC complex and become functional. This process is shown in Figure 1.12 (Barca-Mayo and Lu, 2012).



The 'seed' sequence, located at the 5' end of the mature miRNA is complementary to the 3' untranslated region (3' UTR) of the target mRNA. The complementarity doesn't have to be perfect, therefore, a single miRNA can regulate multiple mRNA targets and multiple miRNAs can regulate one mRNA (Barca-Mayo and Lu, 2012). The level of complementarity of the miRNA controls the mode of inhibition. When the complementarity is perfect, the mRNA is degraded whereas if the complementarity is imperfect, translational suppression occurs (Chauhan *et al.*, 2012).

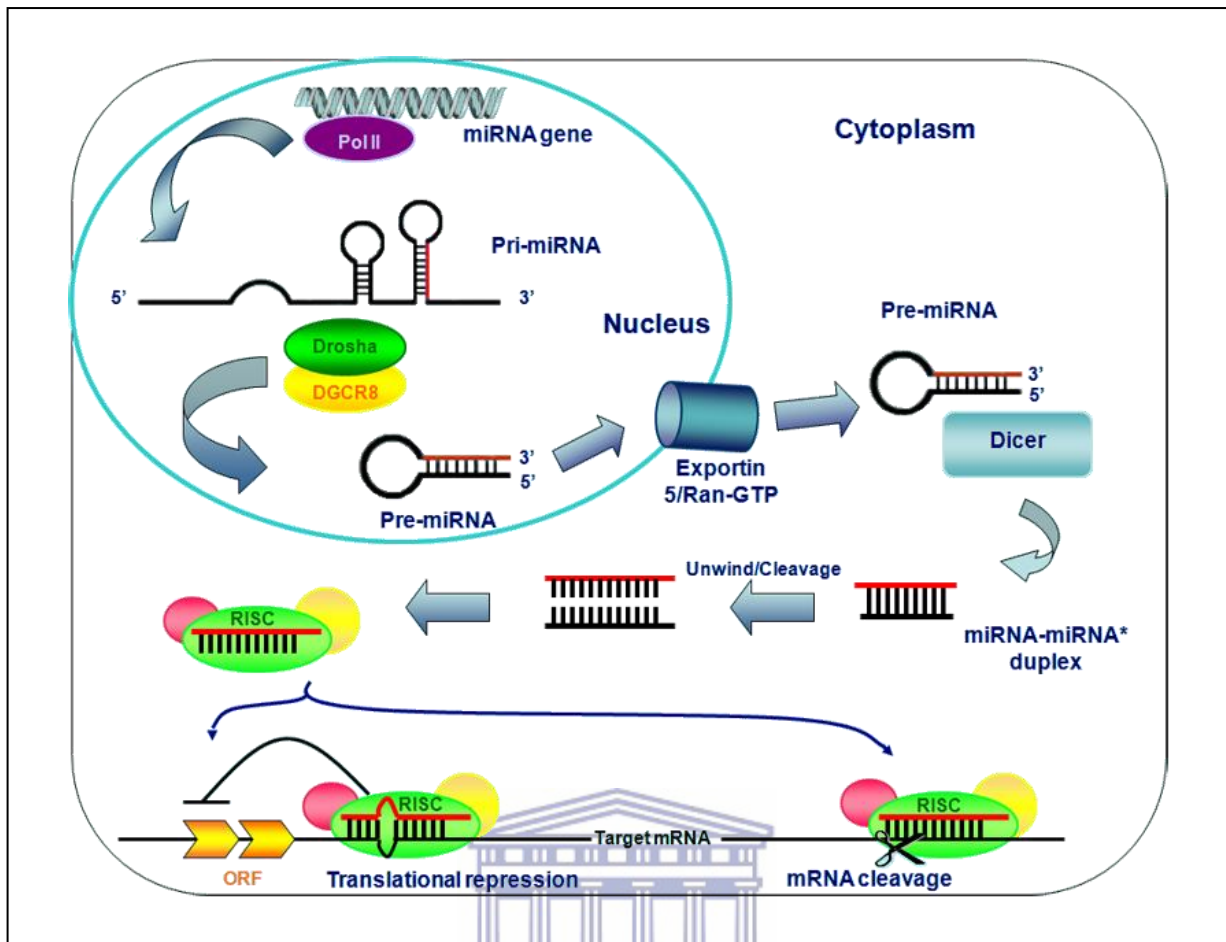
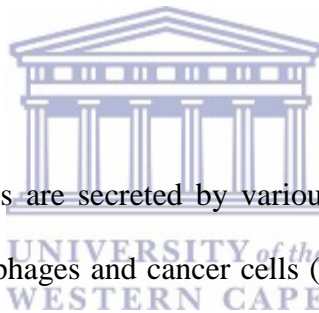


Figure 1.12: Biogenesis of microRNAs. The microRNA gene is transcribed in the nucleus by RNA polymerase II to produce pri-miRNA. The pri-miRNA is processed by Drosha and its cofactor producing pre-miRNA. The pre-miRNA is transported into the cytoplasm by the Exportin 5/Ran-GTP complex where it is further processed by the endoribonuclease Dicer, generating a duplex miRNA-miRNA. The miRNA-miRNA duplex assembles into a RNA induced silencing complex (RISC). Within the complex, the duplex unwinds and the one strand is degraded. The single-stranded miRNA strand which is complementary to the target mRNA remains within the complex, RISC. The miRNA-RISC complex binds to the 3' UTR of the target mRNA resulting in mRNA degradation or translation silencing (Calore and Fabbri, 2011).

1.4.2 Circulating microRNAs

A Large amount of miRNAs has been identified outside of cells including bodily fluids such as urine and blood (Allegra *et al.*, 2012). Many of these miRNAs are quite stable and show distinctive expression profiles, therefore, these miRNAs can potentially be an infinite supply of a non-invasive biomarkers for OC diagnosis (Allegra *et al.*, 2012). Recently, miRNAs have been identified in two types of cell-derived lipid vesicles, microvesicles and exosomes. The circulating miRNAs are most commonly found in exosomes. These exosomes don't only contain miRNAs; it also contains proteins, mRNA and enzymes. It is of an endosomal origin and it serves as a bioactive shuttle vesicle by mediating cell-cell communication (Cardenas-Goicoechea *et al.*, 2013).



According to literature, exosomes are secreted by various cells including B cells, T cells, mast cells, dendritic cells, macrophages and cancer cells (Allegra *et al.*, 2012). A large bulk of these exosomes are secreted by OC cell compared to normal cells. The secretion of exosomal miRNAs in non-malignant disorders is also significantly different compared to miRNA profiles from OC patients (Goy *et al.*, 2012). This inappropriate release of exosomal miRNAs result in a modification, in biological pathways which affect the development of the disease (Enomoto *et al.*, 2012). Other than the increased specificity of circulating miRNAs to the disease, it is also very stable which is a prerequisite for an effective biomarker. According to a study done by Chen and associates in 2008, after treating miRNAs with RNase A, more than half of the microRNAs remained intact after three hours following exposure. These circulating miRNAs also remained stable under harsh conditions, including boiling, high or low pH and an extended storage period (Allegra *et al.*, 2012). Considering all these advantages of circulating miRNAs, they have many characteristics of an ideal biomarker for

early diagnosis of OC, most notably related to specificity, sensitivity, non-invasiveness and stability (Allegra *et al.*, 2012).

1.4.3 The role of microRNAs in the process of Epithelial-to-Mesenchymal Transition (EMT) in Ovarian Cancer

The Epithelial-to-Mesenchymal Transition (EMT) is defined as a biological process that allows the conversion of polarized epithelial cobblestone phenotype to acquire a mesenchymal phenotype. The epithelial cells ordinarily interrelate with the basement membrane via its basal surface to undergo these biochemical changes. This biochemical change also includes enhanced migratory capacity that is required for the metastatic process, increased production of extracellular matrix components and elevated resistance to apoptosis (Kalluri and Weinberg, 2009). Cancer cells lose epithelial cell-cell junction which associates with a decrease in the expression of epithelial proteins and an increase in expression of mesenchymal markers during EMT (Ahmad *et al.*, 2010). These changes are related to augmented activity of matrix metalloproteinases (MMPs) which leads to an invasive phenotype. All of these processes lead to an increased invasion and migration of tumours in many cancers such as OC (Ahmad *et al.*, 2010). It has been suggested that multiple miRNA families play significant roles in controlling the EMT biochemical change (Creighton, Chang and Rosen, 2010).

Table 1.2: Epithelial-Mesenchymal-Transition related miRNAs and their target genes

(Kim, Kim and Wang, 2014).

MicroRNAs	Targets	Effect on rumour metastasis
MiR-20a	PTEN	Enhances EMT process
MiR-34		Inhibit EMT process
MiR-125a	ARID3B	Inhibit EMT process
miR-138	SOX4 and HIF-1alpha	Inhibit EMT process
miR-141		Inhibit EMT process
miR-200b		Inhibit EMT process
miR-429		Inhibit EMT process



The role of most microRNAs in regulating EMT is still not clear with just a few microRNAs being evaluated for their role in cancer, specifically Ovarian Cancer (Guo *et al.*, 2011). The microRNAs 125a and 138 are two microRNAs that seem to play a significant role in EMT in Ovarian Cancer. MicroRNA-125a is a tumour suppressor that has been shown to inhibit the EMT process. Overexpression of this microRNA induces the reversal of EMT in highly invasive Ovarian Cancer cells. Other microRNAs that play a role in the transitioning process are illustrated in Table 1.2 (Kim, Kim and Wang, 2014).

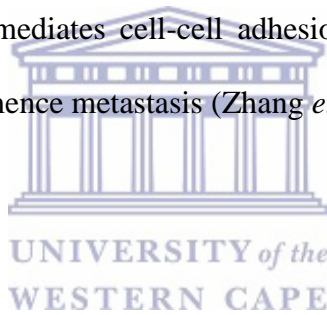
1.4.4 The specific role of microRNAs in Ovarian Cancer

Ovarian Cancer is a complex disease with regards to its multiple histological subtypes documented. Since 2006, there has been an influx in the number of studies demonstrating the

significance of microRNAs in OC. Dysregulation of miRNAs has been document in several human disease including OC. MiRNA profiling studies suggest that they associated with various aspects of OC including tumor histological type, stage, histological grade, prognosis, and therapy resistance. Also, it points to the critical role of miRNAs in OC pathogenesis, tumorigenesis and progression (Croce and Di Leva, 2013; Miska, 2005).

The dysregulation of miRNAs in OC was first described in 2007 by Iorio and associates. In the study the authors examined the dysregulation of miRNAs in OC compared to normal ovaries. They found that mir-141 and miR-200a was upregulated in OC, whereas miR-140, miR-125b, miR-145 and miR-199a was downregulated. They also determined that specific deregulated miRNAs in OC can be used to differentiate the numerous histological subtypes of ovarian carcinoma. For example the miR-200 family is upregulated in mucinous, clear cell and endometriod subtypes; miR-203, miR-21 and miR-205 is upregulated in endometriod carcinomas. Whereas, miR-145 is downregulated in serous and clear cell carcinomas, whilst miR-222 is downregulated in both endometriod and clear cell carcinomas (Croce and Di Leva, 2005; Zhang *et al.*, 2008). One of the most influential studies intergrating miRNAs and Ovarian Cancer was done by Zhang *et al* (2008). In the study the authors utilized an array comparative genomic hybridization approach to identify deregulation in OC. They found that all tumour suppressor miRNA alterations were related to down regulation in late stage ovarian tumours. Also, they showed that down regulation of miRNAs are linked to the loss in DNA copy number and epigenetic silencing. In the case of upregulated miRNAs the chromosomal regions were significantly amplified in multiple cancer samples. Moreover, epigenetic alterations resulted in down regulation of 16 out of 44 miRNAs in late stage Ovarian Cancer (Chauhan *et al.*, 2012; Li *et al.*, 2010; Zhang *et al.*, 2008).

Since then, there have been multiple studies addressing miRNA deregulation in the initiation and progression of OC, using microarray profiling or massive parallel pyrosequencing at the transcriptional level. These studies have shown up- and down regulation in miRNA patterns but down regulated miRNAs were prominent in tumour initiation and progression, especially in high grade OC (Chauhan *et al.*, 2012). During the progression of OC, miRNAs facilitate tumour growth by promoting infiltration of inflammatory cells and inducing tumour angiogenesis or promoting cell adaptation during hypoxia (Chauhan *et al.*, 2012). The most common route of OC metastasis is by means of the transcoelomic route. During the transcoelomic process, the cancer cells undergo EMT, detach from primary tumour, form spheroids and the implant on the peritoneum. The miR-200 plays an important role in this process for example miR-200c mediates cell-cell adhesion and initiate the production of other miR-200 abrogates to commence metastasis (Zhang *et al.*, 2008).



1.5 Problem identification

The development of Ovarian Cancer involves a sequential progression from normal ovarian epithelial cell to preneoplastic ovarian intraepithelial neoplasia and finally invasive Ovarian Cancer. A number of biomarkers have been under investigation for diagnosing OC. Currently, the gold standard for diagnosis is CA-125. However, it's non-specific and it's not sensitive enough. It has moved from being a diagnostic biomarker to a prognostic biomarker. The FDA approved OVA1 test is now used for diagnosis of OC. Regardless of the increased specificity, sensitivity and positive predictive value, its still not performed on its own. It is used in addition to, not in place of, other diagnostic tools. Other diagnostic tools including diagnostic imaging, are invasive or too expensive. The failure of conventional diagnostic approaches to detect OC from an early onset has revealed the need for a novel diagnostic tool

that will enable the detection of OC at its initial stages therefore ensuring a better prognostic outcome. Considering the importance of the issue, the current study was employed to identify potential miRNA biomarkers that can aid in early OC diagnosis, with high specificity and sensitivity using both an *in silico* and molecular approach.

1.6 Aims

The first aim of the research is to characterize the putative miRNA-gene targets identified in a previous study using an *in silico* approach to clearly define their role in cancer, apoptosis and cell differentiation; through gene ontology clustering, transcription factor analysis, co-expression analysis, determining interacting proteins and pathway analysis.



The second aim of the research will be to determine the prognostic significance of the putative genes previously identified through various *in silico* algorithms.

The third aim is to validate which of the six miRNAs – shortlisted by the *in silico* approach – to be specific to OC by generating expression profiles across an array of ovarian cell lines, non cancer cell lines and other cancer cell lines. This will be done using various molecular techniques including cell culture, total RNA extraction, cDNA synthesis and qPCR.

1.8 Reference List

1. Adel, R. and Banerjee, S.K. (2016) Novel Biomarkers to Understand Cardiovascular Complications in Diabetes. *Biochemistry, Genetics and Molecular Biology*. Book article. <http://www.intechopen.com/books/export/citation/EndNote/role-of-biomarkers-in-medicine/novel-biomarkers-to-understand-cardiovascular-complications-in-diabetes>.
2. Agarwal, A., Dada, R., Gupta, S., Michener, C., Rein, B.J.D and Safi, J. (2011) Potential Markers for Detection and Monitoring of Ovarian Cancer. *Journal of Oncology* 10, 1 – 17. doi:10.1155/2011/475983.
3. Ahmad, A., Banerjee, S., Kong, D., Sethi, S. and Wang, Z. (2010) Epithelial-to-Mesenchymal Transition is mechanistically linked with stem cell signatures in prostate cancer. *Plos One*, 5. P. 1-10. doi.org/10.1371/journal.pone.0012445.
4. Al Bakir, M. and Gabra, H. (2014) The molecular genetics of hereditary and sporadic ovarian cancer: implications for the future. *British Medical Bulletin*, 112. P. 57-69. doi: 10.1093/bmb/ldu034.
5. Allocati, N., De Laurenzi, V., Di Ilio, C., Favaloro, B. and Graziano, V. (2012) Role of apoptosis in disease. *Aging*, 4. P. 330-349.
6. Ashton, J. (2015) Understanding Gynecologic Cancer – Female Reproductive Cancers. Gynecologic Cancer Awareness. <http://www.drjenniferashton.com/>.
7. Amant, F., Cadron, I., Daemen, A., De Moor., Despierre, E., Leunen, K., Timmerman, D. And Van Gorp, T. (2011) HE4 and CA125 as a diagnostic test in ovarian cancer: prospective validation of the risk of Ovarian Malignancy Algorithm. *British Journal of Cancer*, 104. P. 863-870.
8. American Cancer Society (2014) What are the key statistics about ovarian cancer. <https://www.cancer.org/cancer/ovarian-cancer/about/key-statistics.html>.

9. Baker, K., Cherry-Bohannon, J. and Francis, H, (2012) VEGF and cholangiocarcinoma: Feeding the tumour. *Translational Gastrointestinal Cancer*, 1. P. 95-102. DOI: 10.3978/j.issn.2224-4778.2011.12.06.
10. Barca – Mayo, O. and Lu, Q.R. (2012) Fine – tuning oligodendrocyte development by microRNAs. *Frontiers in Neuroscience* 6. doi:10.3389/fnins.2012.00013.
11. Berchuck, A., Bentley, R.C., Calingaert, B., Halabi, S., Iversen, E.S., Moorman, P.G., Palmieri, R.T. and Schildkraut, J.M. (2008) Hormonal Risk Factors for Ovarian Cancer in Premenopausal and Postmenopausal Women. *American Journal of Epidemiology*, 167.P.1059-1069.
12. Berk, M. (2015) The Classification of Biomarkers. *Journal of Applied Medical Association of Psychiatry*, 72. 1056-1057. doi:10.1001/jamapsychiatry.2015.1113.
13. Birrer. M.J., Dizon, D., Vathipadiekal, V. And Wei, W. (2013) Ovarian cancer: genomic analysis. *Annals of Oncology*, 24. P. 7-5. doi: 10.1093/annonc/mdt462.
14. Buyers, L. A., Gibbons, D. L. and Kurie, J. M. (2014) Smoking, p53 mutation, and lung cancer. *Molecular Cancer Research*, 12. P. 3-13.
15. Cardenas – Goicoechea, S.J., Chuang, L., Curtin, C., Fishman, D., Gordon, P., Momeni, M. and Nguyen, L. (2013) Biomarkers for early detection of ovarian cancer. *Women’s Health*. (9). P. 171-185. doi: 10.2217/whe.13.2.
16. Calore, F. and Fabbri, M. (2011) MicroRNAs and Cancer. *Atlas of genetics and cytogenetics in oncology and haematology* 16, 51 – 69. doi: 10.4267/2042/47272.
17. Calvert-Joshua, T. (2013) *An in silico identification of putative biomarkers in cervical cancer*. Honours Thesis, University of the Western Cape.
18. Cancer Research UK. (2016) Stages of ovarian cancer. <http://www.cancerresearchuk.org/about-cancer/type/ovarian-cancer/treatment/stages-of-ovarian-cancer>.

19. Cardenas – Goicoechea, S.J., Chuang, L., Curtin, C., Fishman, D., Gordon, P., Momeni, M. and Nguyen, L. (2013) Biomarkers for early detection of ovarian cancer. *Women's Health*.(9). p. 171 – 185. doi: 10.2217/whe.13.2.
20. Casimiro, M. C., Crosariol, M., Loro, E., Li, Z. And Pestell, R. G. (2012) Cyclins and Cell cycle control in cancer and disease. *Genes and Cancer*, 3. P. 649-657.
21. Causters, W., Kumar, A., Rosse, C. and Smith, B. (2005) On carcinomas and other pathological entities. *Comparative and Functional Genetics*, 6. P. 1-8.
22. Chauhan, S., Jaggi, M., Khan, S., Maher, D.M. and Zaman, M.S. (2012) Current status and implications of microRNAs in ovarian cancer diagnosis and therapy. *Journal of Ovarian Research*. [Online] 5 (44). Available from: <http://www.ovarianresearch.com/content/5/1/44> [Accessed 4/03/2017].
23. Chen, G., Shi, L., Yap., Y.L. and Zhang, X. (2007) Integrative Omics Technologies in Cancer Biomarker Discovery. *Landes Biosciences*. P. 129-137.
24. Chen, X., Lui, J., Song, F. and Zheng, H. (2013) Advances in circulating microRNAs as diagnostic and prognostic markers for ovarian cancer. *Cancer, Biology and Medicine* 10, 123-130. doi: 10.7497/j.issn.2095-3941.2013.03.001.
25. Cho, K.R. and Shih, L. (2010) OVARIAN CANCER. *Annual review of pathology*, 4. P. 287-313.
26. Chow, A. Y. (2010) Cell cycle control by oncogenes and tumour suppressors: driving the transformation of normal cells into cancerous cells. *Nature Education*. 3. P.
27. Croce, C.M. and Di Leva, G. (2013) The role of microRNAs in the tumouringensis of ovarian cancer. *Frontiers in Oncology*, 3. P. 1-10. doi: 10.3389/fonc.2013.00153.
28. Corella, D. and Ordovas, J.M. (2015) Biomarkers: background, classification and guidelines for applications in nutritional epidemiology. *Nutrition Hospital*, 31. P 177-188.

29. Correia, N., Dickman, P.W., Magnusson, C.M., Nilsson, S., Nordlinder, H. and Persson, I.R.(2002) Risk Factors for Invasive Epithelial Ovarian Cancer: Results from a Swedish Case –Control Study. *American Journal of Epidemiology* ,156. P.363-373. doi: 10.1093/aje/kwf048.
30. Creighton, C.J., Chang, J.C. and Rosen, J.M. (2010) Epithelial-Mesenchymal-Transition (EMT) in tumour-initiating cells and its clinical applications in breast cancer.*Journal of mammary gland biology and neoplasia*, 2. P. 253-260.
31. Cunha, T.M. and Horta, M. Sex cord-stromal tumours of the ovary: a comprehensive review and update for radiologists. *Diagnostic and interventional Radiology*, 21. P. 277-286. doi: 10.5152/dir.2015.34414.
32. Dai, X., Xiang, L., Li, T. And Bai, Z. (2016) Cancer Hallmarks, Biomarkers and Breast Cancer Molecules. *J Cancer*, 7. P. 1281-1294.
33. DeBerardinis, R.J., Hatzivassiliou, G., Lum, J.J. and Thompson, C.B. (2007) The Biology of Cancer: Metabolic Reprogramming Fuels Cell Growth and Proliferation. *Cell Press*
34. Don, A., Lu, B. And Lu, Y. (2016) Genomic/ Epigenomic Alterations in Ovarian Carcinoma: Translational Insight into Clinical Practice. *Journal of Cancer*, 11. P. 1441-1451. doi: 10.7150/jca.15556.
35. Dorak, M.T. and Karpuzoglu, E. (2012) Gender Differences in Cancer Susceptibility: An Inadequate Addressed Issue. *Frontiers in Genetics*, 3. P. 268-279.
36. Diamandis, E.F. and Kulasingam, V. (2008) Strategies for discovering novel cancer biomarkers through utilization of emerging technologies. *Nature Clinical Practice Oncology*, 5. P. 588-599.
37. Enomoto, T., Fujita, M., Kimura, T., Kimura, T., Kobayashi, E., Matsuzaki, S., Yokoyama, T., Yoshino, K. and Ueda, Y. (2012) Biomarkers for Screening,

- Diagnosis and Monitoring of Ovarian Cancer. *Cancer, Epidemiology, Biomarkers and Prevention*, 21. P. 1902-1912. doi: 10.1158/1055-9965.EPI-12-0646.
38. Evans, P. and Roett, M.A. (2009) Ovarian Cancer: An Overview. *American Family Physician*, 80. P. 609-616.
39. Fernald, K. and Kurokawa, M. (2013) Evading apoptosis in cancer. *Trends in Cell Biology*, 23. P. 620-633.
40. Ghosh, D. and Poisson, L.M. (2008) 'Omics' data and levels of evidence for biomarker discovery. *Genomics*, 93. P. 13-16.
41. Gmeiner, W. H., Hellmann, G. M. and Shen, P. (2008) Tissue-dependent and – independent gene expression changes in metastatic colon cancer. *Oncology reports*, 19. P. 245-251.
42. Goy, A., Li, S., Lim, H., Pecora, A., Sarojini, S., Suh, S., Tamir, A. and Zhang, S. (2012) Early Detection Biomarkers for Ovarian Cancer. *Journal of Oncology*. [Online] 2012 (11). P. UNIVERSITY of the 15. Available from: <http://www.hindawi.com/journals/jo/2012/709049/>. [Accessed 4/03/2015]
43. Groten, J., Borner, C. And Mertelsmann, R. (2016) Understanding and Controlling Cancer: The Hallmark Concept Revisited-Chance, Evolution and Entropy. *Journal of Science, Humanities and Arts*, 3. P. 1-59.
44. Guo, W., Lui, T., Peng, X., Tu, X., Wang, X. and Xiong, D. (2011) Identification of miRs-143 and -145 that is associated with bone metastasis of prostate cancer and involved in the regulation of EMT. *Plos One*, 6.
45. Gupta, S., Rastogi, M. And Sachan, M. (2016) Biomarkers towards Ovarian Cancer Diagnostics: Present and future Prospective. *Brazilian Archives of Biology and Technology*, 59. P. 1678-4324.

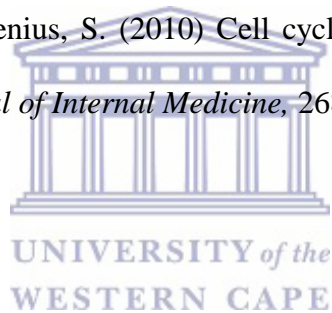
46. Hayes, N.L. and Hayes, D.F. (2012) Cancer biomarkers. *Molecular Oncology*, 6. P. 140-146.
47. Helm, C. W. (2012) Ovarian Cancer Staging. Medscape. <http://emedicine.medscape.com/article/2007140-overview>.
48. Hoff, J.T. and Van Nagell, J.R. (2013) Transvaginal ultrasonography in ovarian cancerscreening: current perspectives. *International journal of women health* 6, 25 – 33. doi:10.2147/IJWH.S38347.
49. Huang, C., Cheng, W., Lee, C., Su, Y., Chien, S., Tzeng, Y., Hsieh, C. And Chen, C. (2006) Serume Mesothelin in Epithelial Ovarian Carcinoma: A New Screening Marker and Prognostic Factor. *Anticancer Research*, 26. P. 4721-4728.
50. Huncharek, M. S. and Muscat, J. E. (2009) Perineal Talc Use and Ovarian Cancer: A Critical Review. *European Journal of cancer prevention*, 17. P. 139-146.
51. Isaacs, H. (2013) Germ Cell Tumours. *Tumours of the Fetus and Infant*, 3. P. 5-31.
52. Kalluri R, Weinberg RA (2009): The basics of epithelial-mesenchymal-transition. *Journal of Clinical Investment*, 119.P.1420-1428.
53. Kelland, L. (2007) Targeting the limitless reproductive potential of cancer: the telomerase/telomere pathway. *Clinical Cancer Research*, 13. P. 4960-4963.
54. Kim, I., Kim, S. and Wang, Y. (2014) Regulatio of metastasis by microRNAs in ovarian cancer. *Frontiers in Oncology*, 4.
55. Kimura, T., Kinose, Y., Nakamura, K., Nakatsuka, E., Sawada, K. and Yoshimura, A. (2016) Clinical relevance of circulating cell-free microRNAs in ovarian cancer. *Molecular Cancer*, 15. P. 1-10.
56. KK Women's and Children's Hospital.(2014) Ovarian Cancer surgery. SingHealth. <http://www.singhealth.com.sg/PatientCare/ConditionsAndTreatments/Pages/Ovarian-Cancer-Surgery.aspx>.

57. Kristjansdottir, B., Levan, K., Partheen, K. and Sundfeldt, K. (2013) Diagnostic performance of the biomarkers HE4 and CA125 in type I and type II epithelial ovarian cancer. *Gynecology Oncology*, 131. P. 52-58. doi: 10.1016/j.ygyno.2013.07.094.
58. Konishi, I., Koshiyama, M. and Matsumura, N. (2014) Recent Concepts of Ovarian Carcinogenesis: Type I and Type II. *BioMed Research International*.
59. Kurman, R.J. and Shih, I. (2010) The Origin and Pathogenesis of Epithelial Ovarian Cancer: A Proposed Unifying Theory. *American Journal of Surgery Pathology*, 34. P. 433-443.
60. Lewis, J. (2008) *Computing genomic sciences: bioinformatics and standardised in proteomics*. PhD Thesis, Cardiff University.
61. Li, S., Wan, X., Wang, Y. And Zhang, J. (2010) The role of microRNAs in ovarian cancer initiation and progression. *Journal of Cellular and Molecular Medicine*, 16. P. 2240-2249.
62. Lockwood, S. and Ritzert, B. (2013) Cost-effectiveness of serum CA-125 compared to transvaginal ultrasound as a screening test for ovarian cancer: a systemic review protocol. *JBI Database of Systemic Reviews and Implementation Reports* 11, 1-10.
63. Lottering, S (2015) *Identification of miRNAs as biomarkers for early diagnosis of Ovarian Cancer: An in silico approach*. Honours thesis. University of the Western Cape.
64. Mayeux, R. (2004) Biomarkers: Potential Uses and Limitations. *The Journal of the American Society for Experimental NeuroTherapeutics*, 1. P. 182-188.
65. MD Guidelines (2010) Cancer, Ovary. ReedsGroupMDGuidelines. <http://www.mdguidelines.com/cancer-ovary>.

66. Mishra, A. and Verma, M. (2010) Cancer Biomarkers: Are We Ready For The Prime Time?. *Cancer*, 2. P. 190-208. doi:10.3390/cancers2010190.
67. Moore, R.G., Brown, A.K., Miller, M.C., Skates, S., Allard, W.J. and Verch, T. (2008) The use of multiple novel tumour biomarkers for the detection of ovarian carcinoma in patients with a pelvic mass. *Gynecology Oncology*, 108. P. 402-408.
68. Mutch, D.G. and Prat, J. (2014) 2014 FIGO staging for ovarian, fallopian tube and peritoneal cancer. *Gynecological Oncology*, 133. P. 401-404. doi: 10.1016/j.ygyno.2014.04.013.
69. National Cancer Institute (2015) Ovarian Epithelial, Fallopian Tube, and Primary Peritoneal Cancer Treatment (PDQ®)-Patient Version. National Institute of Health. <https://www.cancer.gov/types/ovarian/patient/ovarian-epithelial-treatment-pdq>.
70. Ngcoza, N. (2013) *Identification of biomarkers in breast cancer as potential diagnostics and therapeutics agents using a combined in silico and molecular approach*. MSc Thesis, University of the Western Cape.
71. Ohmichi, M. and Tanaka, T. (2012) Recurrent ovarian cancer presenting in the rightsupraclavicular lymph node with isolated metastasis: a case report. *Journal of medical casereports* 6, 1 – 4. doi: 10.1186/1752-1947-6-176.
72. Pezzella, F., Harris, A. L., Tavassoli, M. And Gatter, K. C. (2015) Blood vessels and cancer much more than just angiogenesis. *Cell Death and Discovery*, 1. P. 1-2.
73. Pollack, J.R. (2007) A perspective on DNA microarrays in pathology research and practice. *American Journal of Pathology*, 171. P. 375-385.
74. Sadikovic, B., Al-Romaih, K., Squire, J. A. and Zielenska, M. (2008) Causes and Consequences of Genetic and Epigenetic Alterations in Human Cancer. *Current Genomics*, 9. P. 394-408.

75. Sallam, R.M. (2015) Proteomics in cancer biomarkers discovery: challenges and applications. *Disease biomarkers*, 5. P. 1-12. <http://dx.doi.org/10.1155/2015/321370>
76. Shamma, M. A. (2012) Telomeres, lifestyle, cancer, and aging. *Current opinion in Clinical Nutrition and Metabolic Care*, 14. P. 28-34.
77. Shay, J. W. and Wright, W. E. (2000) Hayflick, his limit, and cellular ageing. *Nature Reviews. Molecular cell biology*, 1. P. 72-76.
78. Simundic, A. (2009) Measure of Diagnostic Accuracy: Basic Definitions. *The Journal of the International of Clinical Chemistry and Laboratory Medicine*, 19. P. 203-211.
79. Singh, N.M. (2011) What is a Biomarker? From its discovery to clinical application. *Journal of Medical Biochemistry*, 30. P. 186-192.
80. Stefan, D.C. (2015) Why is cancer not a priority in South Africa? *The South African Medical Journal*, 105. P. 103 – 104. <http://dx.doi.org/10.7196/SAMJ.9301>.
81. Srivastava, S., Srinivas, P.R., Verma, M. and Zhao, Y. (2002) Proteomics for cancer biomarker discovery. *Clinical chemistry*, 48. P. 1160-1169.
82. Sung, J. J. And Wu. C. (2013) Colorectal cancer screening: are stool and blood based tests good enough?. *Chinese Clinical Oncology*, 2. P. 1-7.
83. Tinelli, A., Vergara, D., Martignago, R., Leo, G., Malvasi, A., Tinelli, R., Marsigliante, S., Maffia, M. And Lorus, V. (2007) Ovarian Cancer Biomarkers: A focus on Genomic and Proteomic Findings. *Current Genomics*, 8. P. 335-342.
84. Miska, E.A. (2005) How microRNAs control cell division, differentiation and death. *Elsevier*, 15. P. 563-568.
85. Movva, S. (2015) Understanding Cancer-The basics. *WebMD*. Available at: <http://www.webmd.com/cancer/understanding-cancer-basics>.
86. Valastyan, S. and Weinberg, R.A. (2011) Tumour Metastasis: Molecular Insights and Evolving Paradigms. *Cell*, 147. P. 275-292.

87. Weis, S. M. And Cheresh, D. A. (2011) Tumour angiogenesis: molecular pathways and therapeutic targets. *Nature medicine*, 17. P. 1359-1370.
88. Witsch, E., Sela, M. And Yarden, Y. (2010) Roles for growth factors in cancer progression. *Physiology*, 25. P. 85-101.
89. Xio-Ping, W., Wang, Y., Wan, X. and Li, S. (2010) The role of microRNAs in ovarian cancer initiation and progression. *Journal of cell molecular medicine*, 14. P. 2240-2249.
90. Zhang, L., Volinia, S. and Bonome, T. (2008) Genomic and epigenetic alterations deregulate microRNA expression in human epithelial ovarian cancer. *Academic science*, 105. P. 7004-7009.
91. Zhivotovskt, B. And Orrenius, S. (2010) Cell cycle and cell death in disease: past, present and future. *Journal of Internal Medicine*, 268. P. 395-409.



Chapter 2

Characterization of putative microRNA target genes

2.1 Introduction

2.1.1 Functional genomics

2.1.2 Previous study

2.1.3 Aim

2.1.4 Objectives

2.2 Methodology

2.2.1 Gene Ontology (GO) clustering

2.2.2 Protein-protein interaction analysis using STRING and GeneMANIA

2.2.2.1 STRING database

2.2.2.2 GeneMANIA

2.2.3 Transcription factor binding site analysis

2.2.3.1 GeneCards

2.2.3.2 TfactS®

2.2.4 Pathway analysis

2.3 Results

2.3.1 Gene Ontology (GO) clustering

2.3.2 Co-expression analysis using STRING and GeneMANIA



2.3.2.1 STRING database

2.3.2.2 GeneMANIA

2.3.3 Transcription factor analysis

2.3.4 Pathway analysis

2.4 Discussion

2.4.1 Gene Ontology (GO)

2.4.2 Protein-protein network generation

2.4.2.1 STRING

2.4.2.2 GeneMANIA

2.4.3 Transcription factor analysis

2.4.4 Pathway analysis

2.5 Conclusion

2.6 Reference List

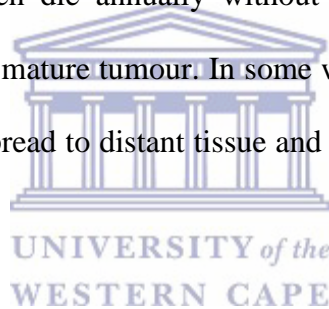


Chapter 2

Characterization of putative microRNA target genes

2.1 Introduction

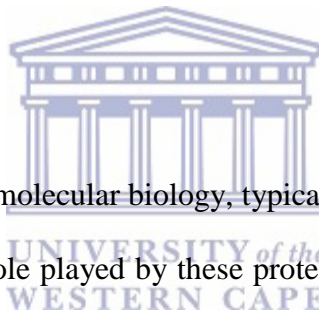
Cancer is a class of diseases characterized by uncontrollable cell division. There are more than 200 different types of cancer, classified by the site of origin (Movva, 2015). For the purpose of this research we will focus on Ovarian Cancer (OC). OC is the most common reproductive and the most lethal gynaecologic cancer in women around the world (Movva, 2015). OC is the eighth most common form of cancer in women world-wide and in South Africa approximately 800 women die annually without diagnosis. OC follows a natural course ultimately resulting into a mature tumour. In some women if OC is not detected while the cancer is localized it might spread to distant tissue and organs, leading to the death of the individual (Movva, 2015).



Presently, OC is diagnosed through pelvic examination, transvaginal ultrasonography, imaging studies and serum protein biomarkers. However, these methods are invasive, lack sensitivity, specificity and positive predictive value and some are expensive (Goy *et al.*, 2012). Therefore, a less invasive method for early detection is required, with the ability to overcome the shortcomings of the current diagnostic tools. Biomarkers which act as biological indicators of the disease have emerged as a viable option for the early detection of the disease including DNA, RNA, proteins and microRNAs (miRNAs) (Goy *et al.*, 2012; Kartha, Subramanian and Sundarbose, 2013). MicroRNA activity has been reported in various diseases including cancer; and studies have shown that miRNAs are often deregulated. These small non-coding RNAs bind to target sequences in mRNAs, generally resulting in repressed

gene expression. The involvement of miRNAs in key cellular processes such as cell death and their negative control over the expression of numerous oncoproteins make them prime candidates as cancer biomarkers (George and Mittal, 2010). Reports have described a one-to-one, one-to-multiple or multiple-to-multiple association between the miRNAs and its target genes in most human cancers (Hashimoto *et al.*, 2013). Characterizing these target genes functionally includes; finding biological processes, pathways, intersection modules or causative network signatures within the onset and progression of a disease and will add function to the miRNAs that could potentially be a candidate biomarker (Hashimoto *et al.*, 2013; Missiaglia *et al.*, 2017).

2.1.1 Functional genomics



Functional genomics is a field of molecular biology, typically described as the study of genes, their resulting proteins, and the role played by these proteins within the bodies' biochemical processes. It integrates a vast wealth of data produced by various molecular methodologies such as genomic and transcriptomic techniques to describe genes/proteins (illustrated in Figure 2.1). The vast amount of data is compiled in databases to ease the increase in genomic, proteomic and transcriptomic data (Cui *et al.*, 2016). Functional genomics enable the understanding of complex relationships between genotype and phenotype on a genome-wide scale. A range of processes such as transcription, translation and epigenetic regulation is investigated in an attempt to answer biological questions including (i) when and where genes are expressed? (ii) how do gene expression levels differ in various cell types and states? (iii) what are the functional roles of different genes and in what cellular processes do they participate? (iv) how genes are regulated and where the active gene promoters are in a particular cell type? (v) how do genes/proteins interact with one another? and (vi) how does

gene expression change in various diseases or following a treatment? (Boehm and Hahn, 2011; Hu *et al.*, 2014).

Functional genomics experiments typically utilize large-scale, high-throughput assays to measure and track many genes and proteins in parallel under different experimental or environmental conditions such as microarray technology (Cui *et al.*, 2016). This ‘genome-wide’ approach allows the function of different parts of the genome to be discovered by combining information from genes, transcripts and proteins (Cui *et al.*, 2016). Various bioinformatics’ visualisation techniques are vital to infer genes/proteins function(s). Functional characterization using bioinformatics enables investigators to simultaneously examine changes in expression, regulation and biological conditions. Functional analysis enables a ‘large’ interesting gene list to be further analysed with the aim of achieving a ‘smaller’ priority gene list; and research of gene functions whereby the miRNAs that target the genes can be indirectly linked to the specific classifications identified (Cui *et al.*, 2016; Masters, McAteer and Merlin, 2002).

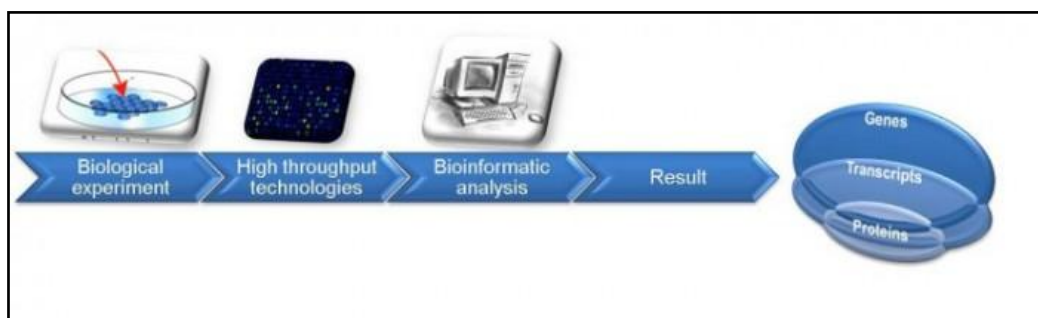


Figure 2.1: The integration of various molecular techniques of how DNA sequence is translated into complex information in a cell (Masters, McAteer and Merlin, 2002).

2.1.2 Previous study

Previously, an *in silico* approach was employed to identify a list of miRNAs potentially implicated in Ovarian Cancer using the mir2disease database (Lottering, 2015). The targets of the miRNAs were identified using databases such as TargetScan Human and miRDB. Furthermore, cell surface microRNA target genes were implicated in the initiation and progression of Ovarian Cancer through various biological databases such as TargetScan Human and MiRDB. TargetScan Human (<http://www.targetscan.org/>) searches for the presence of conserved 7mer and 8mer sites that match the seed region of each miRNA thereby predicting biological targets of miRNAs whilst miRDB (<http://www.mirdb.org/cgi-bin/search.cgi>) searches for conserved and non conserved gene targets, based on the 3'-UTR, by treating target site conservation as an important but not-required sequence feature. Further *in silico* analysis was done including functional annotation, intense literature mining, tissue expression analysis, co-expression analysis and pathway analysis; to identify six miRNAs and seven miRNA target genes not inferred or proven as Ovarian Cancer biomarkers (Table 2.1) at the time of the study as potential biomarkers for the sensitive, accurate and early detection of OC as well as to manage the disease outcomes following treatment(s) (Lottering, 2015). The sequence or ID's of the miRNA's won't be disclosed in the thesis to protect the Intellectual Property (IP) generated through the research as well as future patent applications.

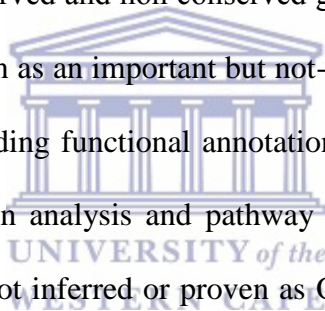



Table 2.1: MicroRNAs and their gene specific targets.

MicroRNAs	Gene targets	Gene description
MIR1, MIR2	CILP	Cartilage Intermediate Layer Protein
MIR3	CRTAP	Cartilage Associated Protein
MIR3	P2RX1	Purinergic Receptor P2X1
MIR4	NPR3	Natriuretic Peptide Receptor 3
MIR3, MIR6	FARP1	FERM, ARH/RhoGEF And Pleckstrin Domain Protein 1
MIR3	STAB2	Stabilin 2
MIR5	NDEL1	NudE Neurodevelopment Protein 1 Like 1



The logo of the University of the Western Cape, featuring a classical building facade with columns and a pediment, with the text 'UNIVERSITY of the WESTERN CAPE' below it.

2.1.3 Aim

The aim of the chapter is to functionally annotate the seven identified miRNA target genes for their involvement in the onset and progression of OC using several *in silico* approaches. The genes implicated in OC will be related back to the miRNAs that regulate those genes subsequently implicating the related miRNAs in OC management.

2.1.4 Objectives

- Gene Ontology (GO) clustering of the target genes to identify molecular and biological processes the genes are involved in
- Protein-protein interaction analysis of the candidate genes to identify association(s) with known cancer causing genes and specifically genes already implicated in OC
- Transcription factor binding analysis of candidate genes to identify the underlying regulatory networks controlling transcription of these genes
- Gene pathway analysis of the candidate genes to place the genes in very specific cancer causing pathways and more specific OC related pathways



2.2 Methodology

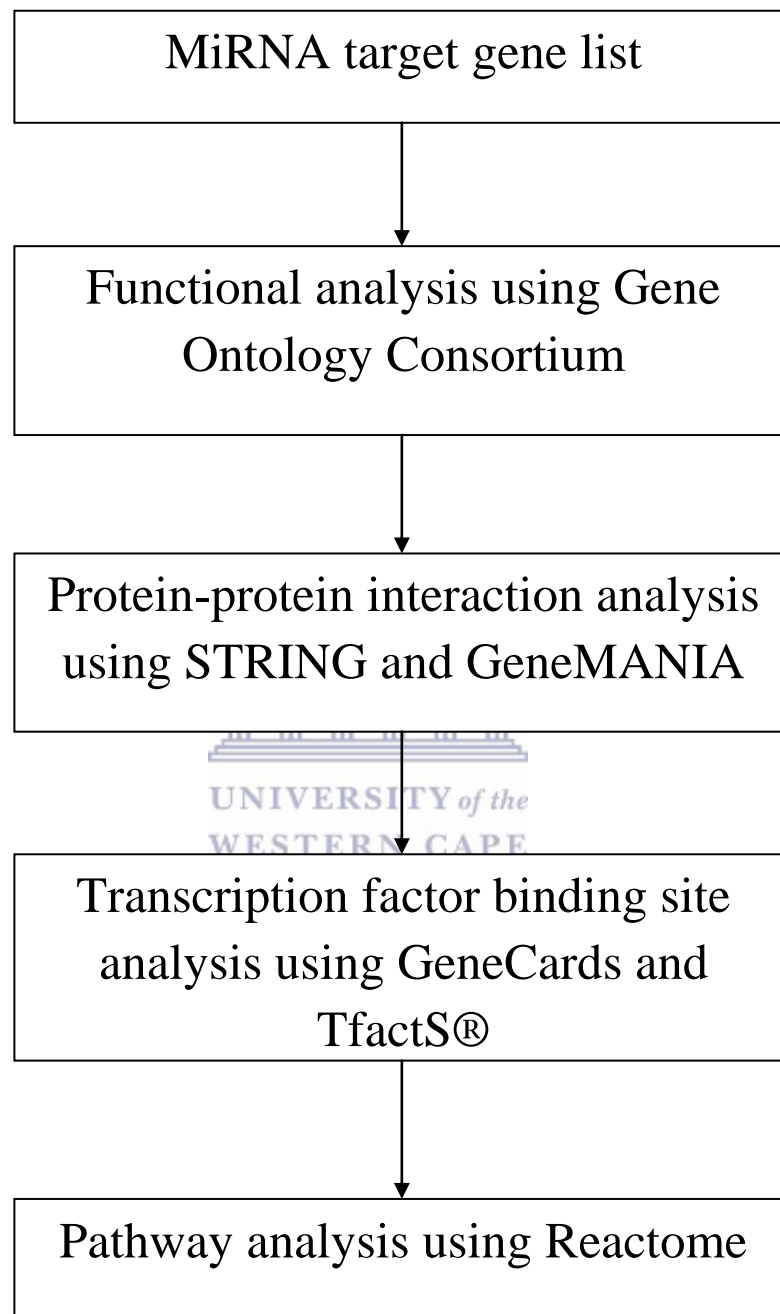


Figure 2.2: The flow chart representing the *in silico* methodologies employed to functionally characterize the miRNA target genes during this study/chapter.

2.2.1 Gene Ontology (GO) clustering

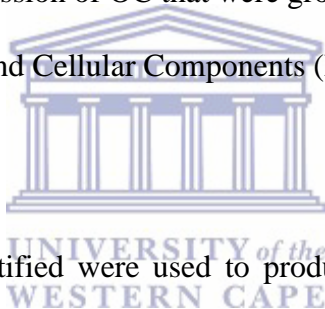
Gene Ontology Consortium is a publically available database, available at <http://www.geneontology.org/>. It is a collaborative effort to address two aspects of information integration including (i) providing consistent descriptors for gene products in different databases and (ii) standardizing classifications for sequences and sequence features in different databases (Coulibaly and Page, 2008). Gene Ontology Consortium describes gene products in terms of their associated Biological Processes (BPs), Cellular Components (CCs) and Molecular Functions (MFs) in a species-independent manner. BPs is a series of events accomplished by one or more assemblies of MFs. CCs include the classification of genes in sub-cellular structures and locations. MFs describe activities that occur at a molecular level such as catalytic or binding activities (Coulibaly and Page, 2008). For the purpose of this research, the MFs and the BPs were the focus. The Gene Ids for the seven target genes were used as an input on the homepage of the database (<http://www.geneontology.org/>). '*Homo sapiens*' was selected as the species of interest and the query was submitted. The BPs and MFs identified were retrieved for further analysis in a Word document.

2.2.2 Protein-protein interaction analysis using STRING and GeneMANIA

Gene IDs for the seven genes, targeted by six miRNAs inferred to be implicated in Ovarian Cancer were used as an input for the generation of a gene network using the Search Tool for the Interacting Gene/Proteins (STRING) database (Bork *et al.*, 2005) and GeneMANIA (Mostafavi *et al.*, 2008).

2.2.2.1 STRING database

STRING is an online database available at <http://string-db.org/> which provides a comprehensive collection of protein-protein interactions for more than 2000 organisms including *Homo sapiens*, *Saccharomyces cerevisiae* and *Escherichia coli* K12 MG1655. These interactions are important for the understanding of Molecular Functions and Biological Processes. The associations between proteins are derived from high throughput experimental data, from the mining of databases and literature; and from predictions based on genomic context analysis (Bork *et al.*, 2005). The bioinformatics tool was used to generate a network of protein-protein interactions between the genes of interest and important proteins implicated in the onset and progression of OC that were grouped according to their Biological Processes, Molecular Functions and Cellular Components (Bork *et al.*, 2005).



The 7 target genes/proteins identified were used to produce expression networks. The url <http://string-db.org/> was launched using the Google search engine. On the homepage, ‘multiple proteins’ was selected and the 7 prioritised target genes IDs were uploaded. Once the genes were uploaded multiple species were identified and ‘*Homo sapiens*’ were selected from which a biological expression network were generated for the seven target genes. To produce the expression networks the parameters were set as follows: (i) a confidence interval of 0.9 and (ii) a limit of 20 interactions shown. The biological expression network generated was exported and saved in a Word document for further analysis.

2.2.2.2 GeneMANIA

GeneMANIA available at <http://pages.genemania.org/> uses a heuristic algorithm derived from ridge regression to predict the function of a set of input genes. It functions by finding directly interrelated/interacting genes and uses functional association from multiple genomics and proteomics network data to link genes/proteins of interest in real-time (Mostafavi *et al.*, 2008). Two genes are linked if their expression levels are similar across a specific condition in a gene expression study. The data is collected from publications within Gene Expression Omnibus (Mostafavi *et al.*, 2008). The url <http://pages.genemania.org/> was launched on the Google homepage. The 7 target gene ID's were uploaded collectively and '*Homo sapiens*' were selected as the species of interest. The expression network generated was exported into a Word document and saved for further analysis.



2.2.3 Transcription factor binding site analysis

Two publically available databases were used to determine the Relevant Transcription Factors (TFs) for each gene namely GeneCards available at <http://www.genecards.org/> and TfactS® available at <http://www.tfacts.org/>. These TFs were validated for their association with OC through published literature following their identification.

2.2.3.1 GeneCards

GeneCards is a publically available database that describes the query genes in terms of proteomics, genetics, transcriptomics, disease and functional information providing a summary of a specific gene. The information is integrated from various data sources in this

database to provide a coherent picture regarding underlying regulatory networks of a subset of genes thus inferring functions to the subset of genes under investigation (Dalah, Stein and Stelzer, 2011). In the Google search engine, <http://www.genecards.org/> was launched. Each of the 7 genes previously identified were individually uploaded using the gene name under ‘search term’ and ‘genomics’ was selected as the analysis type. Transcription factor binding sites were selected and the relevant transcription factors were saved as a Word document for further analysis.

2.2.3.2 TfactS®

TfactS® predicts the regulation, inhibition or activation of the TFs in a biological system based on a list of genes that are up-regulated and down-regulated in microarray experiments (Essaghir *et al.*, 2010). In the TfactS® database, each miRNA target gene was uploaded as TF query using the gene name. The default settings of the software were used. The list containing the regulated TFs was extracted for further analysis. The relevant transcription factors that overlapped in the two databases were subjected to literature mining to determine its regulatory importance in OC for their associated genes.

2.2.4 Pathway analysis using Reactome

Reactome pathway database available at <http://www.reactome.org/> is an open-source, curated and peer reviewed biological pathway database. There are various reactomes that focus on different organisms of which the largest is focused on human biology. Pathway annotations are authored by expert biologists, in collaboration with Reactome editorial staff and cross-referenced to many bioinformatics databases. These include NCBI Gene, Ensembl and

UniProt databases, the UCSC Genome Browser, the KEGG Compound and ChEBI small molecule databases, PubMed, and Gene Ontology (Croft *et al.*, 2011). Reactome provides an intuitive website to navigate pathway knowledge and a suite of data analysis tools to support the pathway-based analysis of complex experimental and computational data sets. Visualisation of Reactome data is facilitated by the Pathway Browser, a Systems Biology Graphical Notation (SBGN)-based interface. It exploits the PSIQUIC web services to overlay molecular interaction data from the Reactome Functional Interaction Network and external interaction databases such as IntAct, ChEMBL, BioGRID and iRefIndex (Croft *et al.*, 2011). The rationale behind the Reactome database is to convey the rich information in a visual representation of biological pathways in a detailed, computationally accessible format (Croft *et al.*, 2011).



On the homepage (<http://www.reactome.org/>), 'Browse pathway' was selected. The 7 target gene ID's were uploaded under search and '*Homo sapiens*' were selected. The networks generated were exported to a Word document for further analysis. Furthermore, the pathways identified were subjected to literature mining to determine the relevant pathways involved in the initiation and progression of cancer including OC.

2.3 Results

2.3.1 Gene Ontology (GO) clustering using Gene Ontology Consortium

The genes of interest previously identified were used to determine Molecular Functions and Biological Processes that the genes play a role in. The seven target genes were associated with Biological Processes such as ‘negative regulation of insulin like growth factor receptor signalling pathway’ and ‘negative regulation of post translational protein modification’ depicted in Figure 2.3. Furthermore, the target genes were associated with Molecular Functions including ‘hormone binding’, ‘protein binding’ and ‘cation binding’ illustrated in Figure 2.3. These terms have been linked to cancer specifically Ovarian Cancer (Smith, Steffen and Williams, 2003; Bach, Fu and Yang, 2013; Harris and Terry, 2016).



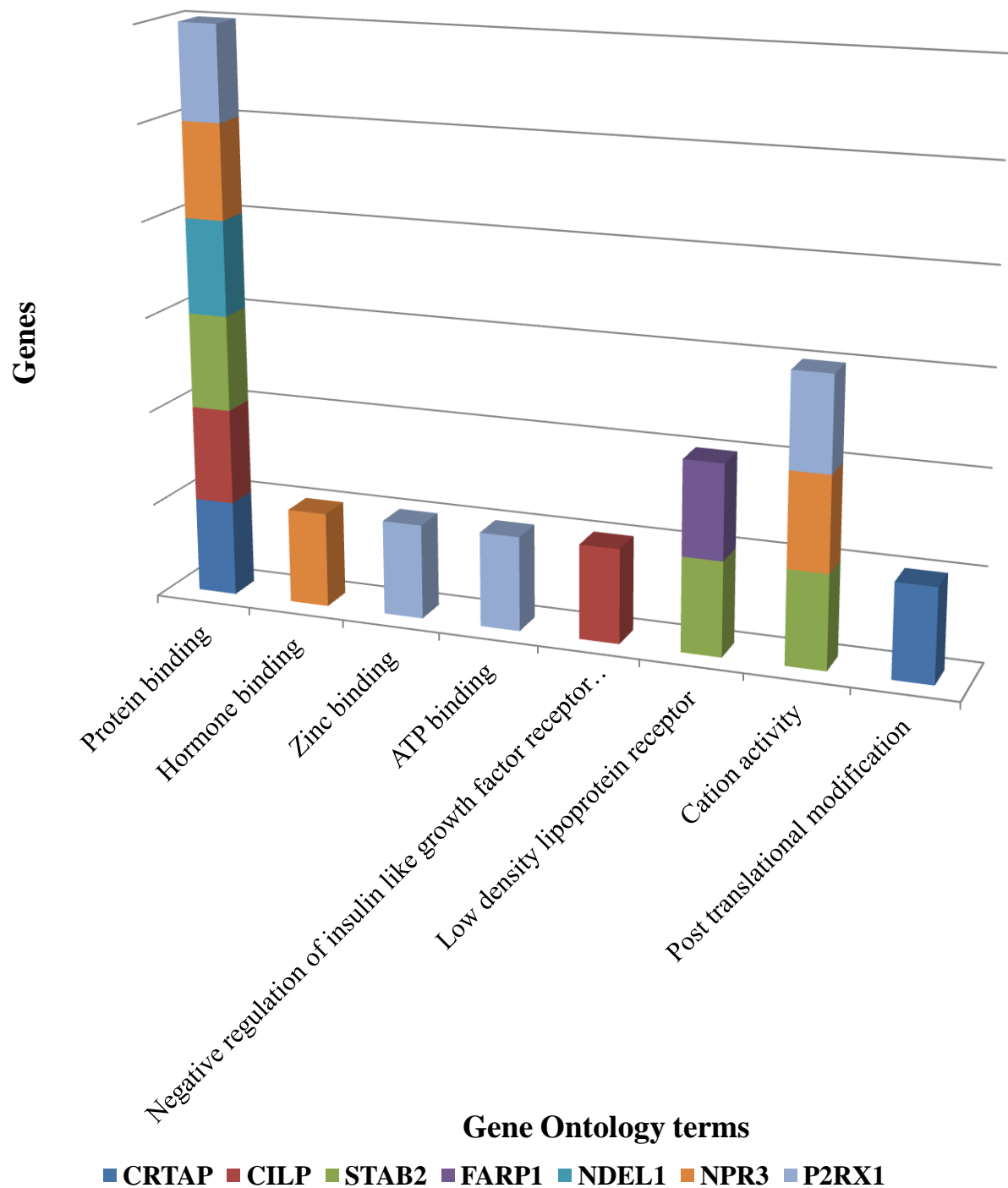


Figure 2.3: Graphical representations of Gene Ontology (GO) clustering of the seven candidate miRNA targeted genes based on their Biological Processes (BPs) and Molecular Functions (MFs). The bars represent the genes associated with specific MF and BP terms.

2.3.2 Co-expression analysis using STRING and GeneMANIA

2.3.2.1 STRING

According to STRING, NDEL1 is present in the main expression network illustrated in Figure 2.4. There are 1 out of the 7 candidate miRNA target genes that are not in the main network but the genes are connected to smaller networks shown in Figure 2.4. Two of the target genes (P2RX1 and FARP1) in the left-hand corner of Figure 2.4 are not present in any of the networks generated by the database (Box B). Networks were generated based on the evidence indicated by the interacting coloured lines decoded by the figure legend in the right-hand corner.



2.3.2.2 GeneMANIA

The 7 candidate genes identified in a previous study were used to produce expression networks shown in Figure 2.5. In Figure 2.5, 5 out of the 7 genes are co-expressed in one of the expression networks and 2 out of the 7 genes are depicted in the second network (right). Based on the networks presented, there is a strong probability that the candidate genes are co-expressing in the same biological/disease processes.

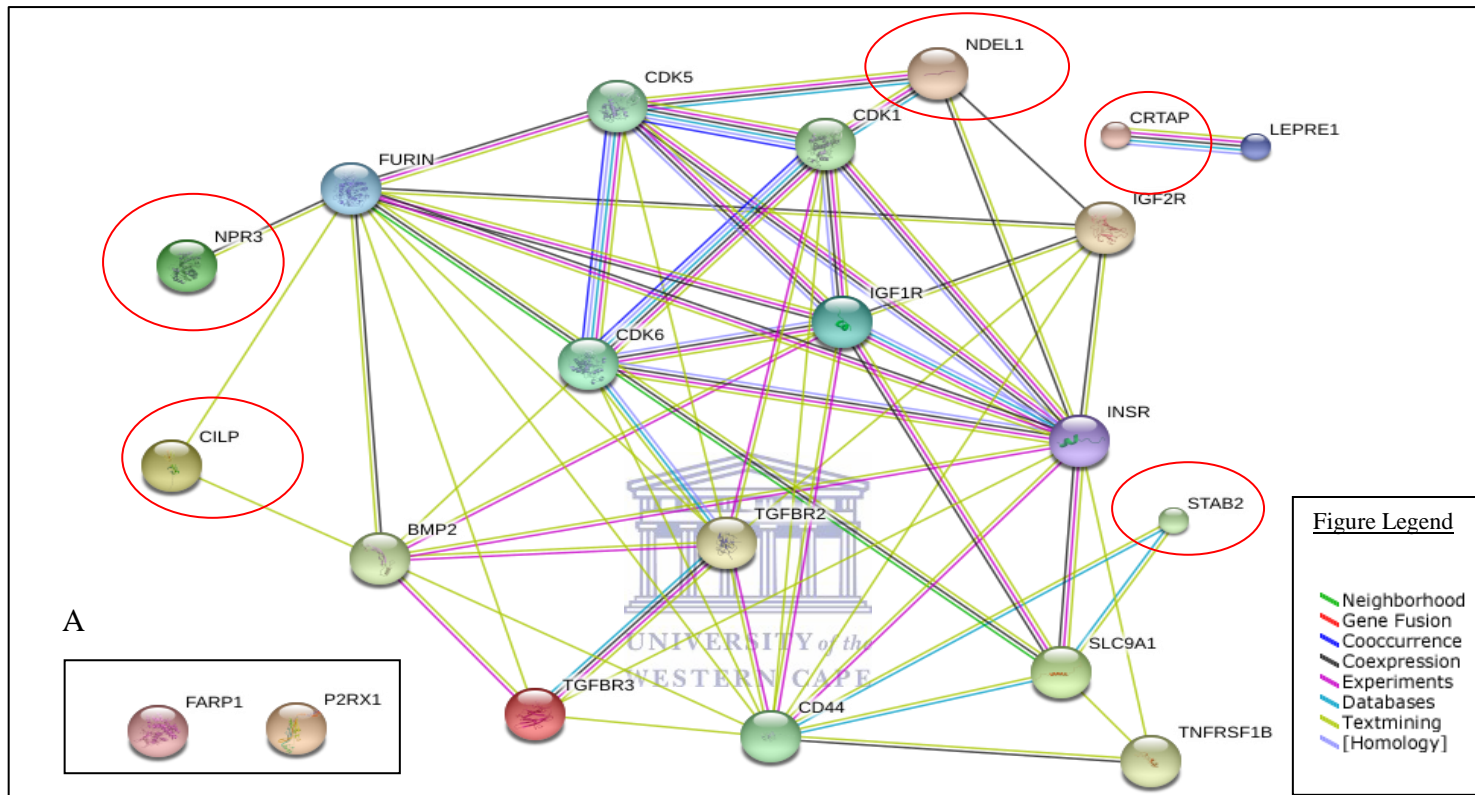


Figure 2.4: The seven candidate genes and their interacting genes produced by (STRING, 2017). STRING analysis shows the interactions of the seven miRNA target genes clustered together using MCL clustering. The nodes represent the genes and the lines joining them represent the evidence available for the connection between the genes as decoded in the legend in the right-hand corner. Box A shows the two target genes not present in the expression network and the 5 out of the 7 target genes were associated with various cancer-related proteins (shown in red ovals).

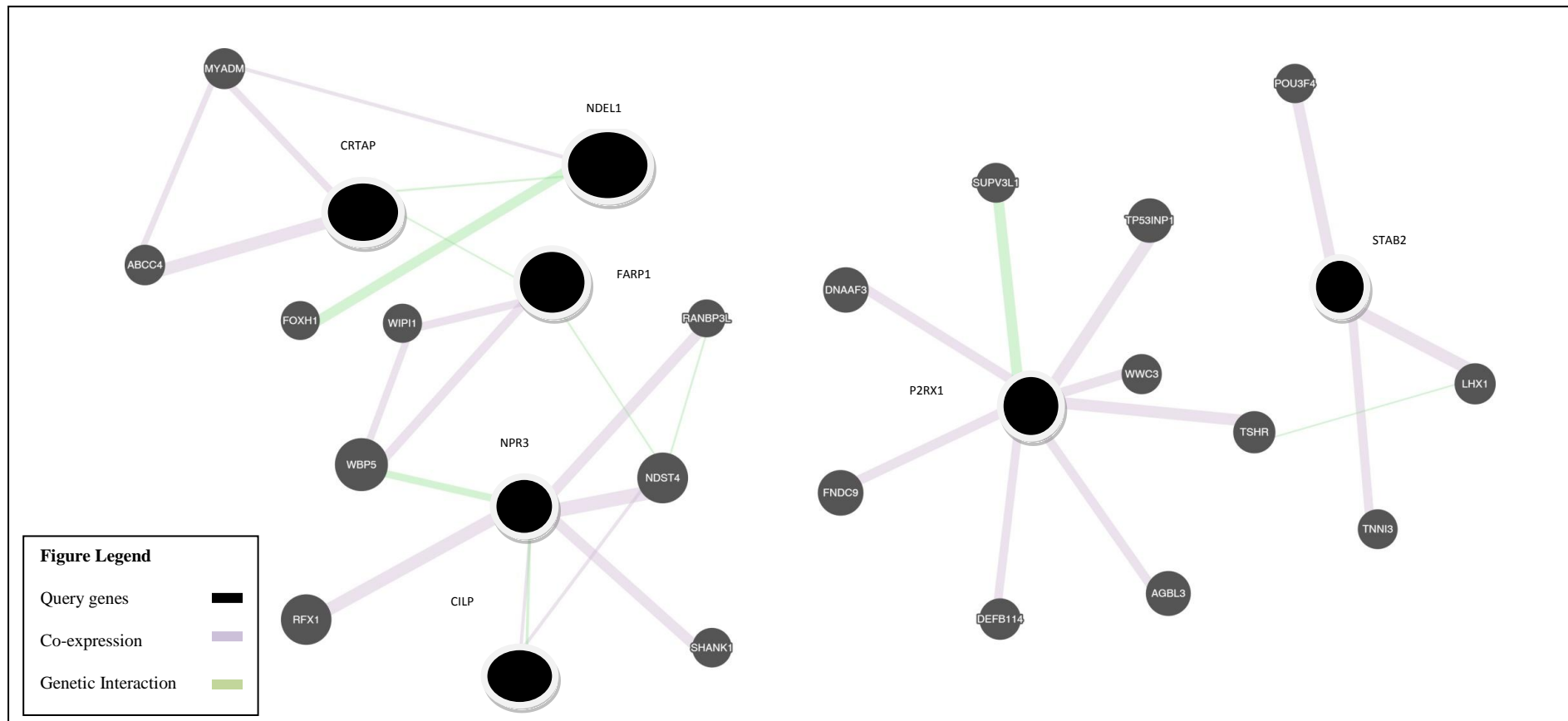


Figure 2.5: Co-expression analysis displaying the expression network of the putative genes and other genes linked to the network (generated by GeneMANIA, 2017). The network includes the predicted miRNA target genes (highlighted in black) and co-expressed genes (shown in grey).

2.3.3 Transcription factor analysis

The seven target genes were used to determine the transcription factors that are linked to their promoter regions, to confirm that the putative genes have a connection to cancer through their underlying regulatory networks and how the function (s) of these transcription factors relates to cancer, specifically Ovarian Cancer. All possible transcription factors regulating the candidate target genes were extracted from GeneCards and TfactS® illustrated in Table 2.2.



Table 2.2: Transcription Factors associated with the seven putative Ovarian Cancer genes produced by GeneCards and TfactS®

Gene	Regulatory Transcription Factor
CILP	PPAR-gamma1, PPAR-gamma2, c-Myc, ER-alpha, AhR
CRTAP	ER-alpha, Elk-1, MEF-2, MEF-2A, AhR
P2RX1	P53, PPAR-gamma1, PPAR-gamma2, TBP, NF-kappaB, NF-kappaB1, c-Myc, ER-alpha, MEF-2, MEF-2A, deltaCREB
NPR3	P53, TBP, NF-kappaB, NF-kappaB1, c-Myc ER-alpha, MEF-2, MEF-2A, AhR, FOXD1, FOXI1, FOXJ2, FOXO1a, FOXO3, FOXO3a, FOXO3b, FOXO4
FARP1	STAT5, c-Fos, c-Jun, c-Myc, Elk-1, deltaCREB, FOXJ2
STAB2	NF-kappaB, NF-kappaB1, c-Fos, c-Jun, deltaCREB, FOXD1, FOXF2, FOXI1, FOXJ2, FOXO1a
NDEL1	P53, PPAR-gamma1, PPAR-gamma2, STAT5, c-Fos, c-Jun, c-Myc, deltaCREB, FOXD1, FOXF2, FOXJ2, FOXO1a, FOXO3, FOXO3a, FOXO3b, FOXO4

2.3.4 Pathway analysis

Pathway analysis were done using the Reactome database, an *in silico* pathway analysis tool. In total 19 pathways were identified that were associated with the seven target genes shown in Table 2.3. Significant pathways associated with these candidate target genes and Ovarian Cancer include RHO GTPase effectors, mitotic pathway and Hyaluronan uptake and degradation pathway. This is further evidence linking the target genes and indirectly linking the miRNAs to cancer, specifically OC.



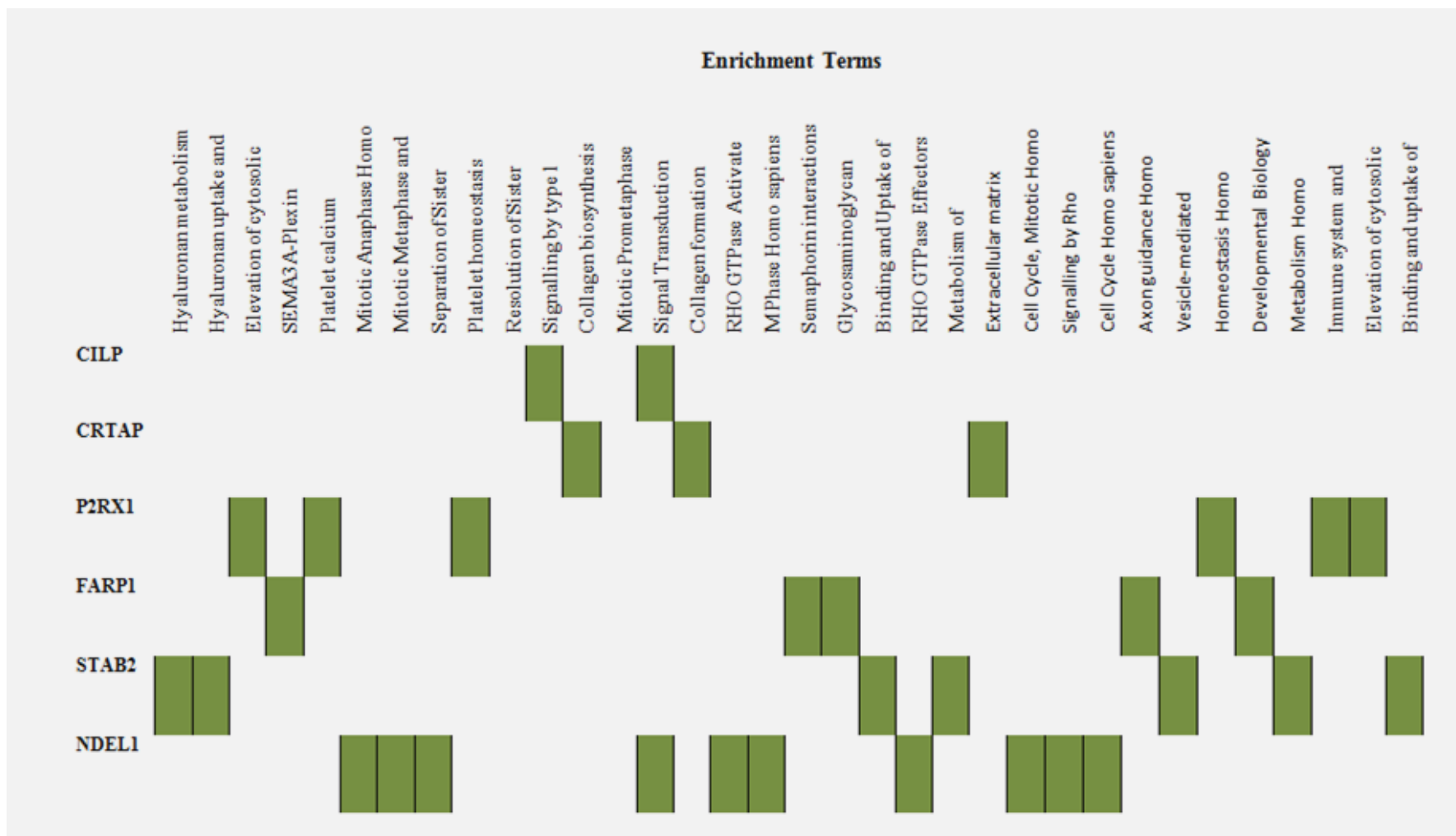


Figure 2.6: Pathways identified associated with the miRNA target genes. The enrichment terms are the columns, the input genes are the rows and the cells (green) in the matrix indicate if a gene is associated with an enrichment term.

2.4 Discussion

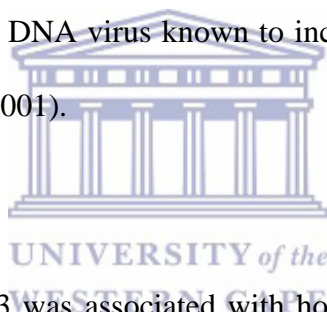
Ovarian Cancer remains one of the most common malignant gynaecologic cancers. The lethality and mortality of the disease is attributed to the fact that more than 85 % of patients are diagnosed in advanced stages (Chen *et al.*, 2013). Many proteins, genes and molecular processes are involved in cancer. These different factors work in conjunction with one another to perform various Biological Processes and Molecular Functions. Modification of these factors could lead to the development of a particular cancer phenotype including OC. The introduction of bioinformatics in the biomarker discovery process will aid in understanding diseases such as cancer in the preliminary stages of development as well as advanced stages; and fast-track the discovery process (Angel *et al.*, 2013)



2.4.1 Gene Ontology

In silico methods were employed to functionally characterize the miRNA target genes previously identified through functional genomics (Lottering, 2015). The ontologies explored consist of Molecular Functions, Biological Processes and Cellular Components. Gene Ontology represents important attributes of genes/proteins across all functionalities. Additionally, the data generated further promotes the understanding of a gene in a disease, based on their associated ontologies. This analysis might assist in the identification of novel diagnostic, prognostic and therapeutic considerations for cancer care (Huntley *et al.*, 2014; Ande *et al.*, 2007). Figure 2.3 shows the target genes clustered into various functional groups using the Gene Ontology Consortium database, as described in Section 2.2.1. Functional annotation using GO terms including Biological Processes and Molecular Functions showed that the gene targets of the miRNAs selected were implicated in processes associated with cancer. The target genes were found to be associated with protein binding, hormone binding,

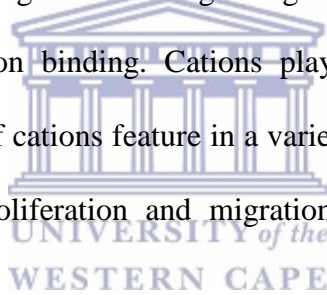
post-translational modification and cation activity shown in Figure 2.3. The transition from a normal cell to a tumour cell requires the deregulation of internal and external factors promoting cell survival, proliferation and tumorigenesis (Ande *et al.*, 2007; Smith, Steffen and Williams, 2003). An example of external factors includes ionized radiation. It causes gene mutations or chromosome aberration and it triggers steps involved in multistage carcinogenesis. An epidemiological survey of carcinogenesis for the atomic bomb victims of Hiroshima showed an increased incidence of leukemia and lung cancer within the population. The incidence of lung cancer is still high in Hiroshima and Nagasaki (Goodman *et a.*, 1994; Listwa, 2012). Another example is hepatitis B virus that results in chronic hepatitis and hepatic cirrhosis. It has been epidemiologically proven to be related to the occurrence of liver cancer. The hepatitis B virus is a DNA virus known to incorporate its DNA into hepatocyte genomes (Saeki and Sugimachi, 2001).



In Figure 2.3 it shows that NPR3 was associated with hormone binding. Hormone binding imbalances are an important risk factor for Ovarian Cancer initiation. Multiple hormonal conditions such as polycystic ovary syndrome and endometriosis associated with OC is directly correlated to hormonal imbalances (Daniilidis and Dinas, 2009; Harris and Terry, 2016). The disease is generally considered as a hormone-dependent cancer as the main function of the ovaries is the production of reproductive hormones namely oestrogen and progesterone. There are several hormonal hypothesis including the progesterone, oestrogen and insulin-like growth factor theories (Huang and Jan, 2014; Kaaks and Lukanova, 2005). Figure 2.3 shows that CILP is involved in the negative regulation of insulin-like growth factor. This growth factor emerges as a hormone directly involved in the pathogenesis of Ovarian Cancer (Huang and Jan, 2014). Studies suggest it has mitogenic and anti-apoptotic properties. Deregulation of the insulin-like growth factor binding proteins has been directly

related to various cancers including cancer of the breast and ovary (Lukanova, 2005; Bach and Yang, 2013; Huang and Jan, 2014). Furthermore, over expression of insulin-like growth factor can induce malignant transformation of ovarian epithelial cells (Lukanova *et al.*, 2002; Li *et al.*, 2017). These hormones play a central role in regulating cell proliferation, differentiation and apoptosis thus deregulation of insulin-like growth factor binding protein may allow mutated proto-oncogenes and tumour suppressor genes to survive (Kaaks and Lukanova, 2005; Ho, 2003; Bach, Fu and Yang, 2013).

Furthermore, Figure 2.3 indicates that NPR3, P2RX1 and STAB2 is associated with the Molecular Function, cation binding. This finding is significant as many diseases including cancer are associated with cation binding. Cations play a major role in many cellular processes thus the deregulation of cations feature in a variety of diseases including important processes in cancer such as proliferation and migration (Davis, Monteith and Roberts-Thomas, 2012).



2.4.2 Protein-protein network generation

Studies on protein interactions have become exceedingly important in an effort to understand human diseases on a system-wide level (Moore and Pattin, 2009). Protein-protein interactions are very important in understanding the functions of proteins and their behaviour, understanding biological processes that provides insight into the function of the proteins, we may assume ‘guilt by association’ for example proteins with unknown function may associate with a protein of known function therefore these proteins should cluster together in network maps; and proteins of unknown function (s) can be characterized based on these network

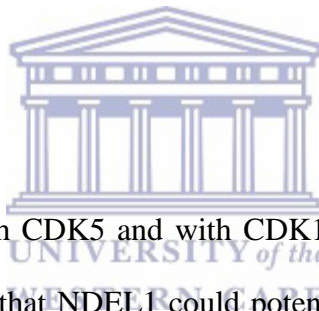
interactions produced. Additionally, it has been shown that proteins within an expression network regulate and support each other (Arga and Sevimoglu, 2014; De Las Rivas and Fontanillo, 2009). Two search tools were used to determine the interacting genes/proteins of our seven putative target genes.

2.4.2.1 STRING

STRING analysis illustrated in Figure 2.4 showed that 2 target genes (FARP1, P2RX1) are not in the expression network generated (Figure 2.4) based on the criteria used in STRING to establish the interacting network. The database might not contain enough information regarding the link between these proteins at the time of analysis. It has been reported that a set of genes for a particular disease including cancer may not directly interact with one another, thus they may not be placed within the same expression network or their intermediary protein may not yet been identified (Barabasi *et al.*, 2011; Barabasi and Oltvai, 2004). The target genes (CILP, NDEL1, NPR3, STAB2, CRTAP) that showed strong association with one another in the expression network (Figure 2.4) were associated with cell proliferation, mitotic cell cycle, the apoptotic pathway and cyclin-dependent protein serine/threonine kinase activity (Barabasi *et al.*, 2011).

One of the target genes, CILP, is associated with BMP-2 and FURIN (Figure 2.4). NPR3 is also associated with FURIN. Both proteins have been linked to multiple cancers including Ovarian Cancer. A study done by Le Page and associates (2009) observed up-regulation of BMP-2 in Ovarian Cancer cell lines compared to normal cell lines (Le Page *et al.*, 2009). They went further to investigate the effect of the over-expression on the tumorigenesis on the

disease by treating OC cell lines with BMP-2 and assaying it to test for cell signalling events, cell migration and cell adhesion. They found that BMP-2 induced mitogenic signalling through the activation of extracellular signal-regulated kinase-mitogen-activated protein kinase (Erk-MAPKs). Additionally, BMP-2 also induced the SMAD signalling pathway (Le Page *et al.*, 2009). Ovarian Cancer is often referred to as a signalling pathway disease because diverse signalling pathways are triggered during the initiation and progression process, depending on the histological subtype of the disease. Furthermore, BMP-2 increases motility and cell proliferation (Le Page *et al.*, 2009; Smolle *et al.*, 2013). Studies indicate that increased expression of BMP-2 as well as FURIN results in a poor prognostic outcome of Ovarian Cancer patients (Page *et al.*, 2007; Le Page *et al.*, 2009; Ma *et al.*, 2010; Jaaks and Bernasconi, 2017).



NDEL1 showed associations with CDK5 and with CDK1; and an indirect association with CDK6. The association suggests that NDEL1 could potentially be key regulators of the cell cycle and regulate tumour growth because CDK5, CDK1 and CDK6 are important protein kinases involved in the cell cycle (Figure 2.4) (Cho *et al.*, 2016; Xi *et al.*, 2015). One target gene (CRTAP) is not part of the main expression network but it is associated with LEPRE1 (also referred to as P3H1). LEPRE1 has not been experimentally linked to Ovarian Cancer but it has been associated with other cancers including breast cancer (Shah *et al.*, 2009). The result (Figure 2.4) also suggests that some of the miRNA target genes (P2RX1, FARP1) may not directly modulate cancer-related pathways and processes because these proteins are absent in expression network generated by the STRING database. The proteins (FARP1 and P2RX1) may interact with intermediary proteins in cancer thus catalyzing reactions in metabolic pathways, regulate important MFs and BPs on a transcriptional level or contribute

through structural assemblies without any direct interactions to the disease (Kulikowski, 2012).

2.4.2.2 GeneMANIA

According to Figure 2.5, all of the putative target genes are present in the two main expression networks produced by GeneMANIA. Therefore, it can be deduced that there is a strong probability that the candidate genes co-function in the same biological and molecular networks. Additionally, CRTAP associates with NDEL1 and FARP1; and NPR3 and CILP are linked. This contrasts with the expression network provided by the STRING database which showed that CILP and P2RX1 had no association with the other miRNA target genes. Also, the proteins weren't directly linked to one another (Figure 2.5). As mentioned above, the STRING database might not contain sufficient information regarding the association between these proteins (Barabasi *et al.*, 2011; Barabasi and Oltvai, 2004). The two databases also used different algorithms to predict the protein-protein interactions to generate a biological/expression network. STRING database uses experimental data, computational prediction methods and public text collections whereas GeneMANIA uses data from Gene Expression Omnibus, BioBRID, Pathway Commons, I2D as well as organism-specific functional genomics data sets (Bork *et al.*, 2005; Mostafavi *et al.*, 2008; Warde-Farley *et al.*, 2010).

According to Figure 2.5, P2XR1 is directly linked to TSHR. The expression of TSHR in cancer correlates with the state of differentiation of tumours, with loss of differentiation resulting in the loss of mRNA expression. In 2014, Gyftaki and associates showed that TSHR

is abundantly expressed in several other tissues apart from the thyroid (Gyftaki *et al.*, 2014). Additionally they studied the alterations of TSHR expression in OC on a protein level, where they found that the TSHR protein expression was significantly increased compared to normal ovarian tissues. Epidemiological studies have proposed that patients with over expressed TSHR have an 80 % increased risk of OC, thus implying that the deregulation of TSHR is a contributing factor to ovarian tumorigenesis (Huang *et al.*, 2016). A study done by Huang and associates (2016) demonstrated that the deregulation of TSHR in OC is involved in multiple signalling pathways important in the onset and progression of the disease. The study further showed that TSHR interacts with G alpha proteins, which in turn activates adenylate cyclase, activating the cAMP-dependent pathway. This crosstalk then induces activation of the epidermal growth factor receptor (EGFR), extracellular signal-regulated kinases (ERKs) and Protein kinase B (AKT) pathways which results in an increase in phosphorylation of the G alpha proteins. Additionally they showed that deregulation of TSHR is also linked to the activation of the phosphatidylinositol 3-kinase/Protein kinase B (PI3K-AKT) cascade/signalling pathway, which is known to coordinate a complex signalling network involved in cell proliferation in various cancer types (Huang *et al.*, 2016; Goel *et al.*, 2011). TSHR has also been reported to be involved in other proliferative pathways including Wnt and MAPK pathways (Goel *et al.*, 2011; Garcia-Jimenez and Santisteban, 2007).

Other candidate genes have also been linked to the initiation and progression of cancer including Ovarian Cancer such as TP53INP1, FOXH1 and ABCC4. It's suggested that genes that show similar expression patterns, generally are controlled by the same regulatory systems (Heyer *et al.*, 1999; Goel *et al.*, 2011).

2.4.3 Transcription factor analysis

Transcription factors are proteins that control the expression of genes by binding to specific DNA sequences within promoters and thereby control the transcription of the genes via up- or down-regulation of RNA polymerase or other regulatory proteins (Sirotkin, 2014). Understanding the transcription factors that directly and/or indirectly interact with a particular gene/protein one can confirm if the putative genes are connected to cancer through their regulatory network. It would also provide insight into cellular mechanisms and functions of these putative genes. Bajic *et al.*, (2010) hypothesized that transcription factors, especially those mediating the expression of disease causing genes, could be a key factor in understanding the genesis of the particular disease (Bajic *et al.*, 2010). Transcription factor analyses were performed using the Qiagen regulatory elements and epigenetic data from the GeneCards database and TfactS®. The relevant transcription factors identified were validated for their association with cancer especially OC through published literature as described in Section 2.2.3. A total of 89 transcription factors were identified for the seven candidate genes of which multiple transcription factors were implicated in OC.

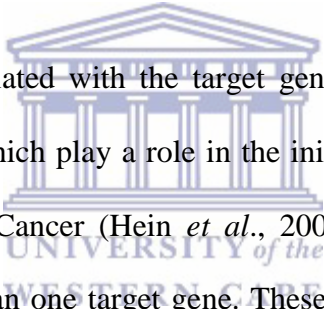
The relevant transcription factors identified, Table 2.2, have been reported in the initiation and progression of cancer; and cell growth and differentiation of Ovarian Cancer. One of these genes includes p53, a tumour suppressor that has been found to play a critical role in multiple cancers. Deregulation of this TF is one of the most frequent genetic lesions in human tumours as p53 plays a role in regulating cell cycle progression, DNA repair and cell death (Berchuck, 1994; Daemen *et al.*, 2005). This transcription factor shows association with two of the putative target genes namely NPR3 and NDEL.

Table 2.2 showed three of the seven genes (NPR3, STAB2 and NDEL1) to be associated with transcription factors of forkhead transcription factor (FOX) family. The superfamily controls a wide range of Biological Processes thus their loss of function can alter cell fate and promote tumorigenesis as well as cancer progression. The FOX family is evolutionary conserved and; play a critical role in the development and progression of cancer including Ovarian Cancer (Myatt and Lam, 2007). FOXO is a family member of the FOX family. It has been found to be involved in tumour suppression (Auguste, Cheaib and Leary, 2015). FOXO's are regulated in response to the activation of the PI3K/AKT signalling pathway. PI3K/AKT signalling pathway is deregulated in more than 70 % of ovarian cancers. Additionally, the pathway contributes to the development and tumouringensis of the disease (Auguste, Cheaib and Leary, 2015; Pavlidou and Vlahos, 2014).



The transcription factor c-Jun (associated with FARP1, STAB2, NDEL1) play an important role in the regulation of cell proliferation and progression, carcinogenesis, apoptosis and angiogenesis. The TF play an important role in the carcinogenesis and progression of several tumour types. According to Eckhoff *et al.*, (2013) c-Jun influence carcinogenesis and tumour progression of human ovarian carcinoma thus it plays a significant role in OC prognosis (Eckhoff *et al.*, 2013). The transcription factor, c-Myc was linked to CILP, P2RX1, NPR3, FARP1, NDEL1 (Table 2.2). It is a multifunctional nuclear phosphoprotein that plays a role in cell cycle progression, cell differentiation, apoptosis and cellular transformation. It is often over expressed in cancer as a result stimulating gene expression. The over-expression of c-myc are found in 37.3 % of all ovarian tumour tissues and in 63.5 % of serous adenocarcinoma tissues. However, c-myc over-expression is commonly found in advanced stage OC (Baker *et al.*, 1990; Reyes-Gonzalez *et al.*, 2015).

Table 2.2 shows c-Fos (a transcription factor associated with NDEL1, STAB2) which is a proto-oncogene and plays an important role in many cellular functions. It has been shown to be over-expressed in various cancers including Ovarian Cancer (Hein *et al.*, 2009). It plays a role in cell proliferation and differentiation of normal tissue as well as malignant transformation and tumour progression (Hein *et al.*, 2009). A study done by Mahner *et al.*, (2008) investigated the function of the c-Fos transcription factor in OC. It was determined that the loss of function of the TF is associated with tumour progression in ovarian carcinoma thus it plays a critical role in the prognosis of OC by changing the adhesion of ovarian tumour cells (Mahner *et al.*, 2008; Oliveira-Ferrer *et al.*, 2014).

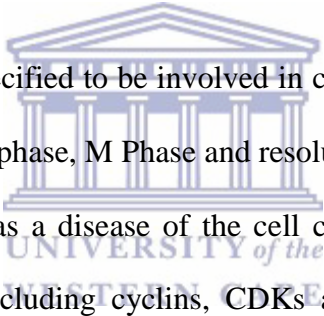


Other transcription factors associated with the target genes include AhR, PPAR-gamma1, deltaCREB and STAT5; all of which play a role in the initiation and progression of various cancer types including Ovarian Cancer (Hein *et al.*, 2009). Multiple transcription factors identified associate with more than one target gene. These shared transcription factors were annotated to regulate two of the candidate genes, both of which were not in the expression network in Figure 2.4 generated by STRING database. The association of these various transcription factors to the putative miRNA target genes further implicates the identified miRNAs and the target genes in cancer-related processes.

2.4.4 Pathway analysis

Determining the pathways biological molecules are involved in, enables characterization of the candidate target genes at a molecular level. Pathway analysis has become a powerful tool for understanding the biology of differentially expressed genes and proteins (Khatri *et al.*,

2012). Pathway analysis performed using Reactome as described in Section 2.2.4 and presented in Figure 2.6 represents the target genes with their associated pathways. From the pathway analysis result in Figure 2.6, some of the seven target genes showed no direct/specific link to Ovarian Cancer pathways. However, some of the genes are involved in RHO GTPase, immune responses, metabolism of carbohydrates and hyaluronan and extracellular matrix organization. These pathways have been linked to cancer specifically Ovarian Cancer. One of the target genes namely NPR3 were not associated with any pathways in the Reactome database. The database may not contain sufficient information regarding the pathways associated with these target genes (Khatri *et al.*, 2012).



From Figure 2.6, NDEL1 was specified to be involved in cell cycle related pathways namely mitotic prometaphase, mitotic anaphase, M Phase and resolution of sister chromatid cohesion. Cancer is frequently referred to as a disease of the cell cycle. Alteration of the regulatory mechanisms of the cell cycle including cyclins, CDKs and CDK inhibitors; resulting in uncontrolled cell proliferation which is a significant characteristic of human cancers (Bocicelli, D'Andrilli and Giordano, 2008; Kim and Nam, 2008). Most of the cell cycle regulatory genes play a role in Ovarian Cancer tumorigenesis and/or development. Thus deregulation of these pathways should enhance tumour growth, hence the probable role of the target gene NDEL in the onset/progression of Ovarian Cancer (D'Andrilli *et al.*, 2004; Bocicelli, D'Andrilli and Giordano, 2008).

The metabolic processes in cancer cells differ from those in normal cells. Activated oncogenes and loss of tumour suppressors', result in altered metabolism and an induction of aerobic glycolysis (Dang, 2012). There are metabolic adaptations observable including an

increased expression of all enzymes associated with cell division and DNA synthesis and an inability to transcribe genes coding for apoptotic proteins. More resources are used such as nitrogen and glucose; triggering angiogenesis and often stimulating a local anaerobic environment. More energy is consumed and normal cellular functions such as the urea cycle are not executed. Basically all of the normal cells metabolic processes are hijacked for cell division and the evasion of cell death (Dang, 2012).

According to Figure 2.6, STAB2 is involved in metabolic pathways including glycosaminoglycan metabolism and metabolism of carbohydrates. Glycoaminoglycans play an important role in the physiological and pathological conditions of cancer. They are key macromolecules that affect cell properties and functions, acting directly on cell receptors via interactions with growth factors (Annibaldi and Widmann, 2010). Therefore these macromolecules play an important role in cancer progression and treatment. Glucose is a major energy source thus abnormalities of carbohydrate metabolism may be important in the development and progression of cancer. The Warburg effect is the best studied metabolic process observed in cancer cells. In tumour cells, rather than utilizing glucose in the oxidative phosphorylation pathway, glucose is used for aerobic glycolysis thus conferring a selective growth advantage to the tumour cells (Annibaldi and Widmann, 2010; Holroyde and Reichard, 1981; Gevers and Levin, 1981). This suggests that STAB2 potentially play a role in sustaining cancer cells (Kizuka and Taniguchi, 2015; Afratis *et al.*, 2011).

The target gene, CRTAP, is involved in extracellular matrix organization shown in Figure 2.6. The local microenvironment of a cancer cell plays a significant role in its development. A major component of the tumour environment includes the extracellular matrix.

Reorganization of the extracellular matrix is associated with cancer development and progression of cancer (Mouw, Pickup and Weaver, 2014). Recent studies suggest that tumour cells play an essential role in extracellular matrix organization and remodelling. Although the extracellular matrix development is tightly controlled during embryonic development and organ homeostasis, the process of extracellular matrix organisation development is deregulated and disorganized in a disease such as cancer. Additionally, these cancer cell-derived extracellular matrix proteins enhance the survival and promote cell colonization at distant tissues and organs (Lu, Weaver and Werb, 2012; Xiong and Xu, 2016). Cell proliferation and invasion require the physical, biochemical and biomechanical properties elicited from the extracellular matrix. It has also been found that extracellular matrix proteins are highly expressed in cancer cells. These factors suggest that extracellular matrix organization is important in cancer growth, invasion and metastasis (Xiong and Xu, 2016). Furthermore, this suggests that CRTAP (Figure 2.6), being in the same class as the extracellular matrix proteins, could play a significant role in cancer. Thus it could serve as a useful biological marker for the disease.

These findings in combination with protein-protein interactions, transcriptional analysis and pathway analysis further implicates the identified candidate miRNAs and their target genes in the initiation and progression of cancer including Ovarian Cancer by virtue of the fact that they are miRNA regulated thus implicating the regulatory miRNAs.

2.5 Conclusion

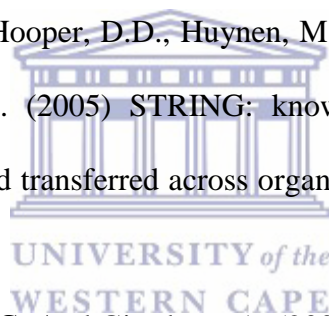
Signature biomarkers are urgently required for Ovarian Cancer diagnosis, prognosis and therapeutics to improve management of the disease. Currently, all the diagnostic tools for OC lack sensitivity, specificity and a positive predictive value; and are quite invasive. Therefore, focus has been shifted to biomarkers such as miRNAs. For the purpose of this study, a list of candidate target genes, previously identified were functionally characterized using *in silico* methodologies. The results observed suggest that the seven target genes are associated with various processes involved in the onset and progression of cancer including Ovarian Cancer. Furthermore, this subsequently links the miRNAs that target these genes to processes implicated in cancer. To further validate the genes, their prognostic value will be analysed using various bioinformatics tools. Also, the miRNAs will be molecularly validated in subsequent chapters.



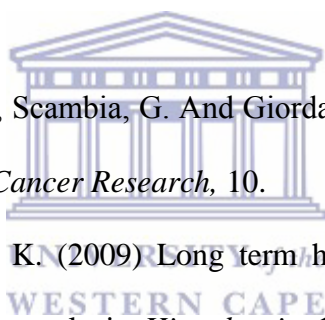
2.6 Reference List

1. Afratis, N., Gialeli, C., Nikitovic, D., Tsegenidis, T., Karousou, E., Theocharis, A.D., Pavao, M.S., Tzanakakis, G.N. and Karamanos, N.K. (2011) Glycoaminoglycans: key players in cancer cell biology and treatment. *The FEBS Journal*, 2.
2. Ande, S.R., Eshraghi, M., Los, M., Maddika, S., Manda, K.D., Panigrahi, S., Paranjothy, Weglarczyk, K., Wiechec, E. and Zuse, A. (2007) Cell survival, cell death and cell cycle pathways are interconnected: implications for cancer therapy. *Drug Resistance Update*, 10. P. 13-29.
3. Angel, P.M., Berenguer, A., Cordero, D., Crous-Bou, M., Moreno, V., Sanz-Pamplona, R. and Serra-Musach, S. (2012) Tools for protein-protein interaction network analysis in cancer research. *Clinical Translation Oncology*, 14. P. 3-14.
4. Annibaldi, A. and Widmann, C. (2010) Glucose metabolism in cancer cells. *Current opinion in Clinical nutrition and metabolic care*, 13. P. 466-470.
5. Arga, K.Y. and Sevimoglu, T. (2014) The role of protein interaction networks in systems biomedicine. *Computational and Structural Biotechnology Journal*, 11. P. 22-27.
6. Auguste, A., Cheaib, B. And Leary, A. (2015) The PI35/Akt/Mtor pathway in ovarian cancer: therapeutic opportunities and challenges. *Chinese Journal of Cancer*, 34. P. 4-16.
7. Baker, V. V., Borst, M. P., Dixon, D., Hatch, K. D., Shingleton, H. M. And Miller, D. (1990) c-myc amplification in ovarian cancer. *Gynecological Oncology*, 3. P. 340-342.
8. Bajic, V. B., Brendan, P. N. K., Choolani, M., Huak, C. Y., Keow, P. B., Kothandaraman, N., Razvi, K. And Salto-Tellez, M. (2010) E2F5 status significantly

- improves malignancy diagnosis of epithelial ovarian cancer. *BMC Cancer*, 10. P. 1-13.
9. Barabasi, A. L. and Oltvai, Z. N. (2004). Network biology: understanding the cell's functional organization. *Nature Reviews Genetics*, 5. P. 101-113.
10. Barabási, A. L., Gulbahce, N. and Loscalzo, J. (2011). Network medicine: a networkbased approach to human disease. *Nature Reviews Genetics*, 12. P. 56-68.
11. Berchuck, A. (1994) Overexpression of p53 is not a feature of benign and early-stage borderline epithelial ovarian tumours. *Gynecological Oncology*, 52. P. 232-236.
12. Boehm, J. S. and Hahn, W. C. (2011) Towards systemic functional characterization of cancer genomes. *Nature Review*, 12. P. 487-498.
13. Bork, P., Foglierini, M., Hooper, D.D., Huynen, M.A., Jensen, L.J., Krupp, M., Snel, B. And Von Mering, C. (2005) STRING: known and predicted protein-protein associations integrated and transferred across organisms. *Nucleic Acids Research*, 33. P. 433 – 437.
14. Bovicelli, A., D'Andrilli, G. And Giordano, A. (2008) Epithelial Ovarian Cancer: The Role of Cell Cycle Genes in the Different Histotypes. *Open Clinical Cancer Journal*, 7. P. 7-12.
15. Chen, X., Lui, J., Song, F. and Zheng, H. (2013) Advances in circulating microRNAs as diagnostic and prognostic markers for ovarian cancer. *Cancer, Biology and Medicine* 10, 123– 130. doi: 10.7497/j.issn.2095-3941.2013.03.001.
16. Cho, H., Chung, J.Y., Kang, E.S., Kim, J.H., Lee, E.J. and Yang, W. (2016) Accumulation of cytoplasmic Cdk1 is associated with cancer growth and survival rate in epithelial ovarian cancer. *Oncotarget*, 31. P. 1-12.



17. Coulibaly, I. and Page, G. P. (2008). Bioinformatic tools for inferring functional information from plant microarray data II: analysis beyond single gene. *International Journal of Plant Genomics*, 8.
18. Croft, D., O'Kelly, G., Wum G., Haw, R., Gillespie, M., Matthews, L., Caudy, M., Garapati, P., Gopinath, G., Jassal, B., Jupe, S., Kalatskaya, I., Mahajan, S., May, B., Ndegwa, N., Schmidt, E., Shamovsky, V., Yung, C., Birney, E., Hermjakob, H., D'Eustachio, P. and Stein, L. (2011) Reactome: a database of reactions, pathways and biological processes. *Nucleic Acid Research*, 39. P. 691-697.
19. Cui, H., Lan, X., Lu, S., Zhang, F. And Zhang, W. (2016) Bioinformatic prediction and functional characterization of human KIAA0100 gene. *Journal of Pharmaceutical Analysis*, 7. P. 10-18.
20. D'Andrilli, G., Kumar, C., Scambia, G. And Giordano, A. (2004) Cell cycle Genes in Ovarian Cancer. *Clinical Cancer Research*, 10.
21. Daniilidis, A. and Dinas, K. (2009) Long term health consequences of polycystic ovarian syndrome: a review analysis. *Hippokratia*, 13. P. 90-92.
22. Daemen, T., Lambeck, A., Nijman, H. W., van der Burg, S. H. and van der Zee, A, G, J. (2005) Immunologic aspects of ovarian cancer and p53 as tumour antigen. *Journal of Translational Medicine*, 3. P. 1-12.
23. Dalah, I., Stein, T.I. and Stelzer, G. (2011) In-silico human genomics with GeneCards. *Human Genomics*, 5. P. 709-717.
24. Davis, F. M., Monteith, G. R. And Roberts-Thomas, S. J. (2012) Calcium Channels and Pumps in Cancer: Changes and Consequences. *The Journal of Biological Chemistry*, 10. P. 1-9.

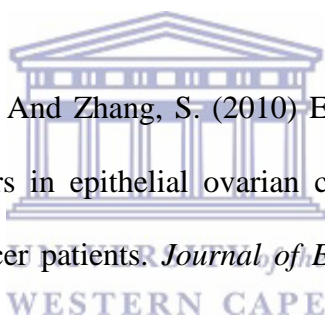


25. De Las Rivas, J. And Fontanillo, C. (2010) Protein-Protein Interaction Essentials: Key Concepts to Building and Analyzing Interactome Networks. *PLoS Computational Biology*, 6. P. 1-8.
26. Eckhoff, K., Flurschutz, R., Trillsch, F., Mahner, S., Milde-Langosch, K. (2013) The prognostic significance of Jun transcription factor in ovarian cancer. *Journal of cancer research and clinical oncology*, 139. P. 1673-1680.
27. Garcia-Jimenez, C. And Santixteban, P. (2007) TSH Signalling and Cancer. *Endocrinology and Metabolites*, 51. P. 654-671.
28. Gevers, W. And Levin, L. (1981) Metabolic alterations in cancer. Part I. Carbohydrate metabolism. *South African medical journal*, 4. P 518-521.
29. George, G.P. and Mittal, R.D. (2010) MicroRNAs: Potential Biomarkers in Cancer. *Indian Journal of Clinical Biochemistry*, 25. P. 4-14.
30. Goel, R., Raju, R., Maharudraiah, J., Kumar, G.S.S., Ghosh, K., Kumar, A., Lakshmi, T.P., Sharma, J., Sharma, R., Balakrishnan, L., Pan, A., Kandasamy, K., Christopher, R., Krishna, V., Mohan, S.S., Harsha, H.C., Mathur, P.P., Pandey, A. and Prasad, T.S.K. (2011) A signalling network of thyroid-stimulating hormone. *Proteomics and Bioinformatics*, 4. P. 238-241.
31. Goodman, M. T., Mabuchi, K., Morita, M., Soda, M., Ochikubo, T., Fukuhara, T., Ikeda, T and Terasaki, M. (1994) Cancer incidence in Hiroshima and Nagasaki, Japan, 1958-1987. *Elsevier*, 30. P. 801-807.
32. Goy, A., Li, S., Lim, H., Pecora, A., Sarojini, S., Suh, S., Tamir, A. and Zhang, S. (2012) Early Detection Biomarkers for Ovarian Cancer. *Journal of Oncology*. [Online] 2012 (11). p. 1 – 15. Available from: <http://www.hindawi.com/journals/jo/2012/709049/>. [Accessed 4/03/2015].

33. Gyftaki, R., Liacos, C., Politi, E., Lontos, M., Saltiki, K., Papageorgiou, T., Thomaskos, N., Haidopoulos, D., Rodolakis, A., Alevizaki, M., Bamias, A., Dimopoulos, A. (2014) Differential transcriptional and protein expression of thyroid-stimulating hormone receptor in ovarian carcinomas. *International Journal of Gynaecological Cancers*, 24. P. 851-856.
34. Harris, H. R. And Terry, K. L. (2016) Polycystic ovary syndrome and risk of endometrial, ovarian, and breast cancer: a systemic review. *Fertility Research and Practice*, 2. P. 1-9.
35. Hein, S., Mahner, S., Kanowski, C., Loning, T., Janicke, F. And Milde-Lanosch, K. (2009) Expression of Jun and Fos proteins in ovarian tumours of different malignant potential and in ovarian cancer cell lines. *Oncology Reports*, 22. P. 177-183.
36. Ho, S. (2003) Estrogen, Progesterone and Epithelial Ovarian Cancer. *Reproductive Biology and Endocrinology*, 1. P. 1-8.
37. Holroyde, C. P. and Reichard, G. A. (1981) Carbohydrate metabolism in cancer cachexia. *Cancer Treatment Reports*, 5. P. 55-59.
38. Hu, Z., Seiler, M., Tian, F. And Wang, Y. (2014) Functional characterization of breast cancer using pathway profiles. *BMC Medical Genomics*, 7. P. 1-13.
39. Huang, X. and Jan, L.Y. (2014) Targeting potassium channels in cancer. *The Journal of Cell Biology*, 206. P. 151-162.
40. Huang, W., Li, Z., Lin, T., Wang, S., Wu, F. And Luo, C. (2016) Thyrostimulin-TSHR signalling promotes the proliferation of NIH:OVCA9 ovarian cancer cells via trans-regulation of the EGFR pathway. *Scientific Reports*, 6. P. 1-13.
41. Huntley, R.P., Sawford, T., Martin, M. J. And O'Donovan, C. (2014) Understanding how and why the Gene Ontology and its annotations evolve: the GO within UniProt. *BMC Structural Biology*, 3. P. 1-10.

42. Jaaks, P. and Bernasconi, M. (2017) The proprotein convertase furin in tumour progression. *International Journal of Cancer*, 141. P. 654-663. DOI: 10.1002/ijc.30714.
43. Kaaks, R. And Lukanova, A. (2005) Endogenous Hormones and Ovarian Cancer: Epidemiology and Current Hypothesis. *Cancer Epidemiology, Biomarkers and Prevention*, 10. P. 98-103.
44. Kartha, R.V., Subramanian, S. and Sundarbose, K. (2013) MicroRNAs as Biomarkers in Cancer. *Diagnostics* 3, 84-104. doi:10.3390/diagnostics3010084.
45. Khatri, P., Sirota, M. and Butte, A. J. (2012). Ten years of pathway analysis: current approaches and outstanding challenges. *Plos Computational Biology*, 8. P. 1-10.
46. Kim, Y.T. and Nam, E.J. (2008) Alteration of the cell-cycle regulation in epithelial ovarian cancer. *International Journal of Gynaecological Cancer*, 18. P. 1169-1182.
47. Kizuka, Y. And Taniguchi, N. (2015) Glycans and cancer: role of N-glycans in cancer biomarker, progression and metastasis, and therapeutics. *Advances in Cancer Research*, 126. P. 11-51.
48. Kulikowski, M. (2012) Importance of Gene-Gene interactions shown in study. *NC State University*, 10. [Online] Available from: <https://news.ncsu.edu/2012/09/mkmackaypnas/>.
49. Le Page, C., Puiffe, M., Meunier, L., Zietarska, M., de Ladurantaye, M., Tonin, P. N., Provencher, D. And Mes-Masson, A. (2009) BMP-2 signalling in ovarian cancer and its association with poor prognosis. *Journal of Ovarian Cancer Research*, 2. P. 1-11.
50. Li, H., Batth, I. S., Qu, X., Xu, I., Song, N., Wang, R. And Lui, Y. (2017) IGF-IR signalling in epithelial to mesenchymal transition and targeting IGF-IR therapy: overview and new insights. *Molecular Cancer*, 16. doi: 10.1186/s12943-016-0576-5.

51. Listwa, D. (2012) Hiroshima and Nagasaki: The Long Term Health Effects. Columbia: Centre for Nuclear Studies.(Online) Available from: <https://k1project.columbia.edu/news/hiroshima-and-nagasaki>.
52. Lottering, S (2015) *Identification of miRNAs as biomarkers for early diagnosis of Ovarian Cancer: An in silico approach*. Honours thesis. University of the Western Cape.
53. Lukanova, A., Lundin, E., Toniolo, P., Micheli, A., Akimmedkhanov, A., Rinaldi, S., Mutt, P., Lenner, P., Biessy, C., Krogh, V., Zeleniuch-Jacquotte, A., Berrino, F., Hallmans, G., Riboli, E. And Kaaks, R. (2002) Circulating levels of insulin-like growth factor-I and risk of Ovarian Cancer. *International Journal of Cancer*, 101. P. 549-554.
54. Ma, Y., Ma, L., GuO, Q. And Zhang, S. (2010) Expression of bone morphogenetic protein-2 and its receptors in epithelial ovarian cancer and their influence on the prognosis of ovarian cancer patients. *Journal of Experimental and Clinical Cancer Research*, 29. P. 1-6.
55. Masters, M., McAteer, S. and Merlin, C. (2002) Tools for characterization of *Escherichia coli* Genes of Unknown Function. *Journal of Bacteriology*, 184. P. 4573-4581.
56. Missiaglia, E., Shepherd, C. J., Aladowicz, E., Olmos, D., Selfe, J., Pierron, G., Delattre, O., Walters, Z., Shipley, J. (2017) MicroRNA and gene co-expression networks characterize biological and clinical behaviour of rhabdomyosarcomas. *Cancer Letters*, 385. P. 251-260.
57. Mostafavi, S., Ray, D., Warde-Farley, D., Grouios, C. and Morris, Q. 2008. GeneMANIA: a real-time multiple association network integration algorithm for



- predicting gene function. *Genome Biology*, 9(S4). Accessed 21/10/2013. Available from <http://genemania.org/pdf/Mostafavi.pdf>.
58. Moore, J.H. and Pattin, K.A. (2009) Role of protein-protein interaction database in human genetics. *Expert Review of Proteomics*, 6. P. 647-659.
59. Mouw, J.K., Pickup, M. W. And Weaver, V. M. (2014) The extracellular matrix modulates the hallmarks of cancer. *EMBO reports*, 15. P. 1243-1253.
60. Movva, S. (2015) Understanding Cancer-The basics. WebMD. <http://www.webmd.com/cancer/understanding-cancer-basics>.
61. Myatt, S. S. and Lam, E. W. F. (2007) The emerging role of forkhead box (Fox) proteins in cancer. *Nature Reviews Cancer*, 7. P. 847-859.
62. Oliveira-Ferrer, L., Robler, K., Hausteil, V., Schroder, C., Wicklein, D., Maltseva, D., Khaustova, N., Samatov, T., Tonevitsky, A., Mahner, S., Janicke, F., Schumacher, U. And Milder-Langosch, K. (2014) c-Fos suppresses ovarian cancer progression by changing adhesion. *British Journal of Cancer*, 4. P. 753-763.
63. Page, R. E., Klein-Szanto, A. J., Litwin, S., Nicolas, E., Al-Jumaily, R., Alexander, P., Godwin, A. K., Ross, E. A., Schilder, R. J. and Bassi, D. E. (2007) Increased expression of the pro-protein convertase furin predicts decreased survival in Ovarian Cancer. *Cellular Oncology*, 29. P. 289-299.
64. Pavlidou, A. And Vlahos, N. F. (2014) Molecular alterations of PI3K/Akt/Mtor pathway: A therapeutic target in endometrial cancer. *The scientific word journal*, 10. P. 1-9.
65. Pengfei, L., Weaver, V. M. And Werb, Z. (2012) The extracellular matrix: A dynamic niche in cancer progression. *JCB: Review*, 196. P. 395-406.
66. Reyes-Gonzalez, J. M., Armaiz-Pena, G. N., Mangala, L. S., Valiyeva, F., Ivan, C., Pradeep, S., Echevarria-Vargas, I. M. Rivera-Reyes, A., Sood, A. K. and Vivas-Mejia,

- P. E. (2015) Targeting c-MYC in platinum-resistant ovarian cancer. *Molecular Cancer Therapy*, 10. P. 2260-2269.
67. Saeki, H. and Sugimachi, K. (2001) Carcinogenic risk factors. *Prevention and Early Detection of Cancer*, 44. P. 245-249.
68. Shah, R., Smith, P., Purdie, C., Quinlan, P., Baker, L., Aman, P., Thompson, A. M. And Crook, T. (2009) The prolyl 3-hydroxylases P3H2 and P3H3 are novel targets for epigenetic silencing in breast cancer. *British Journal of Cancer*, 100. P. 1687-1696.
69. Sirotkin, R. V. (2014) Transcription factors and ovarian functions. *Journal of Cellular Physiology*, 225. P. 20-26.
70. Smith, B., Steffen, S. and Williams, J. (2003) The Ontology of the Gene Ontology. *AMIA Annual Symposium Proceedings Archive*, 3. P. 609-613.
71. Smolle, E., Taucher, V., Pichler, M., Petru, E., Lax, S. and Haybaeck, J. (2013) Targeting signalling pathways in Epithelial Ovarian Cancer. *International Journal of Molecular Sciences*, 14. P. 9536-9555. doi:10.3390/ijms14059536.
72. Warde-Farley, D., Donaldson, S.L., Comes, O., Zuberi, K., Badrawi, r., Choa, P., Franz, M., Grouios, C., Kazi, F., Lopes, C.T., Maitland, A., Mostafavi, S., Montojo, J., Shao, Q., Wright, G., Bader, G.D. and Morris, Q. (2010) The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Research*, 38. P. 214-220.
73. Xi, Q., Huang, M., Wang, Y., Zhong, J., Liu, R., Xu., G., Jiang., L., Wang, J., Fang, Z. And Yang, S. (2015) The expression of CDK1 is associated with proliferation and can be a prognostic factor in epithelial ovarian cancer. *Tumour Biology*, 7. P. 4939-4948.
74. Xiong, G. And Xu, R. (2016) Function of cancer cell-derived extracellular matrix in tumour progression. *Journal of Cancer Metastasis and Treatment*, 2. P. 357-364.

Chapter 3

Prognostic studies of candidate genes

3.1 Introduction

3.1.1 Prognostic markers/factors in Ovarian Cancer

3.1.2 Kaplan-Meier plot

3.2 Aims and objectives

3.2.1 Aims

3.2.2 Objectives

3.3 Methodology

3.3.1 Kaplan-Meier Plotter database

3.3.2 SurvExpress database

3.3.3 PROGgene database

3.4 Results

3.4.1 Kaplan-Meier Plotter database

3.4.2 SurvExpress database

3.4.3 PROGgene database

3.5 Discussion

3.5.1 Kaplan-Meier Plotter database

3.5.2 SurvExpress database



3.5.3 PROGgene database

3.6 Conclusion

3.7 Reference List



Chapter 3

Prognostic studies of candidate genes

3.1 Introduction

Ovarian Cancer (OC) remains the most lethal gynaecological cancer, regardless of the advances in diagnosis and treatment, in part due to the advanced stage presentation in most patients (National Cancer Institute, 2016). OC has the lowest survival rate of all gynaecologic cancers. According to the National Cancer Institute (2016), approximately 1 in every 57 women contracts Ovarian Cancer and 60 % of these women already have advanced stage Ovarian Cancer. For effective cancer treatment, validated prognostic biomarkers are required to predict the course of the disease and the response to specific treatments. Currently, research focuses on identifying novel biomarkers for Ovarian Cancer diagnostics and prognostics. Previously, cancer antigen 125 (CA-125) was identified as a diagnostic biomarker but due to its limited specificity and sensitivity, less than half of cases are diagnosed successfully (Hu, Huang and Sood, 2010). Other biomarkers have been identified as prognostic markers such as glycoprotein human epididymis protein 4 (HE4), transthyretin, apolipoprotein A-1, beta2-microglobulin, transferrin and CA-125 but it presented sub-standard predictive values. Due to the inefficacious nature of the protein biomarkers identified, microRNAs (miRNAs) became the area of interest for research (Gao and Wu, 2015).

MiRNAs are a class of non-coding RNAs approximately 17-22 nucleotides in length (Gao and Wu, 2015). It regulates gene expression through post-translational alterations involving sequence-specific interactions with the 3' untranslated region (3' UTR) of the target messenger RNA (mRNA). MiRNAs have emerged as important regulators of cell

differentiation and proliferation processes. Deregulation of these molecules has been implicated in the pathogenesis of multiple cancers including Ovarian Cancer. MiRNAs can either be classified as oncogenic- or tumour suppressor miRNAs depending on their functions in regulating tumour phenotypes and they are important regulators in several facets of tumorigenesis including proliferation, metastasis and cancer cell functions (Shapira *et al.*, 2014; DiFeo, Joseph and Nagaraj, 2015). MiRNAs are classified as candidate biomarkers for diagnosis and prognosis in cancer because of their inimitable characteristics including (i) tissue-specific expression; (ii) stability in formalin-fixed tissues; and it's (iii) presence in bodily fluids. Due to its practicality for clinical settings, miRNA as prognostic entities in predicting survival outcome in Ovarian Cancer is evolving. Various technologies are available to explore the use of miRNAs as prognostic markers including *in silico* methods (bioinformatics) (Cramer and Elias, 2016) through understanding the prognostic significance of the target genes. Various bioinformatics tools are available to analyse the candidate target gene biomarkers, identified in Section 2, such as Kaplan- Meier Plotter, PROGgene and SurvExpress (Coticchia, Yang and Moses 2008; Goswami and Nakshatri, 2012).

3.1.1 Prognostic markers/factors in Ovarian Cancer

Despite years of research of new tumour biomarkers, the amount of clinically validated prognostic markers is pitifully small. There are many prognostic biomarkers investigated each year and published in cancer journals, however, few find a role in clinical practice. This is due to the complex nature of the biomarkers and the lack of acceptable standards for effectively evaluating and incorporating the newly identified biomarkers in the clinical setting (Mehta *et al.*, 2010). In oncology, prognostic biomarkers aim to predict a patient's outcome prior to treatment. Also, it is used to elicit an individual patient's risk of a future

outcome such as recurrence of the disease post-primary treatment. Prognostic markers play a significant role in clinical practice, distinguishing patients into different risk groups, treatment strategies and patient counselling especially in cancer (Riley, Sauerbrei and Altman, 2009).

Ovarian Cancer is characterised by unambiguous clinical and pathological features for ovarian carcinogenesis including: (i) pre-invasive and even-invasive lesions are difficult to detect, (ii) familial predisposition is significant for the heredity model of carcinogenesis; and (iii) morphological borderline ovarian tumours pose unanswered questions (Friedlander, 1998). Prognostic factors have been defined as phenotypes that correlate to the overall survival of a disease. Generally prognostic factors in Ovarian Cancer include intrinsic factors such as histological subtype, disease extent, age, grade, performance status and residual disease. Other factors that may impact the outcome include the treatment received by the patient and the effect of treatment on the tumour or the patient. Understanding the prognostic factors that result in a poor prognosis can potentially help individualise treatment for patients (Friedlander, 1998).

As previously mentioned, CA-125 is the gold standard marker and it's the most extensively studied molecular marker in Ovarian Cancer (Gupta and Christopher, 2009). CA-125 is expressed in more than 80 % of ovarian cancers and it correlates with the risk of malignancy, stage of disease and histology. The serum level of CA-125 is currently used to monitor response to chemotherapy, relapse and disease progression in OC patients (Gupta and Christopher, 2009). Multiple studies demonstrated that following chemotherapy, serum CA-125 level is a good predictor of overall and progression-free survival of Ovarian Cancer.

Furthermore, these studies suggested that patients with serum CA-125 values within the normal range post-chemotherapy had a significantly longer overall disease-free survival than the patients whose CA-125 levels remained high post chemotherapy (Gupta, 2009).

A potential prognostic biomarker in Ovarian Cancer is human kallikrein 8 (also referred to as KLK8). KLK8 is over expressed in the serum of 62 % of OC patients suggesting its prognostic significance. A study done by Borgono and associates (2006) demonstrated that patients with higher KLK8 mRNA levels are associated with lower grade disease, lower residual tumour left following surgery and they have a longer disease-free survival compared to those patients that have low expression levels of KLK8 based on the Cox regression model. This was confirmed by a Kaplan-Meier survival curve for progression-free survival (PFS) and overall survival (OS) for KLK8-positive and KLK8-negative patients (Figure 3.1). Notably, KLK8 is not expressed in normal ovarian tissue Therefore, it can be concluded that KLK8 is an independent biomarker of favourable prognosis in Ovarian Cancer (Borgono *et al.*, 2006; Magklara *et al.*, 2011).

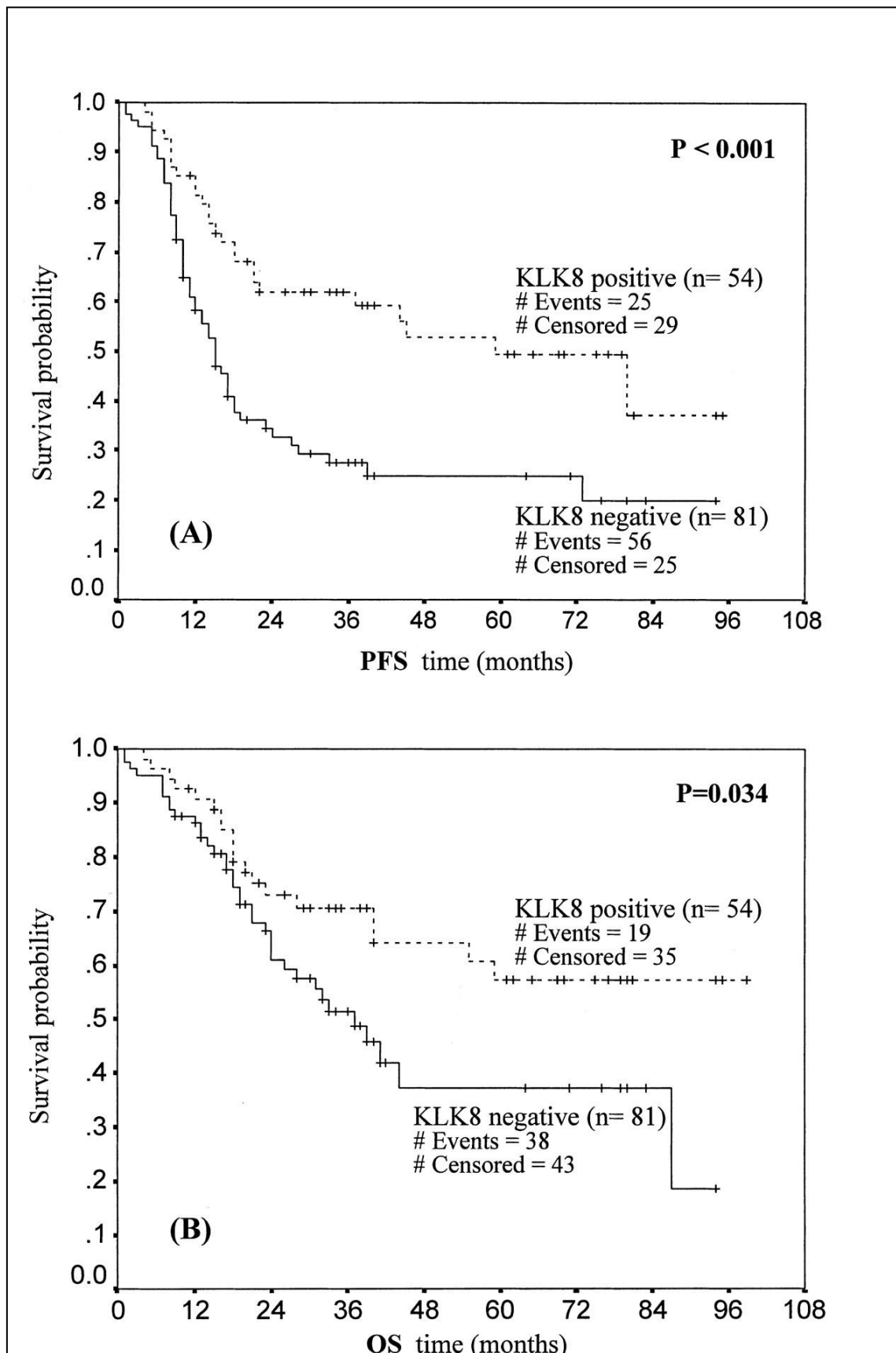
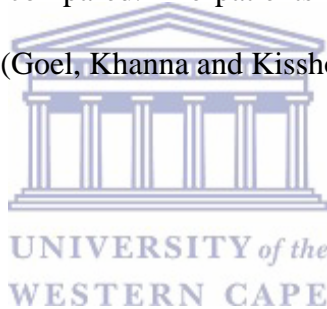


Figure 3.1: Kaplan-Meier plot presenting the association between the expression of KLK8 and Progression-free survival and Overall Survival (Magklara *et al.*, 2011).

3.1.2 Kaplan-Meier plot

The Kaplan-Meier estimator is one of the best options used to estimate empirical hazard, survival and cumulative distribution functions over a period of time. It is defined as a non-parametric statistic and it's one of the simplest ways to compute the survival over-time. The survival curves (Kaplan-Meier survival curves) take into account the patients that withdrew from the study, the subjects that may not experience the same event or die before the end of the study, labelled as right- censored observations, by calculating the occurrences of the event at a certain point in time and multiplying these successive probabilities to get the final estimate. This can be calculated for two groups of subjects and the statistical difference in the survival of these subjects can be compared. The patients are split into groups based on the parameters for the given scenario (Goel, Khanna and Kisshore, 2010).



3.2 Aims and Objectives

3.2.1 Aims

The aim of this chapter is to evaluate the prognostic value of the identified target genes as biomarkers using *in silico* methodologies.

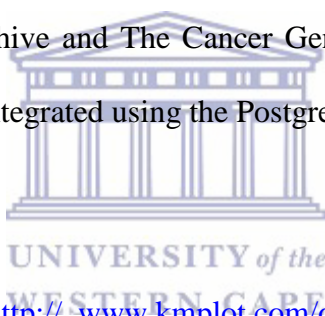
3.2.2 Objectives

- Prognostic analysis of miRNA target genes using Kaplan-Meier Plotter
- Prognostic validation of miRNA target genes using two datasets from SurvExpress
- Prognostic analysis of the miRNA target genes for Ovarian Cancer using two datasets from PROGgene

3.3 Methodology

3.3.1 Kaplan-Meier Plotter database

Kaplan-Meier Plotter (KM plotter), available at <http://www.kmplot.com> was used to determine the prognostic value of the target genes in the recurrence of Ovarian Cancer. The statistical tool has the ability to assess an individual gene or a combination of genes on the survival in breast, ovarian, lung, gastric, colon and prostate cancer patients (Szasz *et al.*, 2016). The KM plotter is capable of assessing the effect of 54 675 genes on survival using 10 461 cancer samples. Gene expression data and relapse free and overall survival information is downloaded from Gene Expression Omnibus (Affymetrix microarrays only), European Genome-phenome Archive and The Cancer Genome Atlas. Gene expression and clinical data are simultaneously integrated using the PostgreSQL server (Szasz *et al.*, 2016).



In the Google search engine, [http:// www.kmplot.com/ovar/](http://www.kmplot.com/ovar/) was launched. Each gene symbol for the target genes was used as an input into the gene space provided. The default settings were used and ‘draw Kaplan-Meier plot’ was selected. The output was downloaded in pdf format for further analysis.

3.3.2 SurvExpress database

SurvExpress, a online database (<http://bioinformatica.mty.itesm.mx/SurvExpress>) was implemented in JSP, JavaScript, MySQL and R. SurvExpress is a cancer-wide gene expression database with clinical outcomes providing survival analysis and risk assessment of cancer datasets. The database contains more than 20 000 samples and 130 datasets; covering

more than 20 tissues. The data is mainly obtained from Gene Expression Omnibus (GEO) and TCGA (Aquirre-Gamboa *et al.*, 2013). SurvExpress facilitate performance comparisons and validations of the prognostic biomarkers for cancer outcomes using various cancer datasets. SurvExpress output include a Kaplan-Meier plot for risk groups, clinical information available for risk groups, heat map representation of the gene expression values, a box plot across risk groups and tables with the summary of the Cox fitting and prognostic indices (Aquirre-Gamboa *et al.*, 2013).

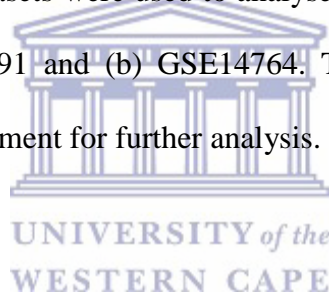
The seven target gene symbols were used as an input in the space provided for the gene list and 'ovarian' was selected as the tissue type. The genes were then analyzed using the SurvExpress dataset containing 784 samples and TCGA dataset containing 578 samples. The dataset contains meta-analysis clinical data and recurrence. The analysis button was clicked and the result summaries, Kaplan-meier plots, box plots of gene expression by risk groups and heat maps were exported in a pdf format for further analysis.

3.3.3 PROGgene database

PROGgene is a web-based application available at <http://www.compbio.iupui.edu/proggene> used for studying the prognostic implication of mRNA biomarkers in a variety of cancers. The tool generates prognostic (Kaplan-Meier, KM) plots for mRNA of interest using R library 'Survival' (Goswami and Nakshatri, 2013). The web application is created using a PHP5 and R Programming environment (v2.15.2), MySQL (v 5.0.95) server at the backend. The database compiles data from various repositories including Gene Expression Omnibus (GEO), EBI Array Express and The Cancer Genome Atlas (TCGA). PROGgene consists of

64 unique patient series accounting for approximately 11 800 samples profiled over a maximum of approximately 24 000 markers in 18 cancer types, providing the most comprehensive resource available for survival analysis. The web based application provides a list of datasets available for analysis of interest and it enables researchers to choose the most pertinent datasets for their study design (Goswami and Nakshatri, 2013).

The online cancer survival tool was accessed at <http://www.compbio.iupui.edu/proggene>. The seven gene symbols were used as queries in the space provided for the genes and ‘ovarian’ option was selected as the cancer type. For the survival measure ‘death’ was selected and the queries were submitted. Two datasets were used to analyse the prognostic value of the seven target genes namely (a) GSE9891 and (b) GSE14764. The output results (Kaplan-Meier plots) were saved in a Word document for further analysis.



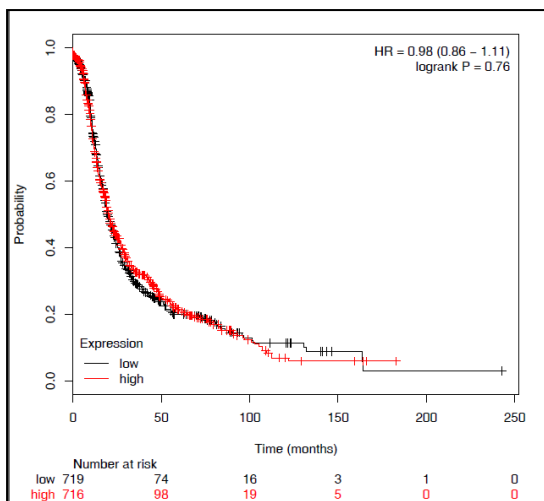
3.4 Results

3.4.1 Kaplan-Meier Plotter database

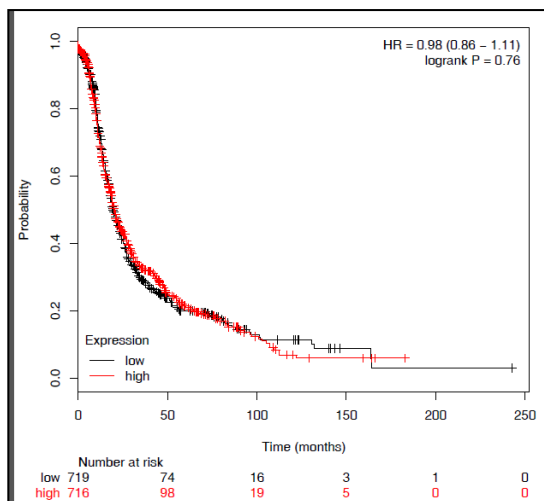
To determine the prognostic significance of the candidate target genes, Kaplan-Meier plotter was used as described in Section 3.3.1. High expression levels were significantly correlated to poor outcome of Ovarian Cancer for all the genes except for FARP1 illustrated in Figure 3.2. However, FARP1 seems to be a statistically significant prognostic marker for survival outcome of Ovarian Cancer (p-value=0.0015).



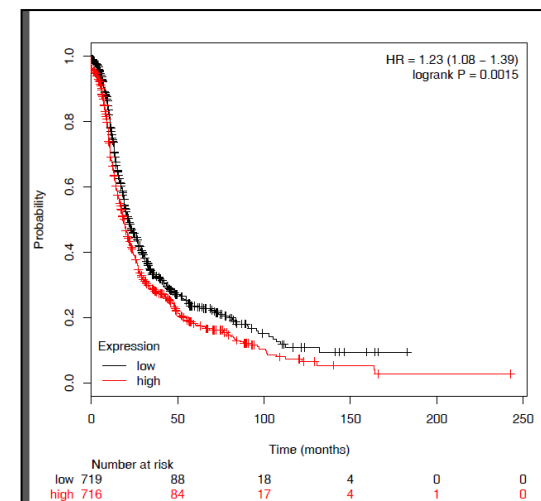
CILP



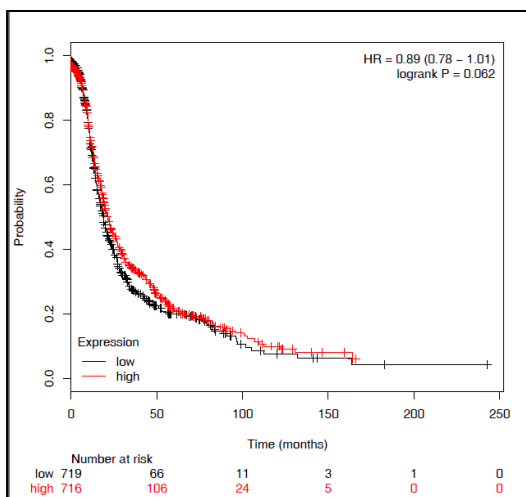
CRTAP



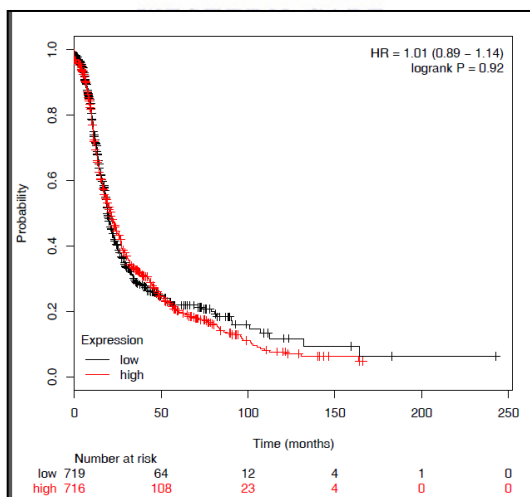
FARP1



NPR3



P2RX1



STAB2

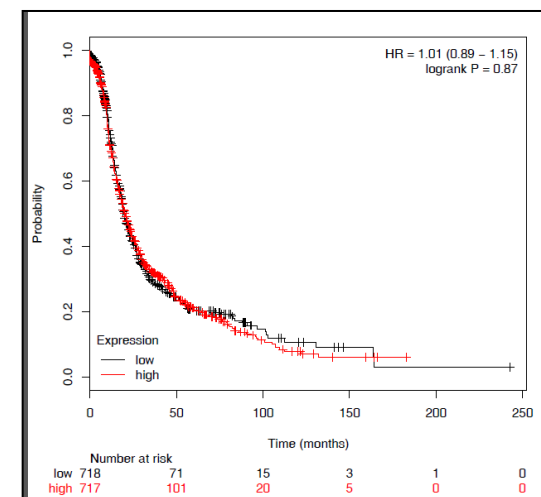


Figure 3.2: Survival curves for the target genes CILP, CRTAP, FARP1, NPR3, P2RX1 and STAB2 using the Kaplan-Meier Plotter database. Low risk is drawn in black and high risk is drawn in red. The p-value is shown in the right-hand corner for each gene

(p-value <0.05 is significant).

3.4.2 SurvExpress database

The prognostic significance of the candidate target genes were determined using two broad datasets in the SurvExpress database. The two broad datasets used were (i) Ovarian Meta-base comprising of 6 datasets consisting of 784 samples and (ii) TCGA comprising of 578 samples. From the results, one significant gene out of the seven target genes showed promise as a good prognostic marker based on the p-value ($p < 0.05$) in the SurvExpress dataset while in the TCGA dataset there were two significant genes (Figure 3.3). Figure 3.4 shows a Kaplan-Meier plot for risk groups, concordance index, and p-value of the log-rank testing equality of survival curves. In combination, the biomarkers seem like good prognostic markers for Ovarian Cancer based on the difference in higher- and lower risk groups for both datasets (SurvExpress dataset: p-value=0.001093; TCGA dataset: p-value=0.003799).

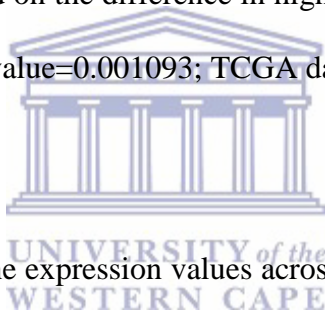


Figure 3.5 shows box plots of gene expression values across gene groups together with the p-value of the corresponding difference using a t-test. The results depicted in the SurvExpress dataset, illustrates that 3/7 target genes were differentially expressed and in the TCGA dataset shows that 5/7 target genes are differentially expressed. Figure 3.6 illustrates a heat map of the target gene (rows) expression values along samples (columns) in the risk groups. Low expression is represented in the green grades and high expression is represented in the red grades.

SurvExpress dataset

	Info	Summary.Result
Biomarker		GILP, CRTAP, FARP1, NDEL1, P2RX1
Database		Ovarian – Ovarian Meta-base: 6 cohorts 22K genes
Genes		5
CI		0.5726306
Log.Rank		0.000990728
Hazard.Ratio		1.380124
C.I.Hazard.Ratio		[1.13746283138607 – 1.67455226223512]
pHR		0.001092938
Correlation		0.02109884
Significant.Genes		1
Marginal.Genes		1
Cox.Interesting		GILP
DEG		0
DEG.Interesting		
Notes		Hazard Ratio was estimated by fitting a CoxPH using risk group as covariate.



	Info	Summary.Result
Biomarker		CRTAP, GILP, P2RX1, STAB2, FARP1, NDEL1, NPR3
Database		Ovarian – Ovarian serous cystadenocarcinoma TCGA
Genes		7
CI		0.5617529
Log.Rank		0.003606863
Hazard.Ratio		1.400143
C.I.Hazard.Ratio		[1.11478104753077 – 1.75855215443099]
pHR		0.003799237
Correlation		0.02780241
Significant.Genes		2
Marginal.Genes		2
Cox.Interesting		GILP, STAB2
DEG		0
DEG.Interesting		
Notes		Hazard Ratio was estimated by fitting a CoxPH using risk group as covariate.

Figure 3.3: The result summaries showing the significant target genes based on the p-values of the individual genes from the two datasets namely (a) SurvExpress dataset and (b) TCGA dataset using the SurvExpress database.

SurvExpress dataset

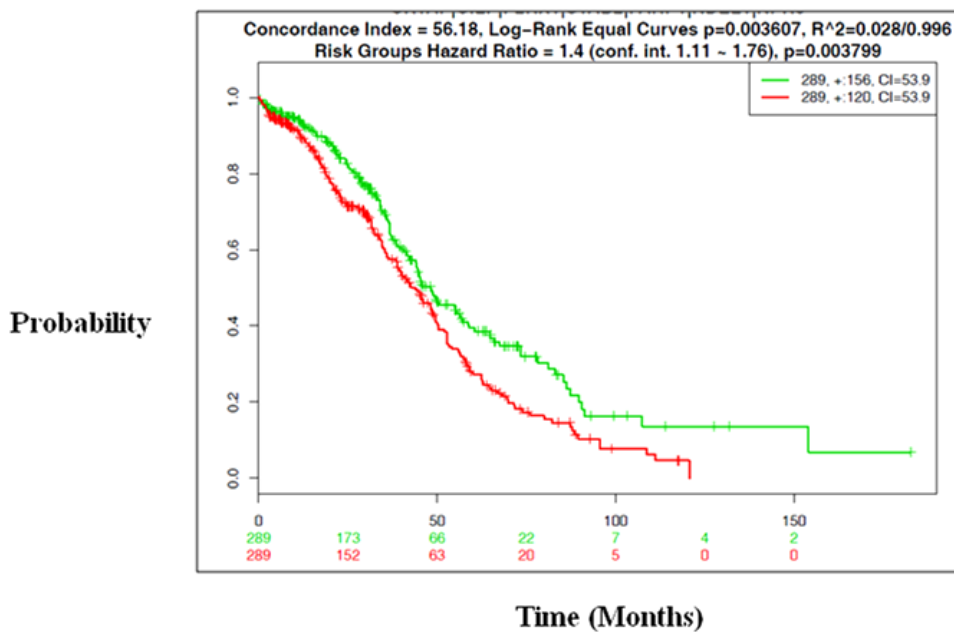
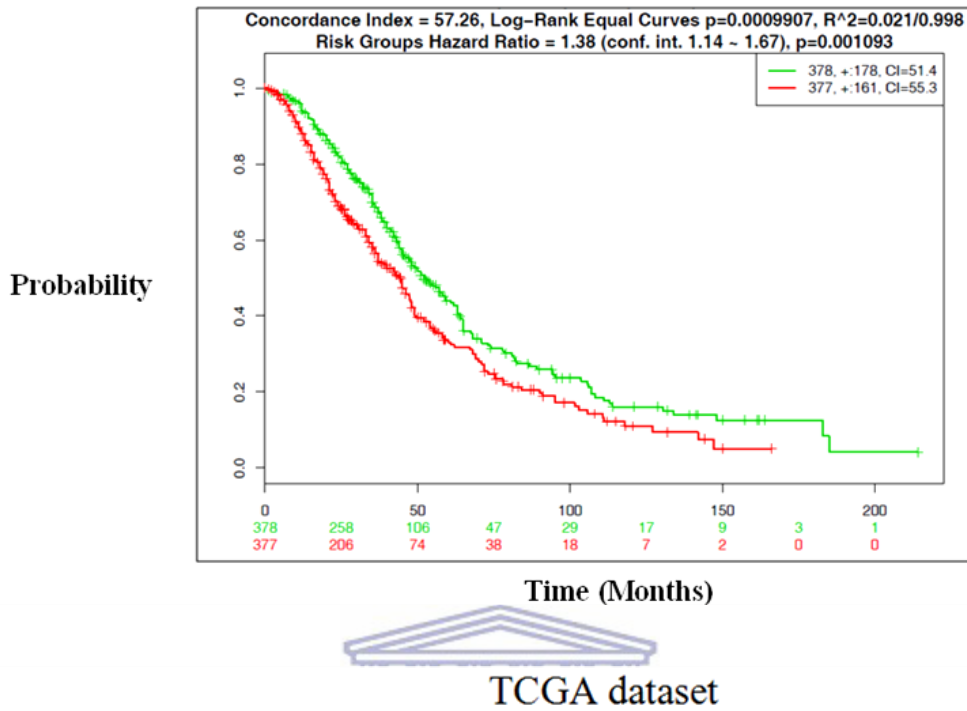
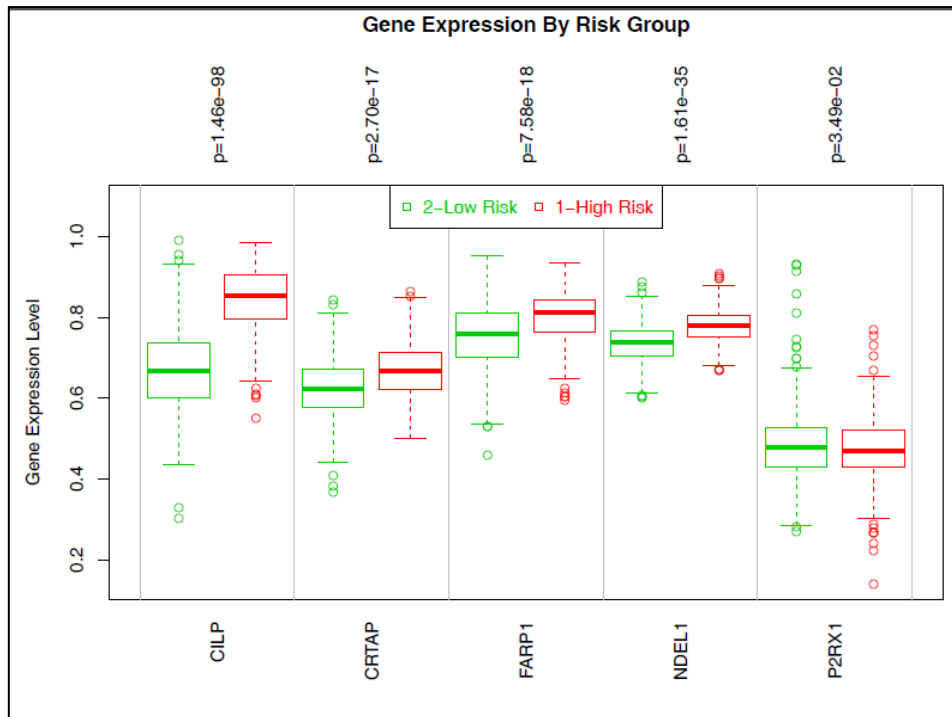


Figure 3.4: Kaplan-Meier analysis of the miRNA target genes for Ovarian Cancer prognostic outcome.

SurvExpress dataset



TCGA dataset

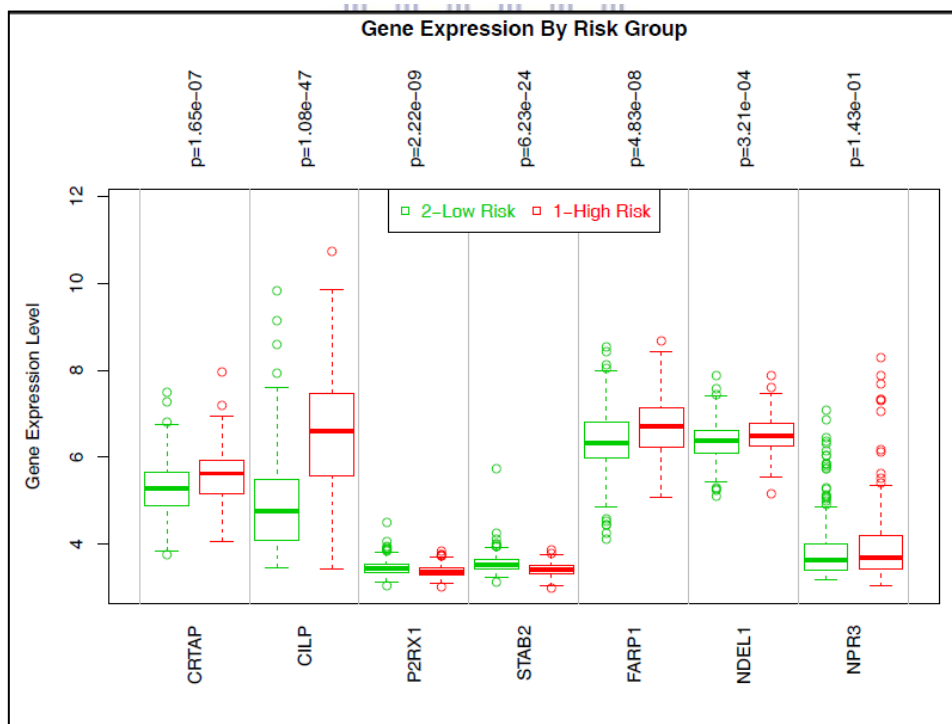
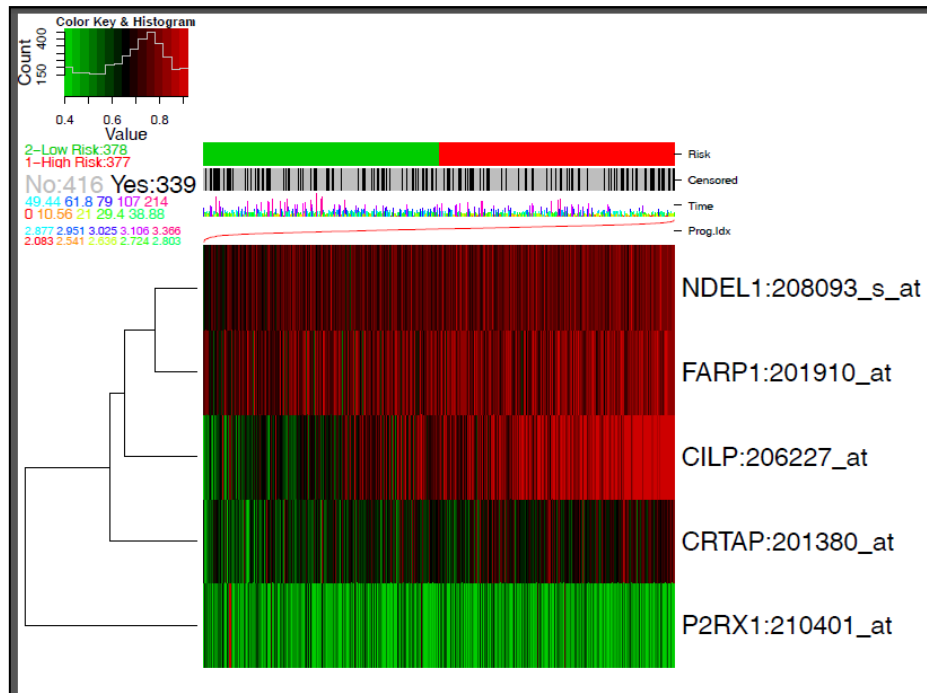


Figure 3.5: Gene expression values across gene groups represented by box plots together with the p-value of the corresponding difference comparing risk groups.

SurvExpress dataset



TCGA dataset

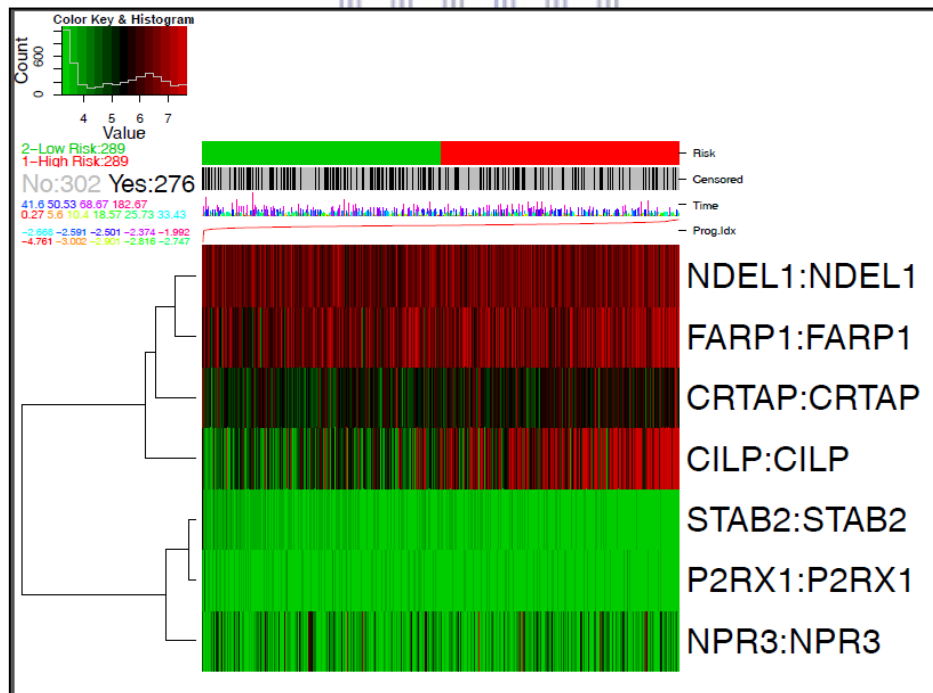


Figure 3.6: Expression profile and gene ranking based on prognostic index. (a) Expression profile from the dataset compiled by SurvExpress comprising 784 samples. (b) Expression profile from TCGA dataset comprised of 578 samples.

3.4.3 PROGgene database

The PROGgene database was used to determine the prognostic value of each target gene as described in Section 3.3.3. Low levels of expression of all the genes were correlated to poor prognostic outcome compared to high expression levels illustrated in Figure 3.7 and 3.8 for the dataset GSE9891 and; figure 3.9 and 3.10 for the dataset GSE14764.



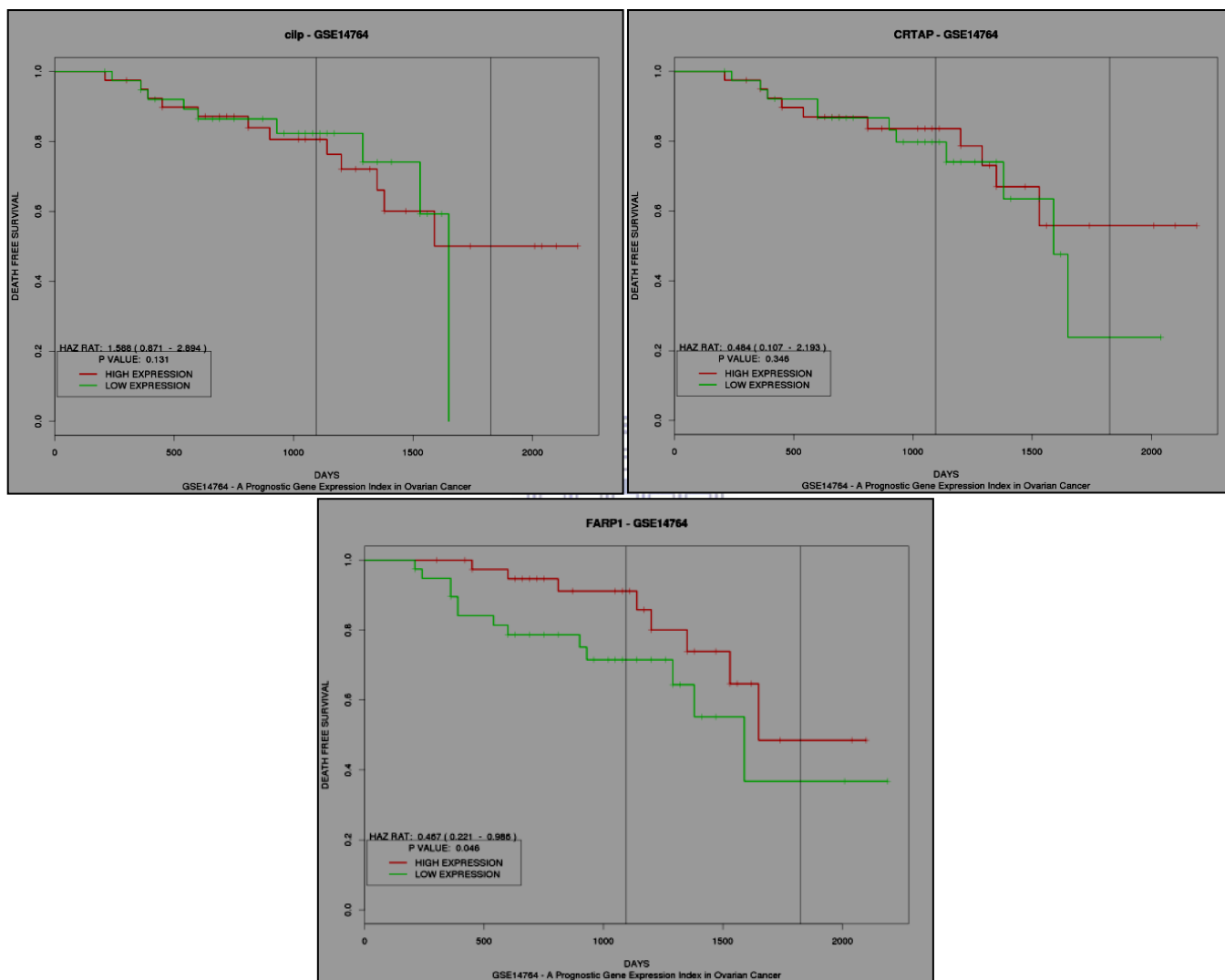


Figure 3.7: Survival curves for the target genes CILP, CRTAP and FARP using the GSE9891 dataset comprising of 285 ovarian tumour samples.

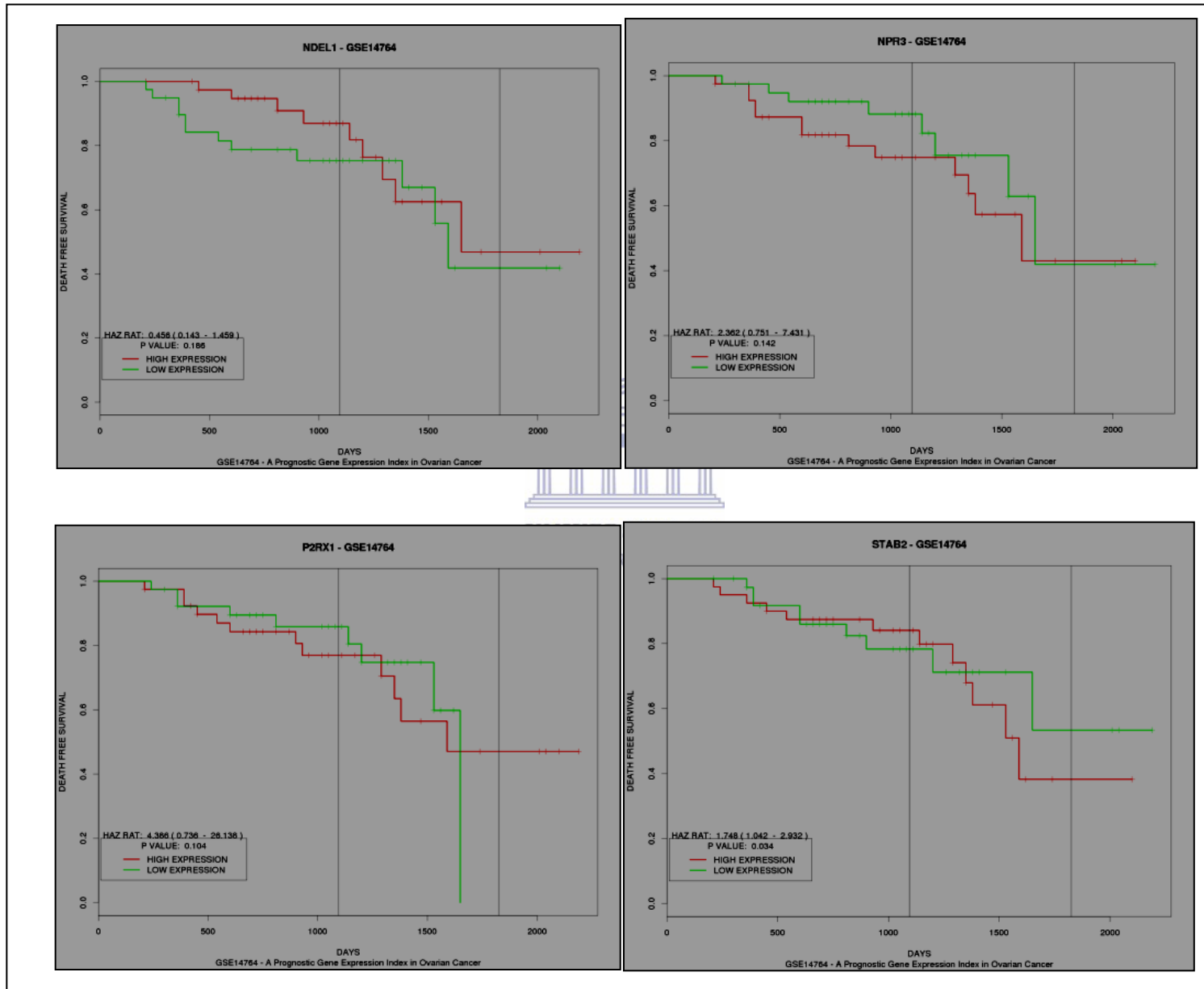


Figure 3.8: Survival curves for the target genes NDEL, NPR3, P2RX1 and STAB2 using the GSE9891 dataset comprising of 285 ovarian tumour samples. 121

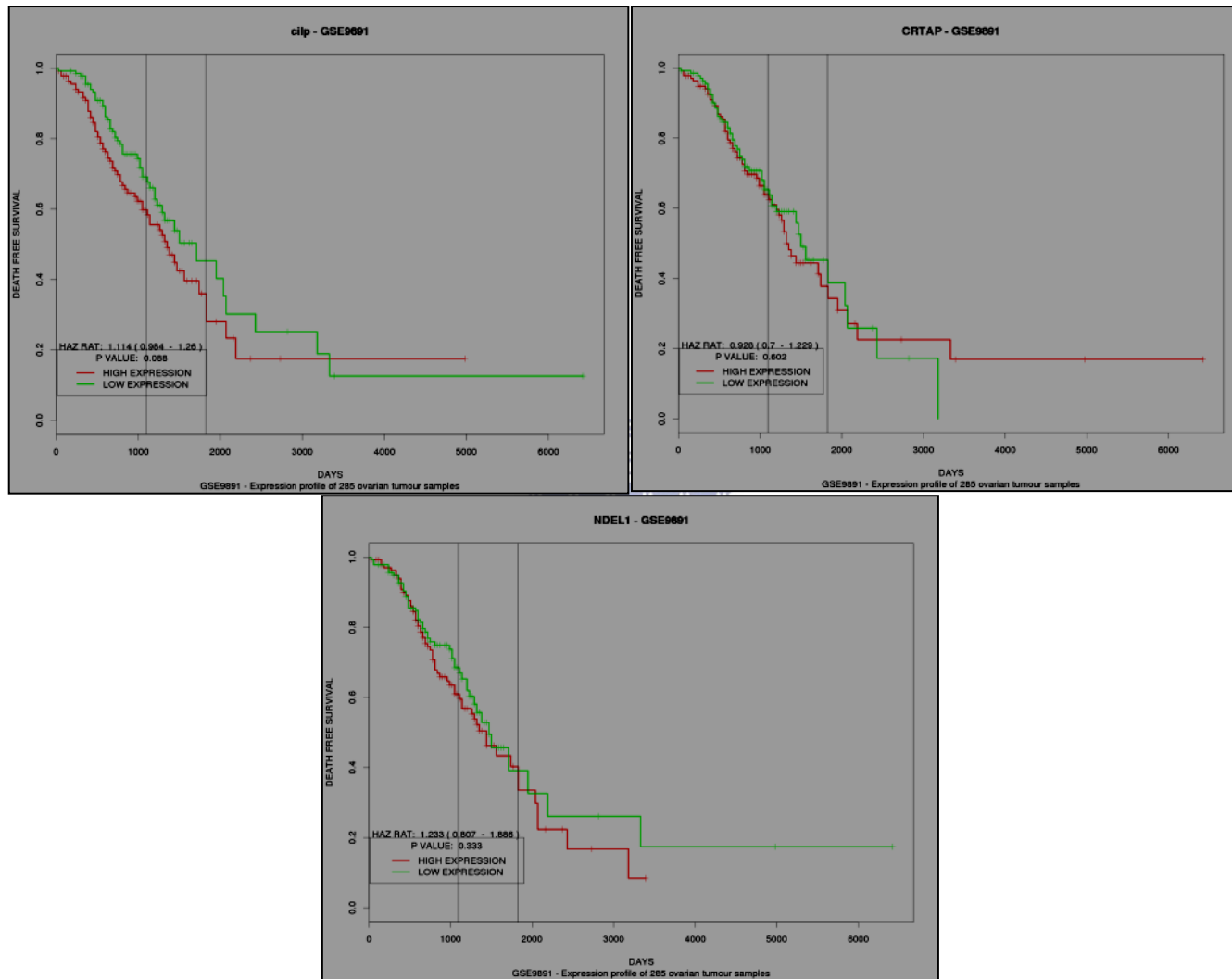


Figure 3.9: Survival curves for the target genes CILP, CRTAP and NDEL1 using the GSE14764 dataset. High expression is represented in red and low expression is represented in green.

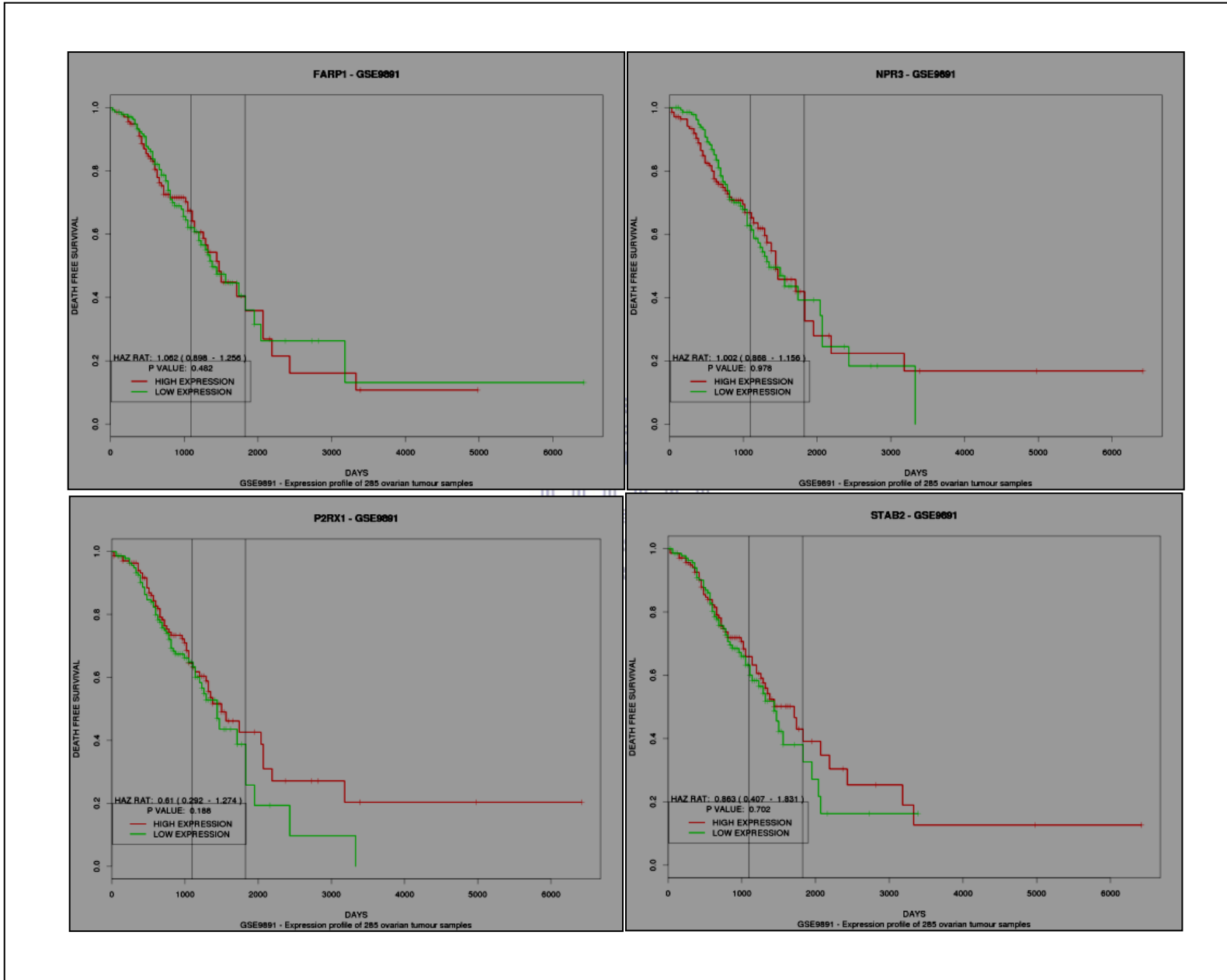


Figure 3.10: Survival curves for the target genes FARP1, NPR3, P2RX1 and STAB2 using the GSE14764 dataset. High expression is represented in red and low expression is represented in green.

3.5 Discussion

In oncology, prognosis plays a vital role in management and clinical decision making for a patient. Prognosis is presented as the time period between the start and the end of the clinical observation in combination with binary status information. Determining prognostic markers are important because it provides insight into the biology and natural history of a particular disease such as cancer (Chen *et al.*, 2014; Halabi and Owzar, 2010). In this study, the miRNA target genes were used to determine the prognostic or predictive value using various prognostic databases, as described in Section 3.3.

3.5.1 Kaplan-Meier plotter database

To determine the prognostic value of the individual target genes in the recurrence of Ovarian Cancer, the Kaplan-Meier plotter database was used. The database uses gene expression microarray data from Gene Expression Omnibus (GEO). Figure 3.2 showed the difference in the survival rate of risk groups based on the expression of the individual target genes. From the survival curves FARP1 (p-value = 0.0015) have a significant p-value in predicting the prognostic outcome based on the differential value of these biomarkers in Ovarian Cancer patients (Figure 3.2). A recent study by Schwaid and colleagues (2015) revealed that the phosphorylation of FARP1 contributes to cytoskeletal rearrangement/disorganization via Mitogen-Activated Protein Kinase Kinase Kinase Kinase 4 (MAP4K4). Furthermore, disorganized cytoskeletal architecture was accompanied by altered patterns of serine/threonine and tyrosine kinase phosphorylation as well as changed expression and enhanced focal adhesion (Crawford *et al.*, 2011; Schwaid *et al.*, 2015). Weichert and associates (2004) further reported that serine/threonine kinase (Polo-like kinases) correlates to poor prognosis of patients with Ovarian Cancer. Additionally, over expression of focal

adhesion kinases (FAK) have also been described as predictors of poor survival in OC patients. These kinases are over expressed in most invasive ovarian cancers and they play a pivotal role in OC progression and invasion (Sood *et al.*, 2004). Given the correlation, FARP1 may be a potential prognostic marker for Ovarian Cancer.

The target gene P2RX1 was not present in the database. The target gene might not be in the database because the database focuses on a particular Affymetrix gene expression platform even though there are more probe sets available to analyse additional genes. Another reason for its absence in the database could be because of the number of samples used in Kaplan-Meier plotter (Aguirre-Gamboa *et al.*, 2013; Chen, Sun and Hoshida, 2014).

3.5.2 SurvExpress database

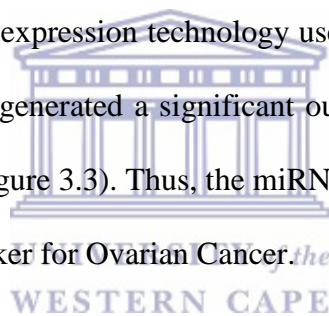


To further determine the prognostic/predictive value of the set of candidate target genes two broad datasets were used: Ovarian Meta-base comprising of 6 datasets consisting of 784 samples (SurvExpress dataset) and (ii) TCGA comprising of 578 samples (TCGA dataset). As shown in Figure 3.3, one out of the seven target genes were found to be significant as a prognostic marker based on its p-value ($p < 0.05$). The target gene that showed promise was CILP ($p < 0.001092936$). In the TCGA dataset two genes were deemed of prognostic significance based on their p-values (Figure 3.3). CILP is common in both datasets. Interestingly, Table 2.3 (Section 2.2.4) shows the target gene is associated with the insulin-like growth factors (IGF's). This hormone is directly involved in the pathogenesis of Ovarian Cancer and over expression is associated with poor prognosis of the disease (Kaaks and Lukanova, 2005; Beauchamp *et al.*, 2010).

A study done by Gu, Shigemasa and Ohama (2004) using 59 epithelial ovarian tumours (8 adenomas, 5 low malignant potential ovarian tumours and 46 adenocarcinomas and 7 normal ovarian tumours examined the expression of IGF-II in epithelial ovarian tumours and identified its' association with a patient's survival. They revealed that the over expression of IGF variants significantly correlated with poor prognosis of patients' (p value=0.0398). Other studies had similar conclusions particularly a study done by Sayer (2005) where they used microarray expression analysis to demonstrate over expression of the insulin-like growth factor is associated with advanced stage Epithelial Ovarian Cancer (EOC). Additionally, they measured IGF in 109 Epithelial Ovarian Cancers and 8 Normal Ovarian Surface Epithelial (NOSE) samples using quantitative real-time polymerase chain reaction (qRT-PCR). The group concluded that IGF expression is significantly higher in ovarian tumours compared to normal surface epithelium. Furthermore, based on multivariate analysis, IGF is an independent predictor of poor prognosis in patients with EOC (Gu *et al.*, 2004; Sayer *et al.*, 2005; Qian *et al.*, 2011). Additionally, CILP has been associated with two proteins (FURIN and BMP-2) that result in a poor prognostic outcome of Ovarian Cancer described in Section 2.4.2.1 (Figure 2.4). Taken together, evidence from this study and other studies supports the hypothesis that CILP might be a potential prognostic marker for Ovarian Cancer (Mahner *et al.*, 2013; Hein *et al.*, 2009; Oliveira-Ferrer *et al.*, 2014).

The survival curves in Figure 3.4 shows the plots by risk group, the log-rank test of differences between risk groups, the hazard-ratio estimate and the concordance index which estimates the probability that subjects with a higher risk prediction (red) and a lower risk prediction (green) will experience the event (death). As shown in Figure 3.3 when combining the miRNA target genes for prediction of prognostic outcome of Ovarian Cancer the prognostic value is more significant for SurvExpress than the TCGA datasets based on the p-

values. The p-value for the panel of biomarkers in the SurvExpress dataset is $p < 0.001093$ and for the TCGA dataset the p-value is 0.003799. In Section 2.3 and Section 2.4 (Chapter 2) the target genes were linked to various cancer-related Molecular Functions (MF's), Biological Processes (BP's), Transcription Factors (TF's) and pathways (Section 2). These different elements have been implicated in poor prognostic outcome of the disease such as variants in the cell cycle, c-Jun and c-Fos. This evidence further supports the hypothesis that the panel of miRNA target genes identified in a previous study (Lottering, 2015) can potentially be prognostic biomarkers for Ovarian Cancer. The Kaplan-Meier plots generated for these two datasets might be different because of the different approaches used when it was compiled. Also, it could be because of the difference in clinical information, populations, probe sets, sample size and gene expression technology used (Aguirre-Gamboa *et al.*, 2013). However, both of these datasets generated a significant output shown by the survival curve for the combined target genes (Figure 3.3). Thus, the miRNA target genes potentially serve as a good prognostic/predictive marker for Ovarian Cancer.



The SurvExpress database provides additional output results, the box plot and the heat map, which shows the expression levels for each of the candidate target genes (Figure 3.4 and Figure 3.5). In the box plot the gene expression of each gene is plotted along the risk groups obtained in the analysis using t-test. Additionally, it indicates whether the gene expression levels are different between the risk groups based on the visual difference. The high risk group is represented in red and the low risk group is represented in green. The heat map enables visualisation of the level of expression (by colour) of each gene and is ranked by their prognostic index (Aguirre-Gamboa *et al.*, 2013).

In the box plots the t-test is used to evaluate the difference between risk groups. In Figure 3.4 the box plot confirmed differential expression of most of the candidate target genes in both datasets. From the SurvExpress dataset three of the seven miRNA target genes were differentially expressed whilst in the TCGA dataset, five of the seven target genes were differentially expressed ($p < 0.05$) between the high and low risk groups.

In Figure 3.5, the heat map ranked the genes based on their prognostic ability. For both of the datasets, the main target genes that are highly differentially expressed in the high risk group ($p < 0.05$) are NDEL and FARP1. This is an indication that the high expression of these genes will result in poor prognosis in Ovarian Cancer patients. Interestingly, based on pathway analysis (Table 2.3) NDEL1 is involved in cell cycle related pathways such as the mitotic pathway. Deregulation of various factors in these pathways has been linked to poor prognosis of the disease (Andrilli, Giordano and Bovicelli, 2008). An example includes CDK1 which is well known to be associated with dysregulation of cancer causing processes/factors. CDK1 promotes the cell cycle alone and is essential for cell cycle progression and cell division. Yang and associates (2016) showed that cytoplasmic CDK1 ($p < 0.001$) is significantly elevated in EOC compared to normal cells and in the nucleus. Furthermore, based on the survival analysis done, over expression of CDK1 correlated with poor prognosis in 5-year overall survival (Log-rank: p -value = 0.028, hazard ratio = 2.016; 95% CI = 1.097 to 4.635) (Yang *et al.*, 2016; Xi *et al.*, 2015).

Additionally, in the SurvExpress dataset two of the seven candidate genes were not present in the dataset. The data might be missing from the

dataset because of the sample size and the patient group (Andrilli *et al.*, 2008; Sanders *et al.*, 2012). The differences in the outputs of these datasets could be due to Ovarian Cancer histological subtypes used prognostic factors such as age of the patients used, tumour size and pathological grade. As previously mentioned it could also be because of clinical information, populations, probe sets, sample size, gene expression technology and statistical methods used for analysis (Aguirre-Gamboa *et al.*, 2013).

3.5.3 PROGgene database

PROGgene database produces Kaplan-Meier plots for mRNAs using datasets for different cancers. Two datasets were used namely GSE9891 and GSE14764 (Figure 3.7- Figure 3.10). The high expression levels are represented in red whilst low expression levels are shown in green. In the GSE9891 dataset (Figure 3.7 and Figure 3.8) there were two significant target genes (the candidate target genes are STAB2 and FARP1) based on the p-value ($p < 0.05$). Interestingly, FARP1 was also rendered as a significant prognostic marker in Ovarian Cancer based on Figure 3.2 (Section 3.5.1). STAB2 hasn't been implicated as a significant prognostic biomarker in any of the other databases used however in Table 2.3 (Section 2.4.4) STAB2 is involved in various metabolic pathways suggesting the role of the miRNA target gene in sustaining cancer cells. A recent study done by Lamkin and associates (2009) investigated the hypothesis that higher pre-surgical glucose levels predict shorter disease-specific survival and recurrence time in OC patients. Based on univariate analysis as well as multivariate analysis, over expression of glucose correlates to poor survival times (p -value = 0.04) for individuals with the disease (Lamkin *et al.*, 2009).

Figure 3.9 and Figure 10 showed no target genes that were significant. However, it showed that two candidate target genes were marginally significant (p -value < 0.10). The contradictory results in the two datasets could be due to the limited sample size, patient cohorts, statistical methods employed and the database uses study specific prognostic plots instead of pooled prognostic plots. Also, the database focuses on transcriptomic profiling technology and thus gene expression profiles cannot be merged with this technology (Goswami and Nakshatri, 2013; Goswami and Nakshatri, 2014)

3.6 Conclusion

Currently the amount of clinically validated prognostic/ predictive biomarkers is pitifully small in cancer especially Ovarian Cancer. Prognostic biomarkers are important because they enables the prediction of a patients outcome or recurrence prior to or post clinical treatment. For the purpose of this study, a list of miRNA target genes, previously identified was used to determine their prognostic significance in Ovarian Cancer. Various bioinformatics tools were used to determine the significance. The results suggest that most of the seven target genes showed prognostic significance. Notably, FARP1 and CILP are common in multiple databases used in the study. These genes have been linked to various factors that result in poor prognosis of the disease as shown in Chapter 2. In the subsequent chapter, the miRNAs regulating the target genes will be molecularly validated using quantitative real-time PCR.

3.7 Reference List

1. Aguirre-Gamboa, R., Gomez-Rueda, H., Martinez-Ledesma, E., Martinez-Torteya, A., Chacolla-Huaring, R., Rodriguez-Barrientos, A., Tamez-Pena, J. G. And Trevino, V. (2013) SurvExpress: An Online Biomarker Validation Tool and Database for Cancer Gene Expression Data Using Survival Analysis. *Plos One*, 8.
2. Andrilli, G. D., Giordano, A. and Bovicelli, A. (2008) Epithelial Ovarian Cancer: The Role of Cell Cycle Genes in the Different Histotypes. *The Open Clinical Cancer Journal*, 2. P. 7-12.
3. Beauchamp, M., Yasmeen, A., Knafo, A. and Gotlieb, W. H. (2010) Targeting Insulin and Insulin-Like Growth Factor Pathways in Epithelial Ovarian Cancer. *Journal of Oncology*, 10. P. 1-11.
4. Borgono, C. A., Kishi, T., Scorilas, A., Harbeck, N., Dorn, J., Schmalfeldt, B., Schmitt, M. And Diamandis, E. P. (2006) Human Kallikrein 8 Protein Is a favourable Prognostic Marker in Ovarian Cancer. *Clinical Cancer Research*, 12. P. 1-8.
5. Crawford SM, Peace J. (2005) Does the nadir CA125 concentration predict a long term outcome after chemotherapy for carcinoma of the ovary? *Ann Oncol*, 16. P. 47–50.
6. Creek, A. L., Silkworth, W. T., Cimini, D., Jensen, R. V., Roberts, P. C. And Schmelz, E. A. (2011) Changes in Gene Expression and Cellulat Architecture an Ovarian Cancer Progression Model. *PLoS One*, 6. P,1-16.
7. Chen, X., Sun, X. and Hoshida, Y. (2014) Survival analysis tools in genomics research. *Human Genomics*, 8. P 1-5.
8. Coticchia, C. M., Yang, J. And Moses, M. A. (2008) Ovarian Cancer Biomarkers: Current Options and Future Promise. *Journal of National Comprehensive Cancer Network*, 8. P. 795-802.

9. Eckhoff, K., Flurschutz, R., Trillsch, F., Mahner, S., Milde-Langosch, K. (2013) The prognostic significance of Jun transcription factor in ovarian cancer. *Journal of Cancer Research and Clinical Oncology*, 139. P. 1673-1680.
10. Friedlander, M. L. (1998) Prognostic factors in ovarian cancer. *Seminars in Oncology*, 25. P. 305-314.
11. Goswami, C. P. and Nakshatri, H. (2013) PROGgene: gene expression based survival analysis web application for multiple cancers. *Journal of Clinical Bioinformatics*, 22. P. 1-9.
12. Goswami, C. P. and Nakshatri, H. (2014) PROGgeneV2: enhancements on the existing database. *Biomed Central*, 22. P. 1-6.
13. Gupta, D. (2009) Role of CA125 in predicting ovarian cancer survival-Areview of the epidemiological literature. *BioMed Central*, 2. P. 1-21.
14. Gupta, D. And Christopher, G. (2009) Role of CA125 in predicting ovarian cancer survival-Areview of the epidemiological literature. *BioMed Central*, 2. P. 1-20.
15. Gu, L., Shigemasa, K. and Ohama, K. (2004) Increased expression of IGF II mRNA-binding protein 1 mRNA is associated with an advanced clinical stage and poor prognosis in patients with ovarian cancer. *International Journal of Oncology*, 10. P. 671-678.
16. Halabi, S. and Owzar, K. (2010) The importance of identifying and validating prognostic factors in oncology. *Seminar Oncology*, 37. P. 9-18.
17. Hein, S., Mahner, S., Kanowski, C., Loning, T., Janicke, F. And Milde-Lanosch, K. (2009) Expression of Jun and Fos proteins in ovarian tumours of different malignant potential and in ovarian cancer cell lines. *Oncology Reports*, 22. P. 177-183.
18. Hu, W., Huang, J. And Sood, A. K. (2010) Prognostic Biomarkers in Ovarian Cancer. *Cancer Biomarkers*, 8. P. 231-251.

19. Lamkin, D. M., Spitz, D. R., Shahzad, M. M. K., Zimmerman, B., Lenihan, D. J., DeGeest, K., Lubaroff, D. M., Shinn, E. H., Sood, A. K. and Lutgendorf, S. K. (2010) Glucose as a prognostic factor in ovarian carcinoma. *Cancer*, 115. P. 1021-1027.
20. Lottering, S (2015) *Identification of miRNAs as biomarkers for early diagnosis of Ovarian Cancer: An in silico approach*. Honours thesis. University of the Western Cape.
21. Magklara, A., Scorilas, A., Katsaros, D., Massobrio, M., Danese, S. and Diamandis, E. P. (2001) The human KLK8 (Neurospain/ Ovasin) Gene: Identification of Two Novel Splice Variants and Its Prognostic Value in Ovarian Cancer. *Clinical Cancer Research*, 7. P. 806-811.
22. Mehta, S., Shelling, A., Muthukaruppan, A., Lasham, A., Blenkiron, C., Laking, G. And Print, C. (2010) Predictive and prognostic molecular markers for cancer medicine. *Therapeutic Advances in Medical Oncology*, 2. P. 125-148.
23. Oliveira-Ferrer, L., Robler, K., Haustein, V., Schroder, C., Wicklein, D., Maltseva, D., Khaustova, N., Samatov, T., Tonevitsky, A., Mahner, S., Janicke, F., Schumacher, U. And Milder-Langosch, K. (2014) c-Fos suppresses ovarian cancer progression by changing adhesion. *British Journal of Cancer*, 4. P. 753-763.
24. Riley, R. D., Sauerbrei, W. And Altman, D. G. (2009) Prognostic markers in cancer: the evolution of evolution from single studies to meta-analysis, and beyond. *British Journal of Cancer*, 100. P. 1219-1229.
25. Sayer, R. A., Lancaster, J. M., Pittman, J., Gray, J., Whitaker, R., Marks, J. R. And Berchuck, A., (2005) High insulin-like growth factor-2 (IGF-2) gene expression is an independent predictor of poor survival for patients with advanced stage serous epithelial ovarian cancer. *Gynecologic Oncology*, 96. P. 355-361.

26. Schwaid, A. G., Chunyan, S., Loos, P., Wu, J., Nguyen, C., Stone, K. L., Kanyo, J., Geoghagan, K. F., Bhattacharya, S. M., Dow, R. L., Buckbinder, L. And Carpino, P. A. (2015) MAP4K4 Is a Threonine Kinase That Phosphorylates FARP1. *American Chemical Society*, 10. P. 2667-2671.
27. Shapira, I., Oswald, M., Lovecchio, J., Khalili, H., Menzin, A., Whyte, J., Dos Santos, L., Liang, S., Bhiuya, T., Keogh, M., Mason, C., Sultan, K., Budman, D., Gregersen, P. K. and Lee, A. T. (2014) Circulating biomarkers for detection of ovarian cancer and predicting cancer outcomes. *British Journal of Cancer*, 110. P. 976-983.
28. Sood, A. K., Coffin, J. E., Schneider, G. B., Fletcher, M. S., DeYoung, B. R., Gruman, D. M., Gershenson, D. M., Schaller, M. D. And Hendrix, M. J. (2004) Biological significance of focal adhesion kinase in ovarian cancer: role in malignant and invasion. *The America Journal of Pathology*, 4. P. 1087-1095.
29. Szasz, A. M., Lanczky, A., Nagy, A., Forster, S., Hark, K., Green, J. E., Boussioutas, A., Busuttil, R., Szabo, A., Gyoryffy, B. (2016) Cross-validation of survival associated biomarkers in gastric cancer using transcriptomic data of 1065 patients. *Oncotarget*, 2. 322-333.
30. Qian, B., Katsaros, D., Lu, L., Canuto, E. M., Benedetto, C., Beeghly-Fadiel, A. and Yu, H. (2011) IGF-II promoter specific methylation and expression in epithelial ovarian cancer and their associations with disease characteristics. *Oncology Reports*, 25. P. 203-213.
31. Weroha, S. J. And Haluska, P. (2012) IGF system on Cancer. *Endocrinology and metabolism clinics of North America*, 41. P. 335-350.
32. Weichert, W., Denkert, C., Schmidt, M., Gekeler, V., Wolf, G., Kobel, M., Dietel, M. And Hauptmann, S. (2004) Polo-like kinase isoform expression is a a prognostic factor in ovarian carcinoma. *British Journal of Cancer*, 4. P. 815-821.

33. Xi, Q., Huang, M., Wang, Y., Zhong, J., Liu, R., Xu, G., Jiang, L., Wang, J., Fang, Z. And Yang, S. (2015) The expression of CDK1 is associated with proliferation and can be a prognostic factor in epithelial ovarian cancer. *Tumour Biology*, 36. P. 4939-4948.
34. Yang, W., Cho, H., Shin, H., Chung, J., Kang, E. S., Lee, E. And Kim, J. H. (2016) Accumulation of cytoplasmic CDK1 is associated with cancer growth and survival rate in epithelial ovarian cancer. *Oncotarget*, 31. P. 49481-49497.



Chapter 4

Molecular validation of identified miRNAs as biomarkers for early diagnosis of Ovarian Cancer (OC)

4.1 Introduction

4.1.1 Quantitative real-time PCR (qPCR)

4.1.2 Quantitative Strategies in qRT-PCR

4.2 Molecular methodologies

4.2.1 Cell culture

4.2.1.1 Thawing of cells

4.2.1.2 Trypsinization

4.2.1.3 Freezing down of cells

4.2.2 Total RNA extraction

4.2.3 Primer design

4.2.4 Poly (A) tailing and Reverse transcription

4.2.5 The validation of miRNA expression levels in Ovarian Cancer and other cell lines using qRT-PCR

4.3 Results

4.3.1 Analysis of amplification curves

4.3.2 Melting curve analysis



4.3.3 Analysis of qRT-PCR expression profiling data

4.4 Discussion

4.4.1 Analysis of Amplification curve and Melting curve

4.4.2 Expression profiling plot analysis

4.5 Conclusion

4.6 Reference List



Chapter 4

Molecular validation of identified miRNAs as biomarkers for early diagnosis of Ovarian Cancer (OC)

4.1 Introduction

MicroRNAs (miRNAs) are a class of short (~22 nucleotides long), single-stranded evolutionary conserved RNA molecules responsible for post-translational modification, either by translational repression or mRNA degradation (Jansson and Lund, 2012). The importance of miRNAs has been long dismissed until its initial discovery in *Caenorhabditis elegans* (Lima and Pasquinelli, 2014). Following its initial discovery, a vast number of miRNAs have been identified through computational and molecular studies in plants, animals and viruses (Jansson and Lund, 2012). Currently, miRNAs have been recognized to play a pivotal role in cell proliferation, apoptosis, neuronal cell fate, stem cell division and metabolism. It has been shown that miRNAs are differentially expressed and play a role in the pathogenesis of various diseases ranging from autoimmune diseases to cancer. The significance of these molecules in diseases such as Ovarian Cancer has prompted research into their diagnostic, prognostic and theranostic value (Ardekani and Naeini, 2010).

The current diagnostic methods for early detection of Ovarian Cancer include pelvic examination, transvaginal ultrasonography, imaging studies and CA-125 (cancer antigen-125 or carbohydrate antigen 125) testing (KK Women's and Children's Hospital, 2014). CA-125 is expressed in most OC tumours but it's also expressed in the normal epithelium of the female reproductive system and in the presence of other benign conditions such as Urinary Tract Infections (UTI), thus its non-specific. The addition of new tumour markers such as

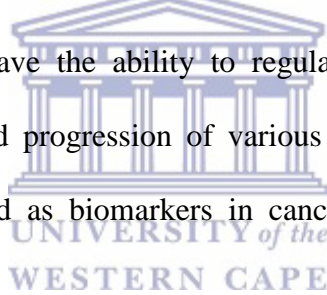
Human Epididymis Protein 4 (HE4) and mesothelin combined with CA-125, have presented an increased specificity and sensitivity (Husseinzadeh, 2011). Studies done by Bast *et al* (2004) and Farias-Eisner *et al* (2005) showed that combining biomarkers, resulted in an improved detection for early stage OC and the specificity and sensitivity increased from 65 % to 74 % and 52 % to 63 % respectively.

The OVA1 blood test, cleared by the FDA, is designed to be a highly sensitive tool (Bristow *et al.*, 2014). It tests for all types of Ovarian Cancers and determines if the mass is low risk Ovarian Cancer or high risk Ovarian Cancer. The OVA1 blood test measures the level of CA-125 in the blood in combination with four other proteins namely beta-2 microglobulin, transferrin, apolipoprotein A1 and transthyretin (prealbumin). The OVA1 blood test has an approximate sensitivity of 96 %, specificity of 35 % and a positive predictive value of 40 %. Regardless of the increased specificity, sensitivity and positive predictive value, the test is still not performed on its own. Another new biomarker for OC includes Risk of Ovarian Malignancy Algorithm (ROMA), approved by the FDA in 2011. The test is used to estimate the risk of OC in women that present a pelvic mass. The ROMA test is preferred over the OVA1 blood test because of its specificity and its cost-effectiveness. The specificity of the ROMA test is 92 % and the sensitivity is 76 % (Li, 2012).

These tests are used in addition to, not in place of, other diagnostic tools (Bristow *et al.*, 2014). However, the current biomarkers are mostly based on proteins and the major limitations of these proteins include low specificity, sensitivity and positive predictive value. Also, these proteins tend to degrade rapidly; therefore, tumour-specific molecules such as

miRNAs are under investigation for early detection, prognosis and therapy of Ovarian Cancer because of the miscellaneous nature of the disease (Kurman and Shih, 2010).

MicroRNAs are endogenous, naturally abundant, relatively stable non-coding RNA molecules and it is responsible for post translational regulation of gene expression in a sequence-specific manner (Makunin and Mattick, 2006). It is estimated that a $\frac{1}{3}$ of the protein-coding genes in the human genome is controlled/regulated by miRNAs (Makunin and Mattick, 2006). More than 1000 miRNAs have been discovered in humans to date. Many of these miRNAs have been implicated in common human diseases such as cancer (Kowdley and Li, 2012). All cancer types share certain characteristics described in Section 1.1.1. It has been suggested that miRNAs have the ability to regulate these processes signifying its involvement in the initiation and progression of various human cancers. Therefore, these miRNAs can potentially be used as biomarkers in cancer diagnosis (Slack and Stahlhut Espinosa, 2006).



In this study, miRNA expression profiling was performed using quantitative real-time polymerase chain reaction (qRT-PCR/qPCR) analysis on the miRNAs identified through computational biology (bioinformatics), in a previous study (Lottering, 2015). Quantitative RT-PCR is a sensitive technique utilized for the estimation of circulating miRNA expression levels (Balet *al.*, 2012). This quantitative technique has numerous benefits including (i) high sensitivity (ii) low amount of starting material is required (iii) high repeatability (iv) high throughput method (iv) it's less time-consuming and (v) quantitation can be performed over several orders of magnitude. Other methods for miRNA quantitative studies include

microarray hybridization and massively parallel/next-generation sequencing (NGS), but these techniques are time-consuming compared to qPCR (Bertone *et al.*, 2010).

4.1.1 Quantitative real-time PCR (qRT-PCR)

Real-time polymerase chain reaction (RT-PCR), commonly referred to as quantitative polymerase chain reaction (qPCR) is a laboratory technique based on the standard polymerase chain reaction (PCR) method (Lai, Rao and Huang, 2013; Guescini *et al.*, 2008). Currently, qRT-PCR is regarded as the gold standard in the quantitative analysis of nucleic acid including DNA, RNA and microRNA molecules, in all areas of molecular biology. The main reason for the success includes the high sensitivity to the single cell level, robustness, high specificity to the disease of interest, good reproducibility, broad dynamic quantification range, not requiring large amount of samples and most importantly, its affordability (Repa and Valasek, 2005; van Rooij, 2011; Bremnes *et al.*, 2014). The assay and primer design can often be automated. Additionally, this method is less time consuming compared to microarrays and NGS and the results do not require analysis or processing by biostatisticians (Bremnes *et al.*, 2014).

Quantitative PCR analysis is typically illustrated using an amplification plot (Figure 4.1). The technique can be broken down in four major phases: linear ground phase, early exponential phase, log linear phase and plateau phase (Mendrano and Wong, 2005). During the first 10-15 cycles, known as the linear ground phase, the PCR is in the initial stage and fluorescence emission at each cycle hasn't risen above background. During this phase the baseline fluorescence is calculated. During the early exponential phase, the amount of fluorescence

has reached a threshold where it is higher than the background levels. At the log linear phase, optimal amplification is reached with PCR product doubling after every cycle in an ideal reaction. Finally, during the plateau phase, the reaction components become limited and the fluorescence intensity is no longer useful for data calculation (Mendrano and Wong, 2005; Pabinger *et al.*, 2014).



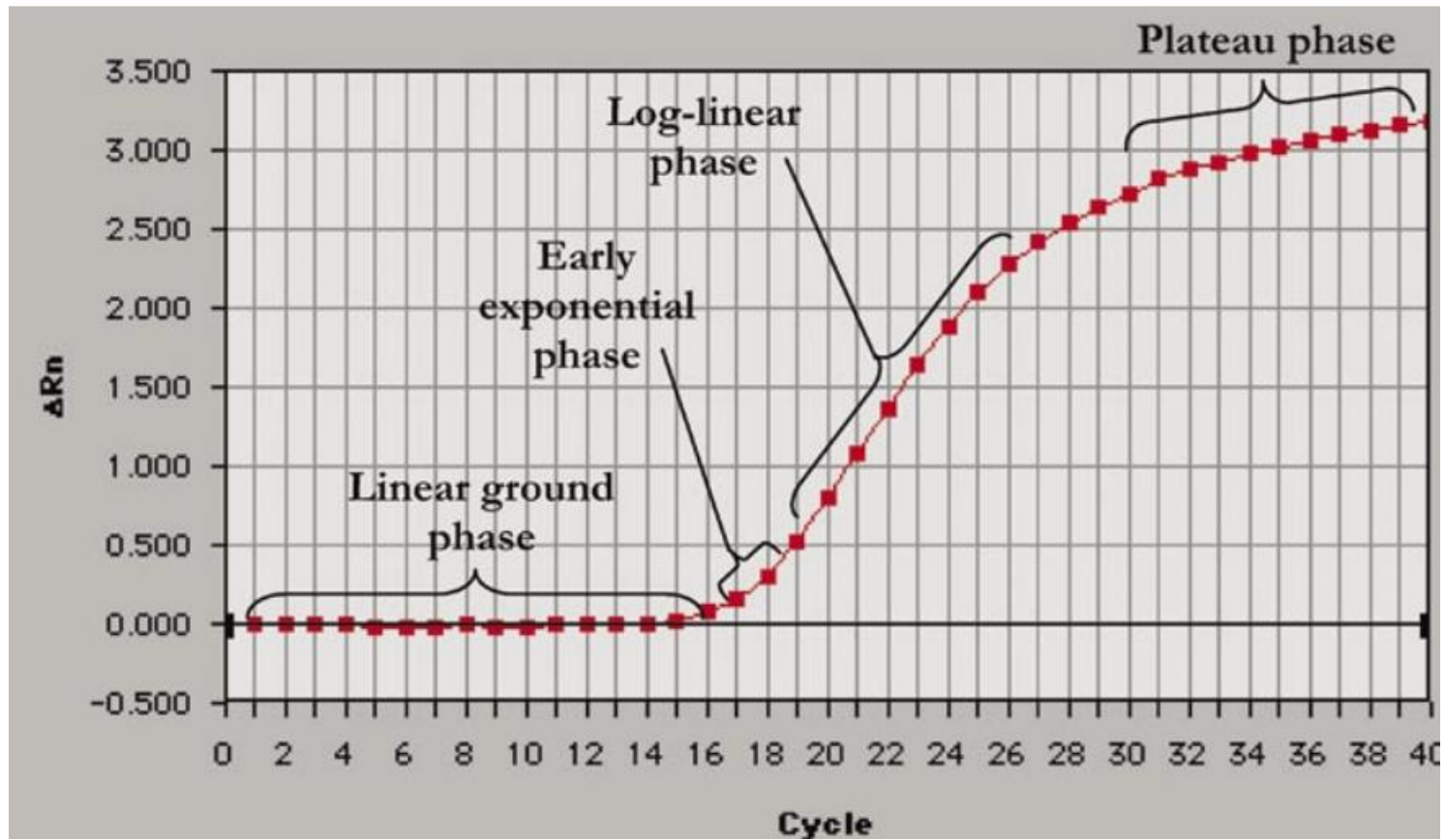
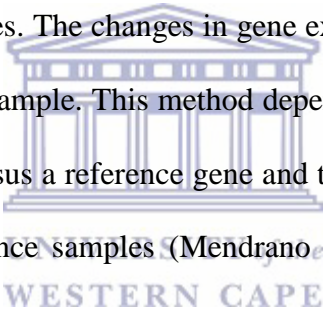


Figure 4.1: The graph illustrates the amplification curve produced following amplification using qPCR (Mendrano and Wong, 2005).

4.1.2 Quantitative Strategies in qRT-PCR

Absolute quantification and relative quantification are the two strategies employed to quantify gene expression in qRT-PCR (Mendrano and Wong, 2005; Yuan *et al.*, 2006). Absolute quantification determines the expression levels in absolute number of copies. It relies on a standard curve which is generated by using a serially diluted sample of a known concentration. The standard curve generates a linear relationship between the cycle threshold (C_T) and the initial amounts of total RNA or cDNA in the sample, enabling the determination of the concentration based on their C_T values (Mendrano and Wong, 2005; Yuan *et al.*, 2006). Relative quantification determines fold exchange in expression between two samples. The changes in gene expression is analysed in a given sample relative to a reference sample. This method depends on the comparison between expression of a target gene versus a reference gene and the expression of the same gene in target sample versus reference samples (Mendrano and Wong, 2005; Yuan *et al.*, 2006).



Expression profiling of the candidate miRNAs was performed in 2 Ovarian Cancer cell lines, 1 non-cancer cell line and an additional 3 non OC cancer cell lines namely (i) OWA28, Caov3, (ii) normal control (KMST-6) and (iii) other cancer cell lines (the cell lines include MCF-7, H157 and Hela). Our main aim is to establish which of the identified miRNA were specific to Ovarian Cancer and whether a definitive expression profile for OC could be established from the rest of the cancer cell lines using qRT-PCR. This was accomplished by using various molecular techniques including cell culture, mRNA extraction cDNA synthesis and qRT-PCR. It is expected that each

cancer cell line will have a unique expression profile thus it can be used to differentiate cancer type from each other and it can potentially be used for diagnostics.



4.2 Molecular methodologies

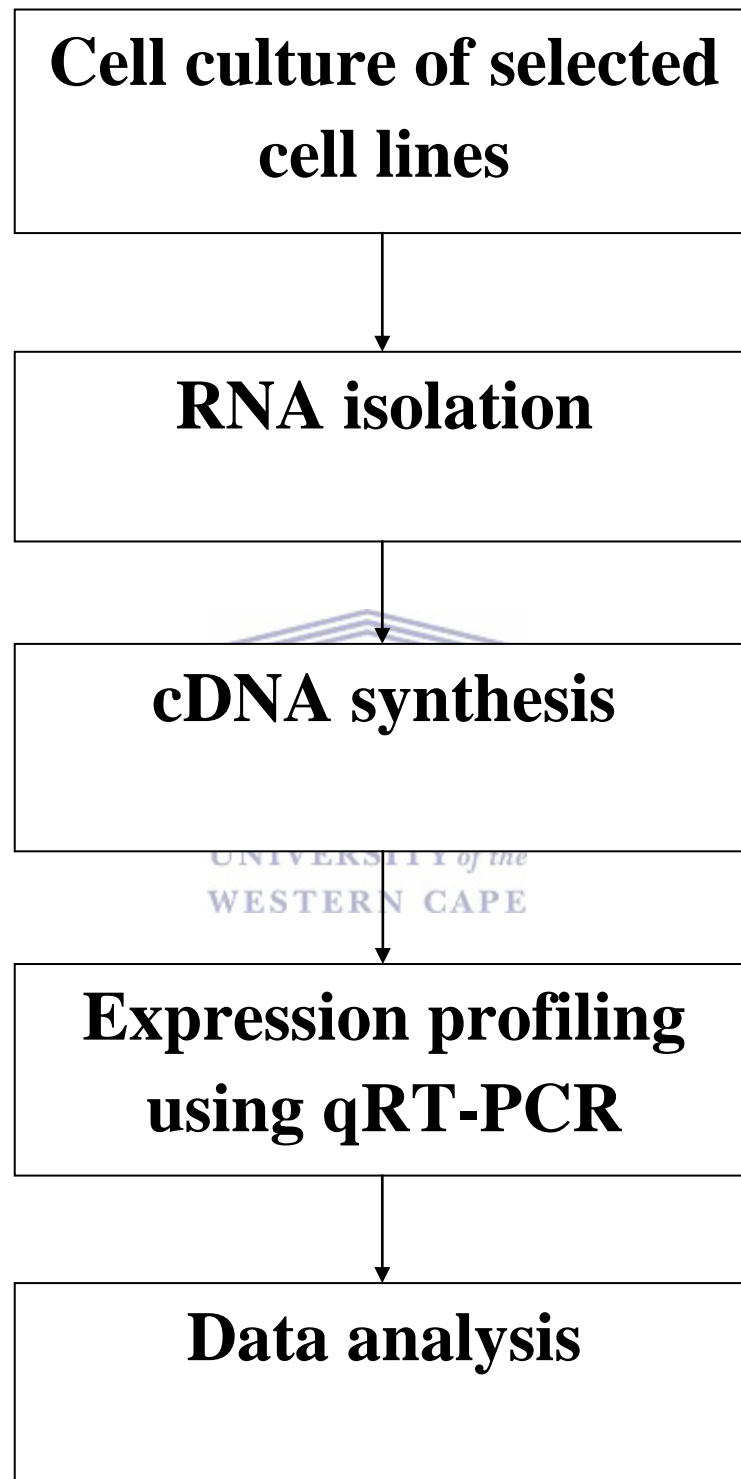


Figure 4.2: Flow diagram representing the molecular validation of miRNA's identified through bioinformatics analysis.

4.2.1 Cell culture

The panel selected (Table 4.1) contains 2 ovarian cell lines, 1 non-cancer cell line and 3 other cancer cell lines. The cell line, KMST-6 served as a control cell line against which expression of the miRNAs will be compared. These cell lines were selected to denote various types of ovarian tissue and to determine if miR1-miR6 were differentially expressed in Ovarian Cancer. The other cell-types (Table 4.1) was selected to assess the differential expression of the five miRNAs. Each cell line was cultured in specific media illustrated in Table 4.1 and supplemented with appropriate concentrations of Fetal Bovine Serum (FBS) and penicillin-streptomycin (Penstrep).



Table 4.1: A list of cell lines used to investigate miR1-miR6.

Name	Description	Status	Cell type	Media used	Supplement	Reference
OAW28	Ovarian adenocarcinoma	Cancer	Ovarian adherent	DMEM	FBS, Penstrep	Hills <i>et al.</i> , 1989
CaOV-3	Ovarian adenocarcinoma	Cancer	Ovarian adherent	DMEM	FBS, Pensrep	Karlan <i>et al.</i> , 1994
MCF-7	Mammary gland, metastatic site	Breast	Breast adherent	DMEM	FBS, Penstrep	Soule <i>et al.</i> , 1973
Hela	Cervix epithelial	Cervix	Cervix adherent	DMEM	FBS, Penstrep	Scherer, Syverton and Gey, 1953
KMST-6	Embryonic fibroblasts	Skin	Skin adherent	DMEM	FBS, Penstrep	Kawashima <i>et al.</i> , 1995
H157	Oral carcinoma/ Buccal mucosal	Oral	Oral adherent	DMEM	FBS, Penstrep	Prime <i>et al.</i> , 1990

4.2.1.1 Thawing of cells

Cryovials containing frozen cells were removed from the -150 °C freezer and allowed to thaw in a 37 °C water bath, until only a small piece of ice remained in the vial. The cryovial was sprayed with 70 % ethanol and transferred to a laminar flow hood. The cells were transferred to a 15 ml tube containing 3 ml of the appropriate complete media. The cell suspension was centrifuged at 3000 xg for 3 minutes. Following centrifugation, the 15 ml tube was sprayed with 70 % ethanol and transferred to the laminar flow hood. The supernatant was aseptically decanted and the pellet was re-suspended in 5 ml of the appropriate complete media, which was then transferred into a 25 cm² cell culture flask and incubated at 37 °C in 5 % CO₂. The cells were cultured with regular media changes until the cells were 90 % confluent.



4.2.1.2 Trypsinization

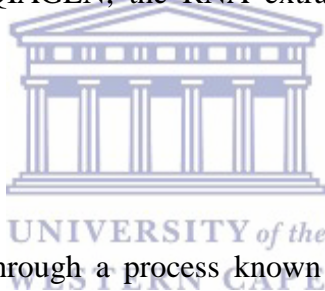
Once the cells have reached 90 % confluency, the cells were trypsinized. This was accomplished by decanting the media from the cell culture flask. The cells were then washed with 5 ml Phosphate Buffered Saline (PBS) to remove any remaining media. After the PBS was decanted, 3 ml of 1 X Trypsin was added to the flask. The flask was then placed into the incubator for approximately 2-5 minutes. Following incubation, the flask was taken to the microscope to assess the detachment process. When the cells were completely detached, 5 ml of appropriate complete media was added to the flask, to inactivate the trypsin. The cell suspension was transferred to a 15 ml tube and centrifuged for 3 minutes at 3000 xg.

4.2.1.3 Freezing down of cells

When the cells reached the desired confluency, the cells were detached through a process known as trypsinization as described in Section 4.2.1.2. The pellet generated after centrifugation, following trypsinization, was re-suspended in the appropriate complete media and 10 % dimethyl sulfoxide (DMSO). The suspension was aliquoted into 2 ml cryovials and stored at -150 °C.

4.2.2 Total RNA extraction

Using the miRNeasy kit from QIAGEN, the RNA extraction process was carried out as follows:



Confluent cells were detached through a process known as trypsinization as described in Section 4.2.1.2. Following centrifugation, the cell pellet was washed with PBS and then collected by centrifugation at 3000 xg for 3 minutes. This step was repeated twice. The cells were then re-suspended in 200 ul of PBS to make sure all the media was removed and then collected by centrifugation at 3000 xg for two minutes. The cells were lysed by adding 700 ul of lysis/binding buffer to the pellet. The sample was incubated for 5 minutes at 15-25 °C. Thereafter, 140 ul of chloroform was added and the sample was vortexed for 15 seconds. The sample was incubated at room temperature for 2-3 minutes after which the sample was centrifuged at 12 000 xg for 15 minutes. The upper aqueous phase was transferred to a new collection tube. Thereafter, 1.5 volumes of 100 % ethanol was added to this phase and mixed by pipetting. A high pure filter tube was combined with a collection tube and 700 ul of the sample was transferred to the upper reservoir of the high pure filter tube. The sample was

spun down at 8000 xg for 15 seconds and the flow through was discarded. The remainder of the sample was transferred to the high pure filter tube and then the filter tube was spun down at 8000 xg for 15 seconds. Following centrifugation, 500 ul of RPE buffer was added to the column, the sample was centrifuged for 15 seconds at 8000 xg and the flow through was discarded. Thereafter, 500 ul of RPE buffer was added and it was centrifuged for 2 minutes at 8000 xg. The flow-through was subsequently discarded. The high pure filter tube was transferred to a sterile 1.5 ml microcentrifuge tube. The RNA was eluted by adding 35 ul RNase free water to the column. The sample was centrifuged at 8000 xg for 1 minute. Following the completion of RNA extraction, the quality and quantity of the RNA was assessed using the Qubit (according to manufacturers' instructions) and a 1 % agarose gel.

4.2.3 Primer design



Mirbase is a publically available database accessed at www.mirbase.org/, described in (Griffiths-Jones and Kazomara, 2013). For the purpose of this experiment, the database was used to extract the miRNA sequences for primer design. The miRNA ID was used as an input for this database. The miRNA sequence was extracted and saved into a Word document for further analysis. The primers against the miRNAs to be analyzed using qRT-PCR were designed using freely available software (miRprimer) that is able to work in the MS Windows platform and in a developer version written in the Ruby programming language. MiRprimer identifies primers specific for the sequences uploaded by using an algorithm that is based on an implementation of previously published rules. It also evaluates the susceptibility of the formation of secondary structures and primer dimers (Busk, 2014). The miRNA sequences obtained from miRBase were used as an input and the best primer pairs generated against each miRNA was selected and saved in a Word document for further analysis.

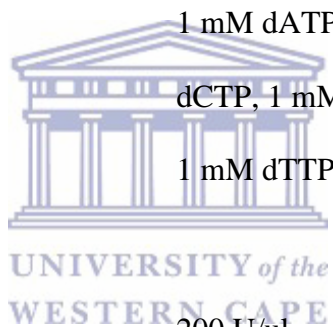
4.2.4 Poly (A) tailing and Reverse transcription

In a total volume of 10 ul, 100 ng of total RNA, prepared in Section 4.2.2, was mixed with reaction buffer *E. coli* poly (A) polymerase, ATP, RT primer, dNTP mixture, M-MuLV reverse transcriptase and *E. coli* polymerase as described in Table 4.2. The reaction was incubated at 42 °C for 1 hour followed by an inactivation step for 5 minutes at 95 °C. The concentration of the cDNA synthesized was determined using the Qubit system and the quality was determined by conventional PCR and a 4 % agarose gel electrophoresis.



Table 4.2: Poly (A) tailing and reverse transcription reagents used for cDNA synthesis

Reagents	Final concentration	Volume required
Reaction buffer <i>E. coli</i> poly (A) polymerase	10 X	2 ul
ATP	1 mM	2 ul
RT primer 5'- CAGGTCCAGTTTTTTTTTTTTTTTTTVN (V is A, C and G; and N is A, C, G, T)	10 uM	2 ul
dNTP mix	1 mM dATP, 1 mM dCTP, 1 mM dGTP, 1 mM dTTP	2 ul
M-MuLV reverse transcriptase	200 U/ul	1 ul
E.coli poly (A) polymerase	5 000 U/ml	0.4 ul
Total RNA	100 ng	variable
RNase free water		variable



4.2.5 The validation of miRNA expression levels in Ovarian Cancer and other cell lines using qRT-PCR

A 2 X LightCycler FastStart SYBR Green I master mix was prepared (Kappa SYBR fast Master Mix) according to the manufacturer's instructions. For all qRT-PCR reactions a standard reaction was prepared, containing 2 X KAPA SYBR FAST qRT-PCR master mix, forward and reverse primers (10 uM); and nuclease free water to a final volume of 9 ul as described in Table 4.3. Experiments were set containing decreasing concentrations starting with 250 ng of cDNA to 0.0025 ng. The reactions were prepared on ice. A 96 well Real Time PCR plate was placed on ice and 9 ul of the master mix and 1 ul of cDNA was aliquoted respectively to each well. Internal controls were prepared by adding nuclease free water instead of cDNA for each of the cell lines assessed (Table 4.3). The Real Time PCR plate was covered with a clear foil seal and centrifuged to make sure all the samples were at the bottom of the wells of the 96 well plate. The PCR plate was transferred to the Light Cycler 480 instrument and incubated at 95 °C for 5 minutes, followed by 40 cycles of 95 °C for 15 seconds and 60 °C for 15 seconds. For melting curve analysis, the samples were denatured at 95 °C, and then cooled to 65 °C at 20 °C per second. Fluorescence signals were collected at a wavelength of 530 nm.

Table 4.3: Reaction prepared for qRT-PCR

Reagents	Final concentration	Volume required
Forward primer	10 uM	0,5 ul
Reverse primer	10 uM	0,5 ul
PCR master mix (SYBR Green)	2 X	5 ul
Nuclease free water		variable

The standard or calibration curves were generated by the LightCycler software using serially diluted cDNA standards (250 to 0.0025 ng). Data on expression levels were obtained in the form of crossing points or cycle threshold. The expression levels were determined relative to the reference miRNA using the Livaks method (Livak and Schmittgen, 2008):

Step 1: Normalise ΔC_T (target miRNA) to ΔC_T (reference miRNA)

1. ΔC_T (control calibrator) = C_T (target calibrator) – C_T (reference calibrator)
2. ΔC_T (tumour test) = C_T (target test) – C_T (reference test)

Step 2: Normalise ΔC_T (tumour test) to ΔC_T (reference test)

Change in expression = ΔC_T (tumour test) - ΔC_T (reference test/ control calibrator)

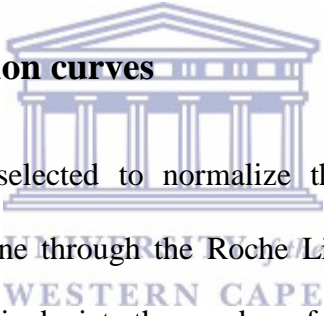
Step 3: Fold difference in expression between the tumour test and the reference test

Normalise expression ratio = $2^{-\Delta\Delta C_T}$

If the first ΔC_T is greater than the second ΔC_T then the value of $2^{-\Delta\Delta C_T}$ will be less than 1, implying that there is a reduction in expression. Therefore, the negative inverse of $2^{-\Delta\Delta C_T}$ will give the fold change reduction in expression.

4.3 Results

4.3.1 Analysis of amplification curves



Internal controls (Let7a) were selected to normalize the relative quantification of the candidate miRNAs. This was done through the Roche LightCycler 480 using the melting curve analysis software. The x-axis depicts the number of cycles while the y-axis shows the fluorescence of the reference dye (SYBR green). Typically, C_T values below 29 cycles show abundant nucleic acids and the C_T values above 39 cycles indicate minimal amounts, and possibly contamination. The amplification plot for the reference miRNA (Let7a) is shown in Figure 4.3. Figure 4.4 shows the amplification plot of miR 1. The optimum dilution is 25 ng at approximately 28 cycles for both Let 7a and miR 1.

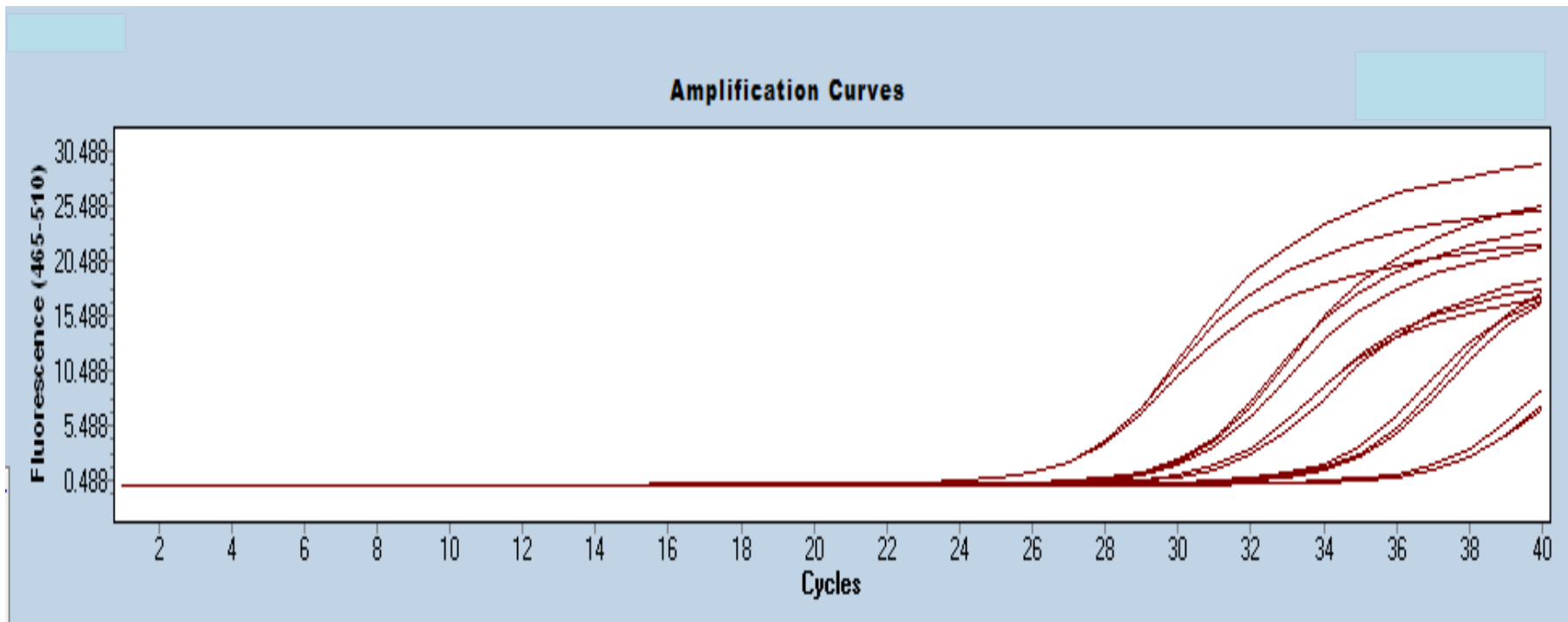


Figure 4.3: Amplification curve of Let7a in KMST-6 cell line. The various amplification curves represent various concentrations of KMST-6 cDNA ranging from 250 ng to 0.0025 ng.

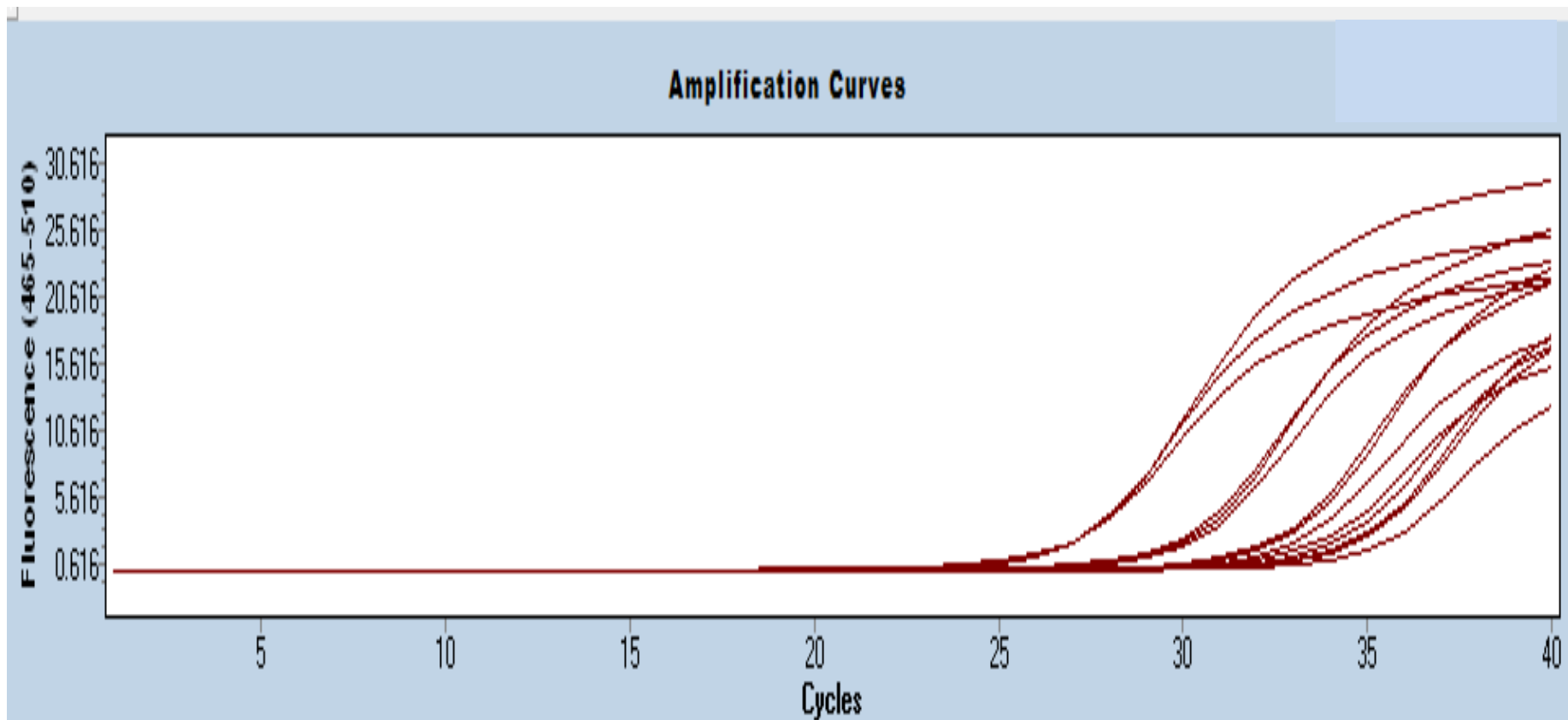


Figure 4.4: Amplification of miRNA 1 in KMST-6 cell line. The various amplifications curves represent different KMST-6 cDNA concentrations ranging from 250 ng to 0.0025 ng.

4.3.2 Melting curve analysis

Melting curve analysis was performed to determine the homogeneity of the PCR product investigated. The melting curves generated were used to determine if there was (a) mis-priming, (b) contamination (c) or any other inconsistencies relating to the amplification process. The curves were generated by plotting the rate of change of the fluorescence units (RFU) with time (T) $(-d(\text{RFU})/dT)$ versus the temperature. The output is a single peak occurring at the desired melting temperature. The melting curve for the reference miRNA and miR1 is shown in Figure 4.5 and Figure 4.6. From the Figures it is evident that only one peak was present for the reference miRNA as well as the target miRNA (miR1) thus the single peak observed presents a pure, single target product (amplicon).



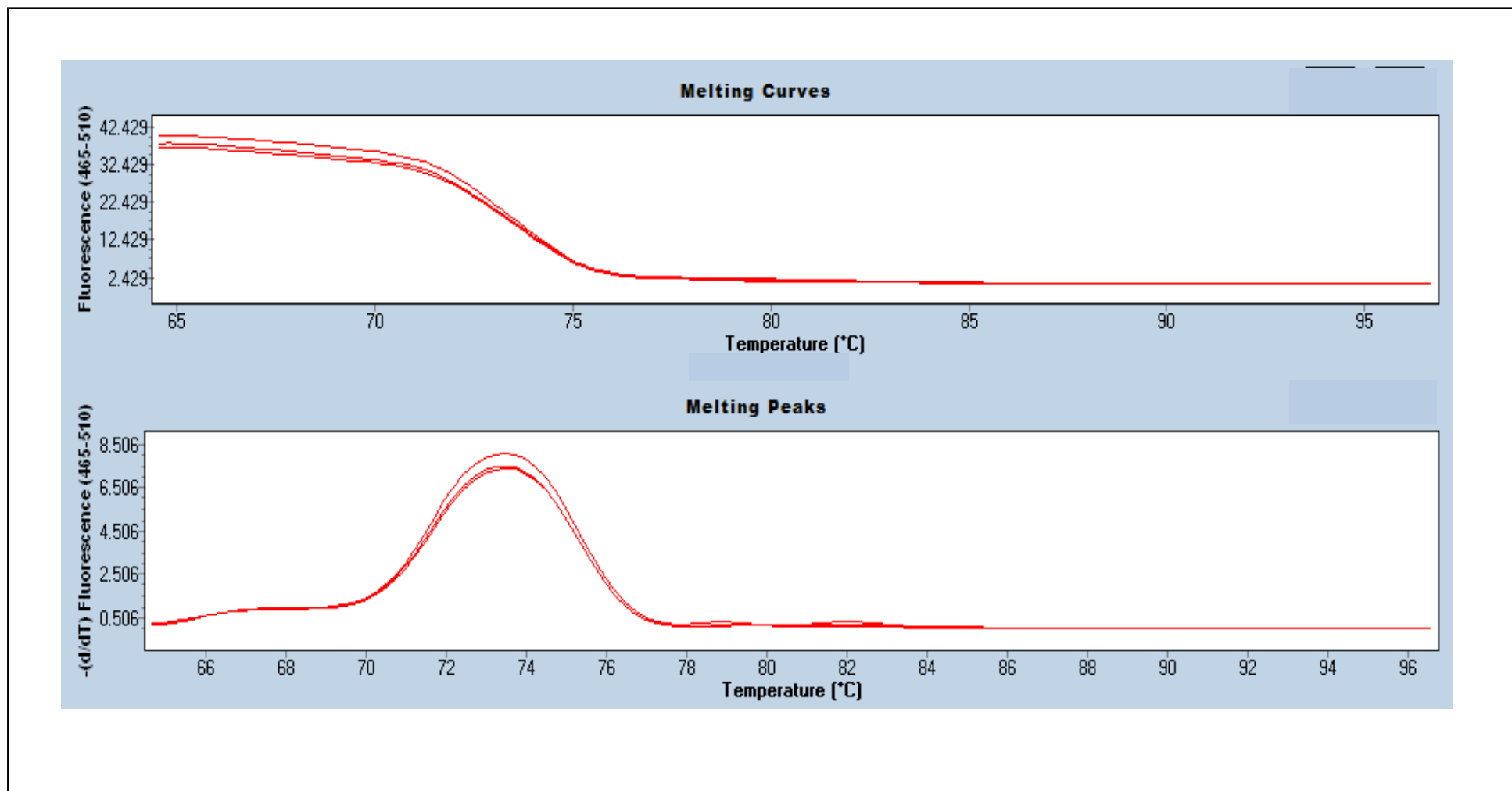


Figure 4.5: Let7a in KMST-6. A prominent peak is seen at T_m of 73°C.

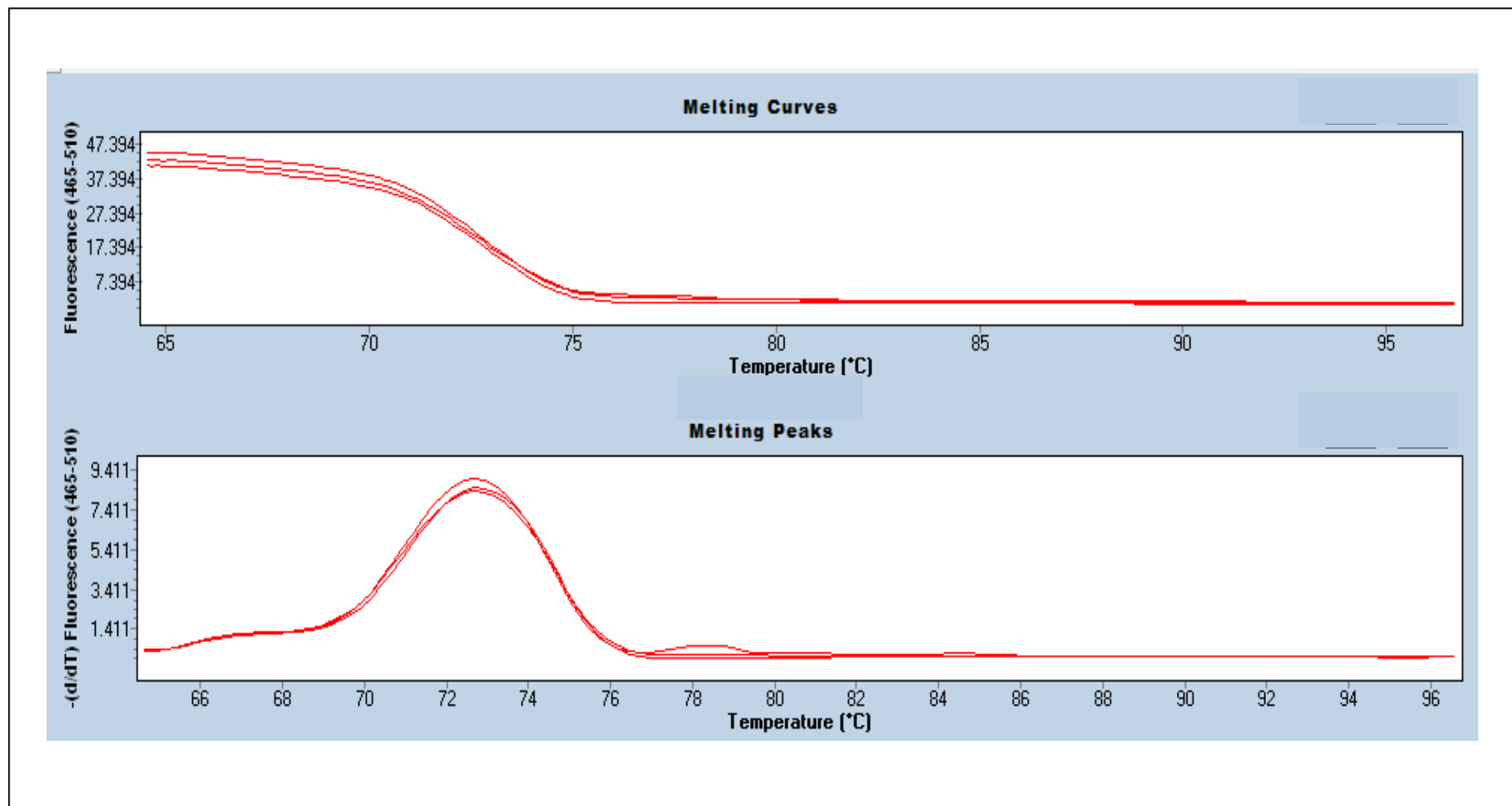
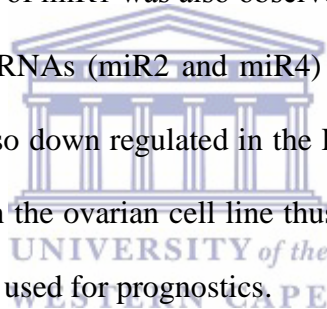


Figure 4.6: MiR1 in KMST-6 cells. A prominent peak is seen at T_m of 72°C.

4.3.3 Analysis of qRT-PCR expression profiling data

Expression profiling was performed to determine the specificity of the candidate miRNAs identified using *in silico* methodologies (Lottering, 2015). Various cancer cell lines were used including Ovarian Cancer cell lines and a normal cancer cell line illustrated in Figure 4.7 to Figure 4.12. The KMST-6 cell line, normal fibroblast, was used for sample normalisation. Differential expression of the six miRNAs was observed across all cancer cell lines shown in Figure 4.12. MiR1 is significantly highly expressed in both Ovarian Cancer cell lines compared to other cancer cell lines investigated (OAW28 and CoaV-3), indicating a unique expression profile. The expression ratios were 8.9 and 10.5 respectively (Figure 4.10 and Figure 4.11). Elevated expression of miR1 was also observed in HeLa and H157 but it wasn't significant. Furthermore, two miRNAs (miR2 and miR4) were under expressed in CoaV-3 (Figure 4.11), however, it was also down regulated in the HeLa (Figure 4.7). The expression ratio was down regulated more in the ovarian cell line thus by combining it with the current prognostic biomarkers it could be used for prognostics.



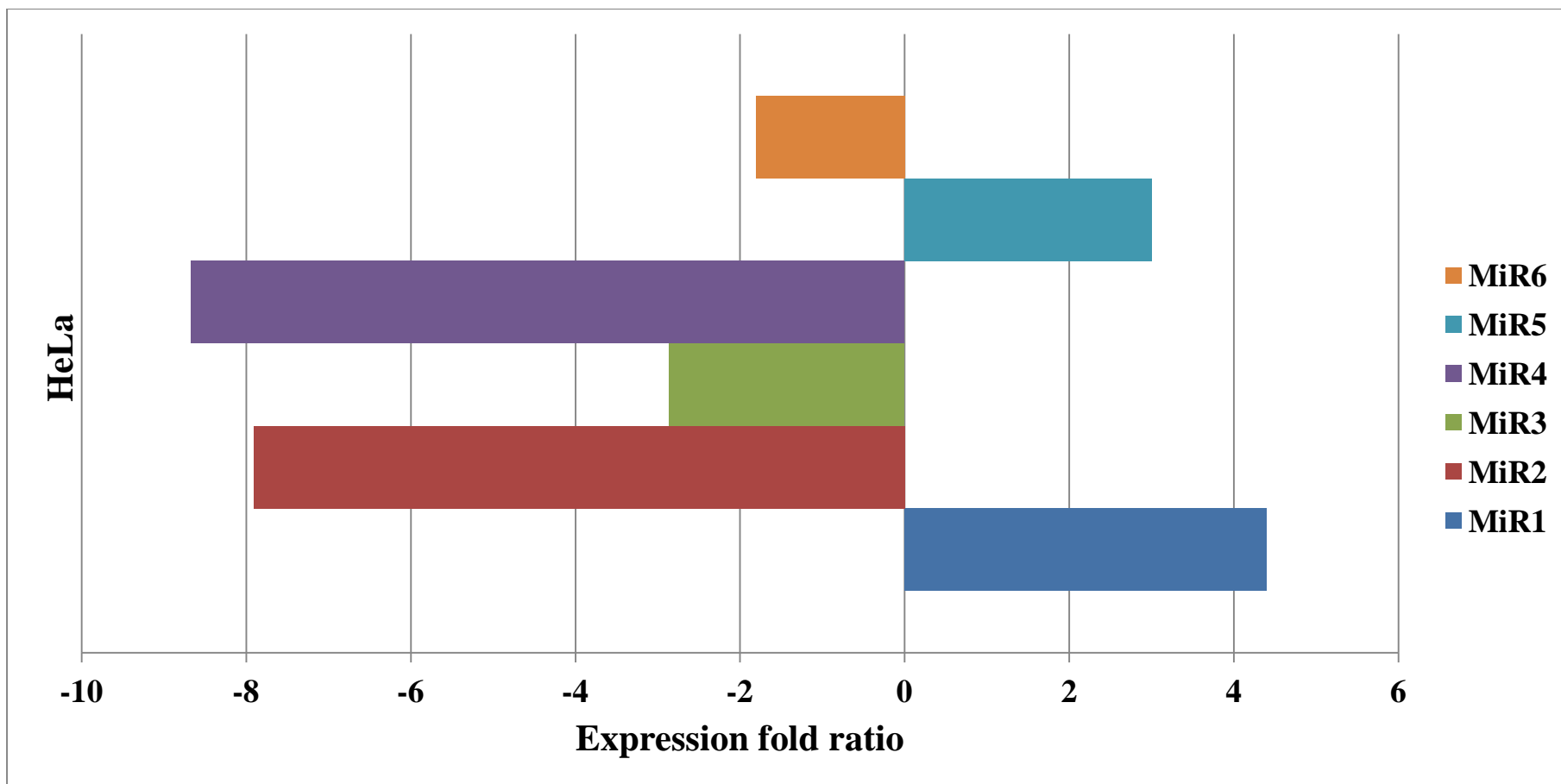


Figure 4.7: Relative expression ratio plot comparing the expression of the six candidate miRNAs in the HeLa (cervical epithelial) cell line.

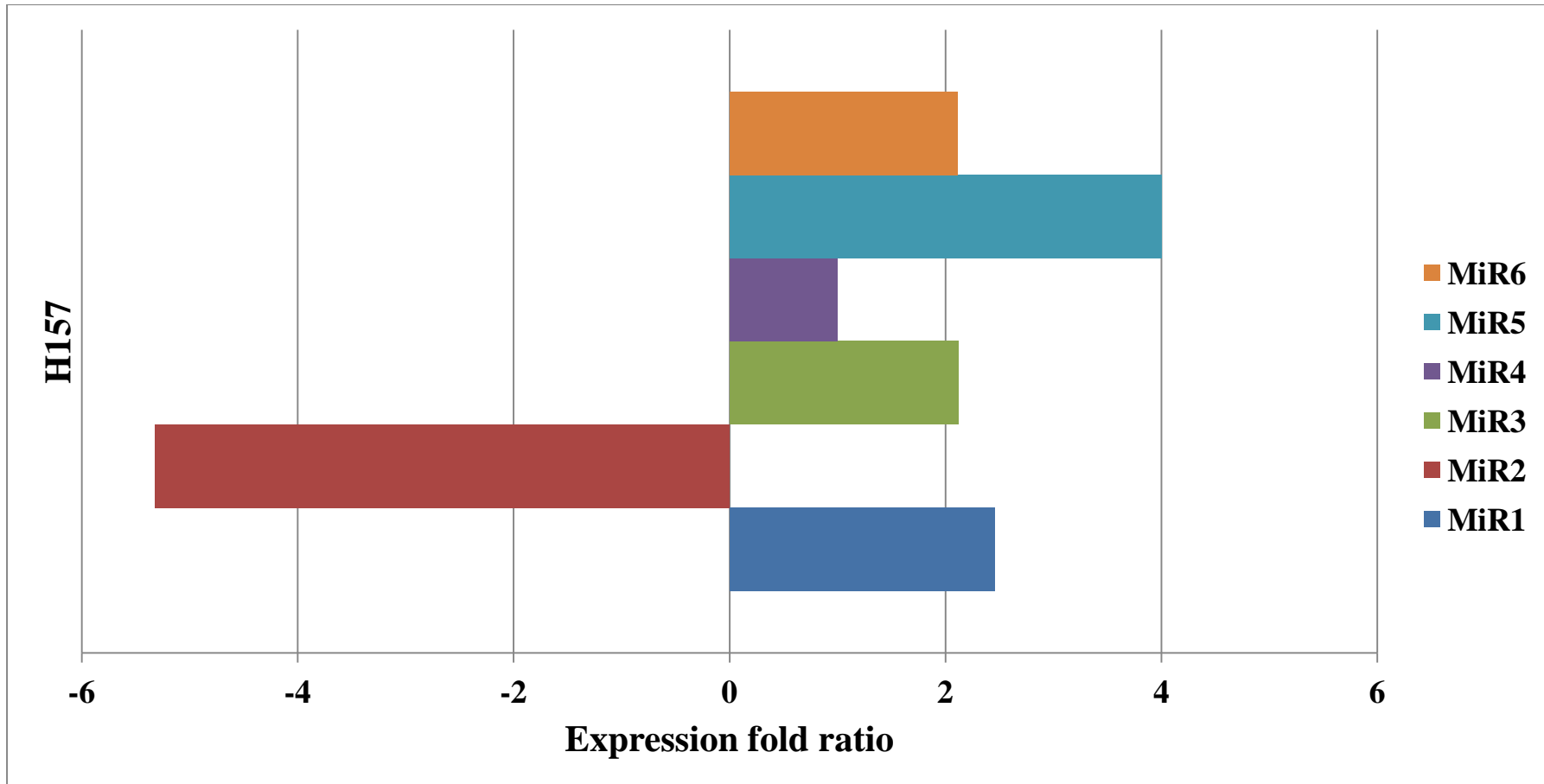


Figure 4.8: Relative expression ratio plot comparing the expression of the six candidate miRNAs in the H157 (oral carcinoma/ buccal mucosal) cell line.

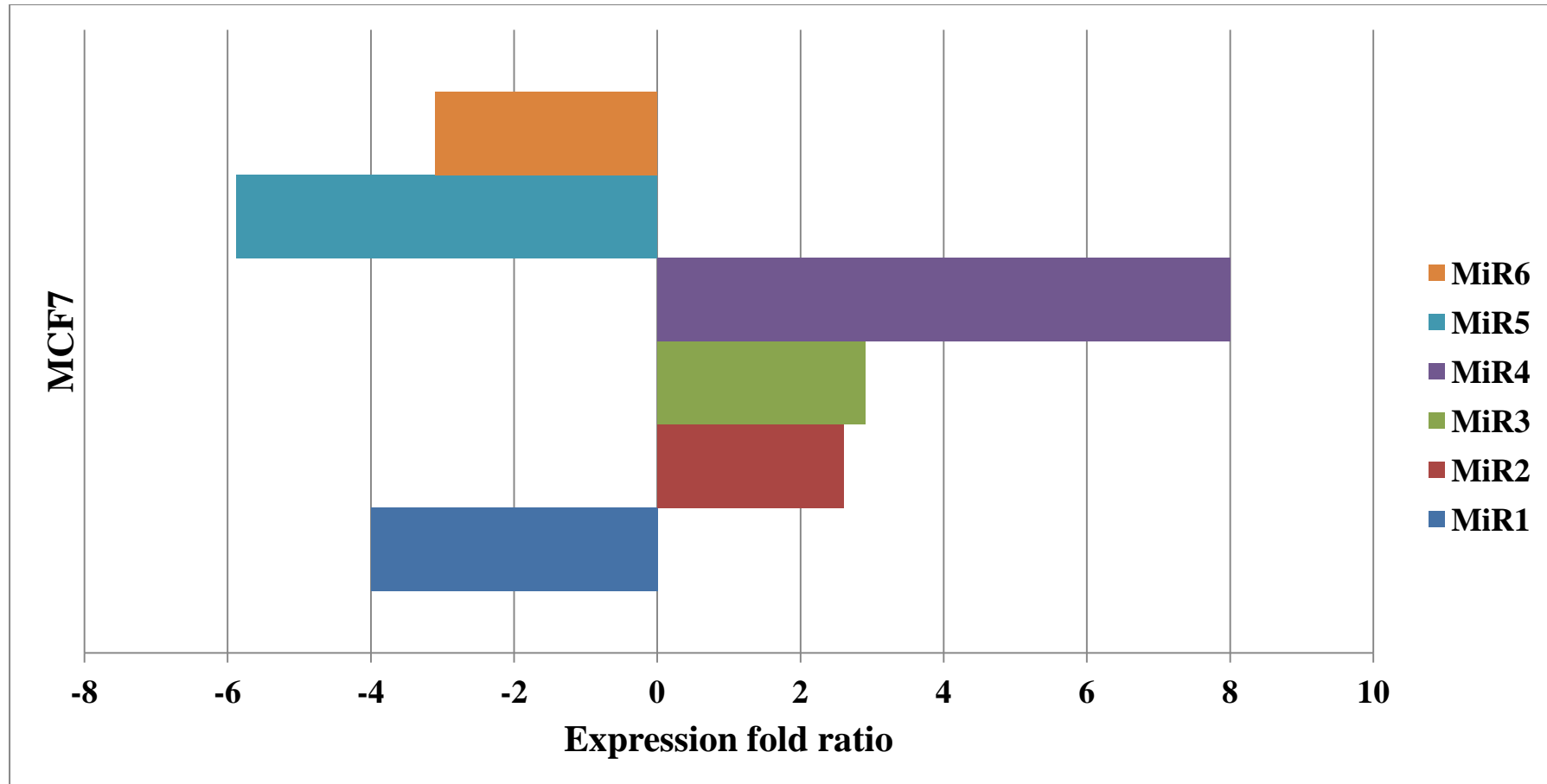


Figure 4.9: Relative expression ratio plot comparing the expression of the six candidate miRNAs in the MCF7 (mammary gland, metastatic site) cell line.

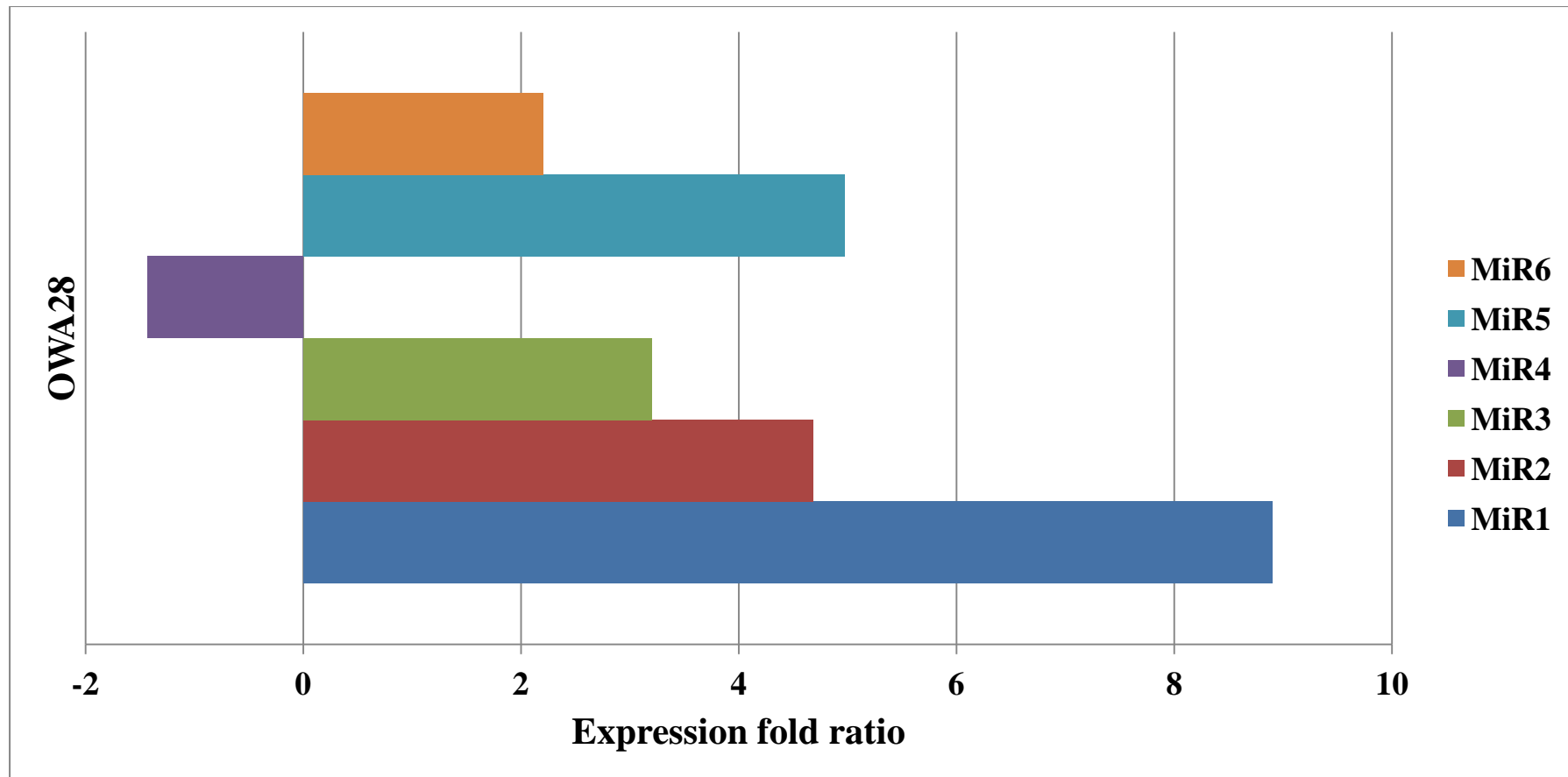


Figure 4.10: Relative expression ratio plot comparing the expression of the six candidate miRNAs in the OAW28 (ovarian adenocarcinoma) cell line.

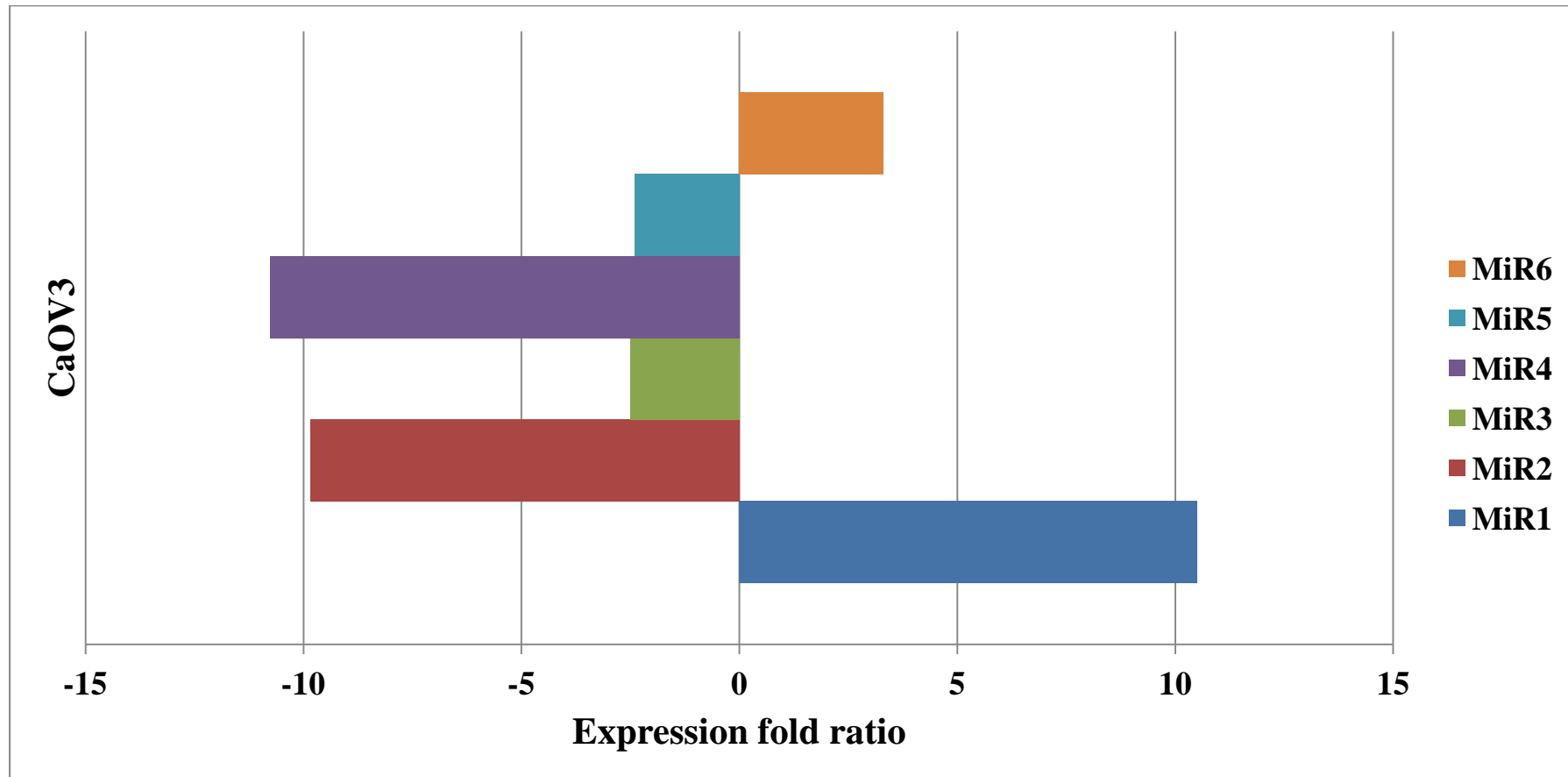


Figure 4.11: Relative expression ratio plot comparing the expression of the six candidate miRNAs in the CaOV-3 (ovarian adenocarcinoma) cell line.

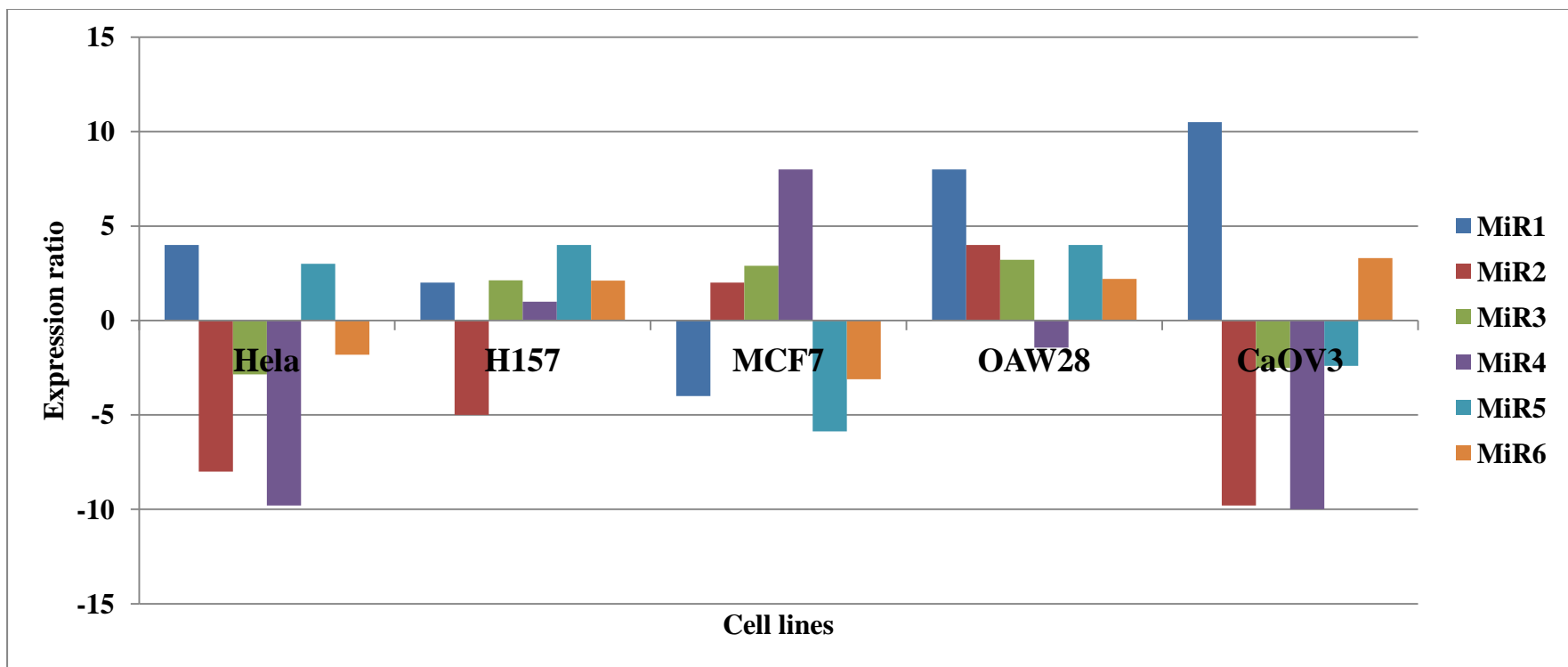
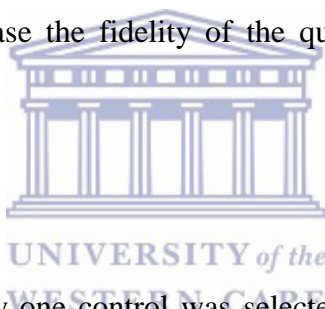


Figure 4.12: Relative expression ratio plot of the six miRNAs in various cancer cell lines including Ovarian Cancer (OAW28 and CaOV-3) cell lines. The bars indicate up-regulation (above 0) and down-regulation (below 0) of the miRNAs.

4.4 Discussion

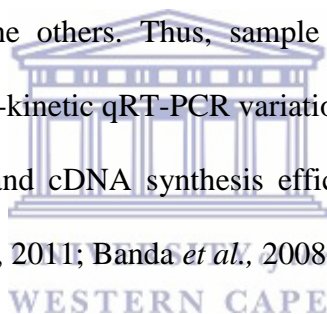
The aim of the chapter was to determine the relative expression of the six candidate miRNAs identified through *in silico* methodologies in a previous study (Lottering, 2015). Based on the strong association of these miRNAs with Ovarian Cancer, they were selected as potential diagnostic biomarkers for early detection of the disease. An array of tumour cell lines including Ovarian Cancer cell lines were selected to molecularly validate the putative miRNAs using qRT-PCR as described in Section 4.2. Normalization using an internal reference miRNA (Let7a) was used to compensate for the variations in expression patterns that could occur as a result of sample preparation and cDNA synthesis to increase the fidelity of the quantification process (Ling and Salvaterra, 2011).



During the course of the study one control was selected namely Let7a because of its stability in all cell lines (Kinose *et al.*, 2014). Let7a precursor miRNA is widely viewed as a tumour suppressor miRNA. A vast majority of its targets tend to have oncogenic properties. It was determined that Let7a target Caspase-3, a pivotal protease activated during apoptosis in many cell lines including OC cell lines. Furthermore, it has been suggested that Let7a regulates various factors that play a role in the cell cycle and cell proliferation which in turn functions as regulators of the initiation and progression of all cancer types (Boyerinas *et al.*, 2010; Kinose *et al.*, 2014 and Rouch and Slack, 2008).

4.4.1 Analysis of Amplification curves and Melting curves

An amplification curve and melting curve was constructed to differentiate between the specific amplicon compared to primer-dimers. The amplification curves demonstrated in Figure 4.3 and Figure 4.4 were obtained from a dilution series (250 ng to 0.0025 ng) of miR1 and the reference miRNA (Let7a). Upon examination of Figure 4.3 and Figure 4.4, clustering of the amplification curves at the respective dilutions was evident. However in Figure 4.4, at lower concentrations reproducibility of the triplicates were not consistent. The inconsistencies could be a result of random errors including variations in template input due to pipetting. This is indicated by a higher C_T values in one triplicate compared to the others. Thus, sample normalisation is important to compensate for inter- and intra-kinetic qRT-PCR variations including sample variations, RNA extraction and quality and cDNA synthesis efficiency (Reboucas *et al.*, 2013; Turabelidze, Guo and DiPietro, 2011; Banda *et al.*, 2008).



Melting curve analysis was performed to check the specificity of the system. It is generated by plotting the rate of change of the relative fluorescence units (RFU) with time (T) ($-d(\text{RFU})/dT$) versus the temperature. The output is a melting curve with a peak signal occurring at the melting temperature (T_m). Strand complementarity, product length, sequence and GC content are responsible for the T_m of a particular sample (Blake, 2006; Lando *et al.*, 2015). Therefore, non-specific amplification products, mis-priming, primer-dimer artefacts and inhibitor binding will result in alterations of the melting temperature and curve of the particular sample. As indicated by Figure 4.5 and Figure 4.6, only one prominent peak was amplified using miR1 and Let7a respectively.

However, in Figure 4.5 there is a small peak at 78.5 °C, in one of the triplicates. One explanation for the two peaks could indicate the presence of contaminating or off-target products (Lando *et al.*, 2015). Furthermore, prominent peak is seen at T_m 72 °C in Figure 4.5 and at T_m 73 °C in Figure 4.6. This corresponds to the melting temperatures calculated from the miRNA sequence obtained from mirBase, indicating that only the target products amplified (Griffiths-Jones and Kazomara, 2013). Hence, it can be concluded that there was no interference in the PCR reactions with regards to non-specific products or primer dimers for all miRNA gene targets (data not shown).

4.4.2 Expression profiling plot analysis

The comparative C_T method (also referred to as the $2^{-\Delta\Delta CT}$ method) was used to calculate relative changes in expression between two samples using a mathematical model. The data generated was imported into an Excel spreadsheet to create a graph illustrating the relative expression ratios of the candidate miRNAs in various cell lines described in Section 4.2.5. The target miRNA was normalized to a reference miRNA (Let7a) to minimise sample to sample variations (Roa *et al.*, 2014; Schmittgen and Livak, 2008). Analysis was performed to evaluate the specificity of the putative miRNAs identified as potential biomarkers for early detection of the disease in two OC cell lines, a non-cancerous cell line and other cancer cell types (illustrated in Table 4.1). This study aimed to identify which miRNAs were differentially expressed during the development of the disease. The six candidate miRNAs were relatively measured in all cancer types against the non-cancer cell line, KMST-6.

Differential expression was observed in all cancer cell lines (Figure 4.7 to Figure 4.12), with miR1 being significantly up-regulated in the two Ovarian Cancer cell lines: OWA28 and CaoV-3 (neoplasia/primary ovarian tumour) (Figure 4.10 and Figure 4.11) in comparison to other cell lines tested (Hills *et al.*, 1989; Karlan *et al.*, 1994). MiR1 was up-regulated with a factor of 8.90 and 10.51 respectively. The qRT-PCR analysis showed that miR1 was over-expressed in Hela and H157 with a factor of 4.40 and 2.45 respectively. However, it was more significantly over expressed in the Ovarian Cancer cell lines thus it could be a good diagnostic indicator for Ovarian Cancer. From the *in silico* studies, miR1 was predicted to regulate Cartilage Intermediate Layer Protein (CILP), illustrated in Table 2.1. The miRNA target gene has been linked to processes, functions and pathways known to be involved in the initiation and progression of Ovarian Cancer (Section 2.4). CILP has been found to have decreased expression in tumour tissue (46.42 digital expression units, DEU) compared to normal tissue (112.33 DEU) in the GeneHub database. Also, in TiGER database, CILP was found to be preferentially expressed in normal ovarian tissue (Lottering, 2015). This suggests that over expression of miR1 results in a decreased expression of the target gene. However, further analysis is required to definitively conclude the association.

Through fluorescence microscopy assay Seki and associates (2005) showed that CILP is co-localized with Transcription Growth Factor beta (TGF β) receptors suggesting that these genes/proteins have a significant statistical biological relationship. CILP may act by antagonizing TGF β 1 functions. Furthermore; deregulation of TGF β receptors was strongly associated with Ovarian Cancer. TGF β has the ability to transform from a tumour suppressor (in normal ovarian surface epithelial cells) to a tumour promoter.

When TGF β acts as a tumour promoter, it enhances tumour cell proliferation and promotes metastasis through the induction of the Epithelial- Mesenchymal-Transition (EMT) process. Thus, over expression of the TGF β receptors in human ovarian tumours results in a poor prognostic outcome of the disease (Cheng *et al.*, 2012; Hirashima *et al.*, 2003; Yeung *et al.*, 2013).

Additionally, as shown in Section 2.4.1, CILP is strongly associated with Insulin-Like growth factor-1 (IGF-1) and Insulin-Like growth factor-2 (IGF-2) signalling. IGF-1 has a significant role in cellular proliferation, cell metabolism, differentiation and survival. Moreover, deregulation of IGF-1 is involved in carcinogenesis of various tumour entities including ovarian tumours (Rohr *et al.*, 2016; Li *et al.*, 2016). Endometriosis in the ovary confers a hormonal imbalance and triggers an up-regulation of growth factors such as IGF-1 to which Ovarian Cancer cells have demonstrated dependency. Moreover, IGF-1 levels are higher in severe cases of endometriosis (Nezhat *et al.*, 2008; Koshiyama, Matsumura and Konishi, 2014; Gianuzzi *et al.*, 2016). A study conducted by Kuroda *et al.*, (2001) showed that over-expression of IGF-1 inhibit apoptosis in normal ovarian surface epithelial cells following a 72 hour exposure to human chorionic gonadotropin (hCG) hormone. They also investigated the effect of IGF-1 independently on apoptosis of normal ovarian surface epithelial cells using Cell Death Detection ELISA (Kuroda *et al.*, 2001). Treatment with 0.1 ug/ml and 0.5 ug/ml IGF-1 decrease apoptosis in ovarian surface epithelial cells by 65.9 % and 70.4 % respectively compared to the control. These findings indicate that deregulation of IGF-1 is significant in the inhibition of apoptosis. Additionally, endometriosis and Ovarian Cancer has shown common genetic alterations suggesting a possible malignant genetic

transition spectrum, from endometriosis to OC (Kuroda *et al.*, 2001; Nezhat *et al.*, 2008; Gianuzzi *et al.*, 2016).

A study done by Dong and associates (2015) investigated the significance of IGF-2 in clinical outcome of Ovarian Cancer patients with regards to overall survival (OS) and progression-free survival (PFS) using a Kaplan-Meier test. They showed that there was a statistical significance of groups with elevated IGF-2 expression levels compared to low expression levels in relation to OS (HR=1.44; p=0.000). Additionally, patients with high IGF-2 expression also had a poorer PFS compared to the low expression group (HR=1.35; p=0.000) (Dong *et al.*, 2015). They also investigated the expression levels of the gene in normal tissue compared to OC tissue; and the prognostic significance in OC patients at different stages, histological grades and following treatment. The expression of IGF-2 was increased in Ovarian Cancer compared to normal tissue at the mRNA and protein level (Dong *et al.*, 2015). The analysis also revealed that elevated expression of IGF-2 results in poor prognostic patient outcome at clinical stages I, II and III, histological grade 2 and 3; and those patients treated with chemotherapy containing platin and Taxol (Dong *et al.*, 2015). This further supports the result that miR1 is over-expressed in both Ovarian Cancer cell lines illustrated in Figure 4.9 and Figure 4.10 respectively.

Additionally, miR2 (CILP) and miR4 (NPR3) were significantly down-regulated in Coav-3, with expression ratios of -9.85 and -10.78 respectively (Figure 4.11). These miRNAs were also under expressed in HeLa as shown in Figure 4.7. The expression

levels were -7.91 and -8.67 respectively. Regardless of these miRNAs being under expressed in HeLa it is more under expressed in the ovarian cell lines. Clinical and basic research investigated the use of multiple biomarkers to overcome the drawbacks of the current diagnostic and prognostic biomarkers. Using a combination of biomarkers has produced promising results and it could further increase the prognostic significance of the current biomarker, CA-125 (Huang, Hu and Sood, 2014; Nozaki *et al.*, 2009).

According to Section 3.5 the genes (CILP and NPR3) these miRNAs target are good prognostic markers in Kaplan-Meier plotter database (Figure 3.2) and SurvExpress database (Figure 3.3) as described in Section 3.5. CILP is associated with various mechanisms involved in the initiation and progression of Ovarian Cancer as previously described (Hirashima *et al.*, 2003; Cheng *et al.*, 2012; Yeung *et al.*, 2013; Rohr *et al.*, 2016; Li *et al.*, 2016). However, the molecular mechanisms underlying NPR3 in the development of cancer is not well defined. Studies have suggested the role of NPR3 in anti-apoptosis via breast cancer type 1 susceptibility protein (BRCA1) and tumour necrosis factor α (TNF- α) (Lin *et al.*, 2016). Numerous studies indicate the dysregulation of these genes in the commencement of cancer, specifically OC. Additionally; studies indicated the involvement of NPR3 in hormone binding, transferrin endocytosis, cAMP biosynthetic process, adenylate cyclase activity, metabolic and growth processes; all of which have been previously described in cancer processes in Chapter 2.

Furthermore, reduced levels of mature miRNAs have been reported in many tumours, including ovarian tumours. It is often associated with poor prognosis of various cancer types. Under-expression is a consequence of genomic copy number loss, epigenetic silencing, and deregulation of their biogenesis pathway as well as transcriptional repression. Reduced DICER and DROSHA expression levels are well documented in an array of cancer types (Jansson and Lund, 2012; Rupaimoole *et al.*, 2016). In 2008, Merritt and associates investigated the effects of deregulated DICER and DROSHA expression in Ovarian Cancer. They measured the mRNA levels of the components in OC patients using qRT-PCR and compared the result to clinical outcomes. Validation was performed using published microarray data from cohorts of OC patients. Their findings indicated that decreased levels of DICER and DROSHA are associated with poor patient survival. Similarly, Rupaimoole and associates (2014) showed that deregulation of the miRNA biogenesis pathway as a consequence of decreased DICER and DROSHA, results in a poor clinical outcome of Ovarian Cancer patients. MiR-503 and its target genes are down regulated by DICER in high grade Ovarian Cancer. The miRNA induces cancer cell growth and migration, ultimately resulting in poor prognostic patient outcome (Park *et al.*, 2013; Li *et al.*, 2015). Moreover, elevated DICER and DROSHA mRNA levels are associated with an increased median survival (Peng and Croce, 2016).

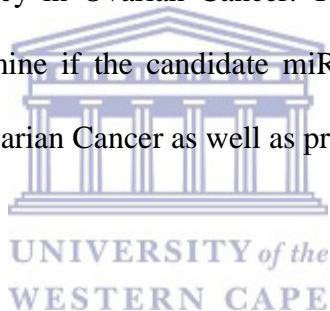
Furthermore, we also examined the expression of miR3, miR5 and miR6 in the 6 cancer cell lines in comparison to the normal cell line as seen in Figure 4.7-Figure 4.11. The data as depicted in the above mentioned figures showed no significant elevation as well as under expression. MiR5 which targets NDEL1 as seen in Table 2.1 was found to be

up-regulated in H157 and OWA28 (Figure 4.8 and Figure 4.10). It was also found to be down regulated in MCF7 (Figure 4.9). NDEL1 functions by anchoring the centrosome. Centrosome aberrations were postulated to cause cancer by promoting genome instability. During mitosis this presents a paradox. Cancer cells possessing extra chromosomes utilize multipolar mitosis to escape death by clustering of supernumerary centrosomes into bipolar arrays. Supernumerary centrosomes are frequently found in cancer cells (Godinho, Kwon and Pellman, 2009; Zyss and Gergely, 2009; Gonczy, 2015). In addition, miR3 and miR6 shows differential expression in the cancer cell lines investigated but the results weren't significant in any of the cell lines. A given miRNA may only be expressed in some cancer tissues but not in others. It may also be present at different stages of cancer development or under certain circumstances (Yue and Tigyi, 2006; Wijnhoven, Michael and Watson, 2007). This suggests that the two miRNAs are dysregulated and regulate the identified target genes (Table 2.1) at different stages of cancer and cancer cell lines, other than the cell lines used for this study illustrated in Table 4.1.

4.5 Conclusion

Molecular validation is a crucial step in biomarker discovery. There are arrays of methodologies that can be employed for biomarker discovery but they have their limitations. In this study we evaluated the expression profiles of the six miRNAs that were predicted via *in silico* methodologies. It is evident that the miRNAs in the study are differentially expressed across all cancer cell lines investigated. MiR1 showed significant up-regulation in the two Ovarian Cancer cell lines used, OWA28 and CoaV-

3, compared to the normal cell line (KMST-6). Also, the miRNA was more over expressed in the OC cell line compared to the other cancer cell lines used in the study. The expression ratio of miR1 was found to be 8.91 and 10.51 in OWA28 and CoaV-3 respectively. This makes it a good candidate biomarker for OC diagnosis. However, additional cell lines and patient samples are required to predict accuracy of the biomarker in the disease. The study also showed that miR2 and miR4 could be a potential prognostic biomarker for Ovarian Cancer. Nonetheless, these miRNAa were also under expressed in HeLa. The expression ratios were -7.91 and -8.67 respectively. Using these miRNAs in combination with the current prognostic tool, CA-125, could improve the predictive accuracy in Ovarian Cancer. The study serves as a basis for future investigations to determine if the candidate miRNAs can be used as potential biomarkers for diagnosis of Ovarian Cancer as well as prognostic markers.



4.6 Reference List

1. Ardekani, A.M. and Naeini, M.M. (2010) The Role of MicroRNAs in Human Diseases. *Avicienna Journal of Medical Biotechnology*, 2. P. 161-179.
2. Bai, Z., Bi, J., Han, S., Han, W., Ma, X., Meng, H., Song, J., Zhang, J. and Zhang, Z. (2012) Identification of suitable reference genes for qPCR analysis of serum microRNA in gastric cancer patients. *Digestive Diseases and Sciences*, 9. P. 1-8.
3. Bast, R.C Jr., Berchuck, A., Cameron, B., Chan, D.W., de Bruijn, H.W., Fung, E.T., Haafte-Day, C., Hacker, N.F., Jacobs, I.J., Li, J., Meng, X.Y., Rai, A.J., Rosenzweig, J.M., Sokoll, L.J., van der Zee, A.G., van Haafte-Day, C. Wang, Y.Y., Yu, Y., Zhang, Z. (2004) Three biomarkers identified from serum proteomic analysis for the detection of early stage ovarian cancer. *Cancer Research*, 64. P. 5882-5890.
4. Banda, M., Bommineni, A., Thomas, R. A. Luckinbill, L. S. and Tucker, J. D. (2008) Evaluation and validation of housekeeping genes in response to ionizing radiation and chemical exposure for normalizing RNA expression in real-time PCR. *Mutation Research*, 649. P. 126-134.
5. Bertone, P., Caldas, C., Dvinge, H., Git, A., Hadfield, J., Kutter, C., Osborne, M. and Salmon-Divon, M. (2010) Systemic comparison of microarray profiling, real-time PCR, and next-generation sequencing technologies for measuring differential microRNA expression. *RNA*, 16. P. 991-1006.
6. Blake, R. D. (2006) Denaturation of DNA. *Reviews in Cell Biology and Molecular Medicine*, 9. DOI: 10.1002/3527600906.mcb.200200014.
7. Boyerinas, B., Park, S., Hau, A., Murmann, A. E. And Peter, M. E. (2010) The role of let-7 in cell differentiation and cancer. *Society of Endocrinology*, 10. P. 19-36.
8. Bristow, R.E., Goodrich, S.T., Miller, R.W., Santoso, J.T., Smith, A., Ueland, F.R. and Zhang, Z. (2014) The effect of ovarian imaging on the clinical interpretation of a

- multivariate index assay. *American Journal Obstetrics Gynaecology*, 211. P. 65-76.
doi:10.1016/j.ajog.2014.02.010.
9. Cheng, J., Auersperg, N. And Leung, P. C. K. (2012) TGF-Beta Serous Borderline Ovarian Tumour Cell Invasion by Activating EMT but Triggers Apoptosis in Low-Grade Serous Ovarian Carcinoma Cells. *PLoS One*, 7. P. 1-9.
10. De Lima Reboucas, E., Costa, J. J. D., Passos, M. J., Passos, J. R. D., van den Hurk, R. And Silva, J. R. V. (2012) Real Time PCR and Importance of Housekeeping Genes for Normalization and Quantification of mRNA Expression in Different Tissues. *Brazilian Archives of Biology and Technology*, 56. P. 145-154.
11. Dong, Y., Li, J., Han, F., Chen, H., Zhao, X., Qin, Q., Shi, R. And Liu, J. (2015) High IGF2 expression is associated with poor clinical outcome in human ovarian cancer. *Oncology Reports*, 10. P. 936-942.
12. Farias-Eisner, R., Faull, K., Kozak, K.R., Reddy, S., Su, F. and Whitelegge, J.P. (2005) Characterization of serum biomarkers for detection of early stage ovarian cancer. *Proteomics*, 17. P. 4589-4596.
13. Gianuzzi, X., Palma-Ardiles, G., Hernandez-Fernandez, W., Pasupuleti, V., Hernandez, A. V. And Perez-Lopez, F. R. (2016) Insulin growth factor (IGF) 1, IGF-binding proteins and ovarian cancer risk: Asystemic review and meta-analysis. *Elsevier*, 94. P. 22-29.
14. Godinho, S. A., Kwon, M. And Pellman, D. (2009) Centrosomes and cancer: how cancer cells divide with too many centrosomes. *Cancer metastasis review*, 28. P. 85-98.
15. Gonczy, P. (2015) Centrosomes and cancer: revisiting a long-standing relationship. *Nature review*, 15. P. 639-652.

16. Hills, C. A., Kelland, L. R., Abel G., Siracky, J., Wilson, A. P. and Harrap, K. R. (1989) Biological properties of ten human ovarian carcinoma cell lines: calibration in vitro against four platinum complexes. *British Journal of Cancer*, 4.
17. Hirashima, Y., Kobayashi, H., Suzuki, M., Tanaka, Y., Kanayama, N. And Terao, T. (2003) Transforming Growth Factor-B1 Produced by Ovarian Cancer Cell line HRA Stimulates Attachment and Invasion Through an Up-regulation of Plasminogen Activator Inhibitor Type-1 In Human Peryoneal Mesothelial Cells. *The Journal of Biological Chemistry*, 278. P. 26793-26802.
18. Husseinzadeh, N. (2011) Status of tumour markers in epithelial ovarian cancer has there been any progress? A review. *Gynecologic Oncology*, 120. P. 152-157.
19. Huang, J., Hu, W. And Sood, A. K. (2014) Prognostic Biomarkers in Ovarian Cancer. *Cancer Biomarkers*, 8. P. 231-251.
20. Jansson, M.D. and Lund, A.H. (2012) MicroRNA and Cancer. *Molecular Oncology*, 6. P. 590-610.
21. Karlan, B. Y., Jones, J., Slamon, D. J. And Lagasse, L. D. (1994) Glucocorticoids stabilize HER-2/neu messenger RNA in human epithelial ovarian carcinoma. *Gynecological Oncology*, 53. P. 70-77.
22. Kawashima, K., Jahan, I., Fushimi, K., Mihara, K. and Namba, M. (1995) Human fibroblasts (KMST-6/RAS cell line) transformed with ⁶⁰Co gamma-rays and c-Ha-ras oncogene produce a large amount of granulocyte colony-stimulating factor (G-CSF); production is enhanced by cAMP, theophylline, and butyrate. *Cell Structure and Function*, 20. P. 41-45.
23. Kinose, Y., Sawada, K., Nakamura, K. and Kimura, T. (2014) The Role of MicroRNAs in Ovarian Cancer. *BioMed Research International*, 10. P. 1-14.

24. Koshiyama, M., Matsumura, N. And Konishi, I. (2014) Recent Concepts of Ovarian Carcinogenesis: Type I and Type II. *BioMed Research International*, 2014. P. 1-11.
25. Kowdley, K.V. and Li, Y. (2012) MicroRNAs in Common Human Diseases. *Genomics Proteomics and Bioinformatics*, 10. P. 246-253.
26. Kurman, R.J. and Shih, L. (2010) The Origin and Pathogenesis of Epithelial Ovarian Cancer-a Proposed Unifying Theory. *American Journal of Surgery Pathology*, 34. P. 433-443.
27. Kuroda, H., Mandai, M., Konishi, I., Tsuruta, Y., Kusakari, T., Kariya, M. And Fujii, S. (2001) Human ovarian surface epithelial (OSE) cells express LH/hCG receptors, and hCG inhibits apoptosis of OSE cells via up-regulation of insulin-like growth factor-1. *International Journal of Cancer*, 91. P. 309-315.
28. Lando, D. Y., Fridman, A. S., Chang, C., Grigoryan, I. F., Galyuk, E. N., Murashko, O. N., Chen, C. And Hu, C. (2015) Determination of melting temperature and temperature melting range for DNA with multi-peak differential melting curves. *Analytical Biochemistry*, 479. P. 28-36.
29. Lang, D. And Salvaterra, P. M. (2011) Robust RT-qPCR Data Normalization: Validation and Selection of Internal Reference Genes during Post-Experimental Data Analysis. *PLoS One*, 6. P. 1-10.
30. Li, A. (2012) New biomarkers for ovarian cancer: OVA1 and ROMA in diagnosis. *ModernMedicine NETWORK, Obstetrics-Gynecology& Women's Health*. Available at: <http://contemporaryobgyn.modernmedicine.com/contemporary-obgyn/news/modernmedicine/modern-medicine-feature-articles/new-biomarkers-ovarian-cance?page=full>.
31. Li, Y., Fang, Y., Lui, Y. and Yang, X. (2015) MicroRNAs in ovarian functions and disorders. *Journal of Ovarian Research*, 8. P. 1-8.

32. Li, Y., Li, Y., Zhang, J., Zhu, C., Yu, H. and Fan, L. (2016) Circulating Insulin-Like Growth Factor-1 Level and Ovarian Cancer Risk. *Cellular Physiology and Biochemistry*, 38. P. 589-597.
33. Lima, S. A. and Pasquinelli, A. E. (2014) Identification of miRNAs and their targets in *C. elegans*. *Systems Biology of RNA Binding Proteins*, 852. P. 431-450.
34. Livak, K. J. And Schmittgen, T. D. (2008) Analyzing real-time PCR data by the comparative CT method. *Nature*, 3. P. 1101-1108.
35. Lottering, S (2015) *Identification of miRNAs as biomarkers for early diagnosis of Ovarian Cancer: An in silico approach*. Honours thesis. University of the Western Cape.
36. Makunin , I.V. and Mattick, J.S. (2006) Non-coding RNA. *Human Molecular Genetics*, 15. P. 17-29.
37. Merrit, W. M., Lin, Y. G., Han, L. Y., Kamat, A. A., Spannuth, W. A., Schmandt, R., Urbauer, D., Pennacchio, L. A., Cheng, J., Nick, A. M., Deavers, M. T., Mourad-Zeidan, A., Wang, H., Mueller, P., Lenburg, M. E., Gray, J. W., Mok, S., Birrer, M. J., Lopez-Berestein, G., Coleman, R. L. And Bar-Eli, M. (2008) Dicer, Drosha and outcomes in Patients with Ovarian Cancer. *The New England Journal of Medicine*, 359. P. 2641-2650.
38. Nezhat, F., Datta, M. S., Hanson, V., Pejovic, T., Nezhat, C. And Nezhat, C. (2008) The relationship of endometriosis and ovarian malignancy: a review. *Fertility and Sterility*, 90. P. 1559-1570.
39. Nozaki, T., Sugiyama, S., Koga, H., Sugamura, K., Ohba, K., Matsuzawa, Y., Sumiday, H., Matsui, K., Jinnouchi, H. and Ogawa, H. (2009) Significance of a Multiple Biomarkers Strategy Including Endothelial Dysfunctional to Improve Risk

- Stratification for Cardiovascular Events in Patients at High Risk for Coronary Heart Disease. *Journal of the American College of Cardiology*, 54. P. 601-608.
40. Pabinger, S., Rodiger, S., Kriegner, A., Vierlinger, K. and Weinhausel, A. (2014) A survey of tools for the analysis of quantitative PCR (qPCR) data. *Biomolecular Detection and Quantification*, 1. P. 23-33.
41. Park, Y. T., Jeong, J., Lee, M., Kim, K., Kwon, Y., Lee, C., Kim, O. J. And An, H. (2013) MicroRNAs overexpressed in ovarian ALDH1- positive cells are associated with chemoresistance. *Journal of Ovarian Research*, 6. P. 1-11.
42. Peng, Y. and Croce, C. M. (2010) The role of miRNAs in human cancers. *Signal Transduction and Targeted Therapy*, 1. P. 1-9.
43. Pfaffl, M.W., Horgan, G.W. and Dempfle, L. (2002) Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. *Nucleic Acid Research*, 30.
44. Prime, S. S., Nixon, S. V., Crane, I. J., Stone, A., Matthews, J. B., Maitland, N. J., Remnant, L., Powell, S. K., Game, S. M. And Scully, C. (1990) The behaviour of human oral squamous cell carcinoma in cell culture. *The Journal of Pathology*, 3.
45. Roa, X., Huang, X., Zhou, Z. And Lin, X. (2013) An improvement of the $2^{-\Delta\Delta CT}$ method for quantitative real-time polymerase chain data analysis. *Biostatistics, Bioinformatics and Biomathematics*, 3. P. 71-85.
46. Rohr, I., Zeillinger, R., Heinrich, M., Concini, N., Vergote, I., Nassir, M., Mahner, S., Nieuwenhuysen, E., Trillsch, F., Cacsire-Tong, D., Chekerov, R., Sehouli, J. And Braicu, E. I. (2016) Role of IGF-I in Primary Ovarian Cancer-Astudy of the OVCAD European Consortium, *International Journal of Cancer Research and Treatment*, 3. P. 1015-1022.

47. Rouch, S. and Slack, F. J. (2008) The let-7 family in microRNAs. *Trends in Cell Biology*, 18. P. 505-516.
48. Rupaimoole, R., Wu, S. Y., Pradeep, S., Ivan, C., Pecot, C. V., Gharpure, K. M., Nagaraja, A. S., Armaiz-Pena, G., N., McGuire, M., Zand, B., Dalton, H. J., Filant, J., Miller, J. B., Lu, C., Sadaoui, N. C., Mangala, L. S., Taylor, M., van den Beucken, T., Koch, E., Rodriguez-Aguayo, C., Huang, L., Bar-Eli, M., Wouters, B. G., Radovich, M., Ivan, M., Calin, G. A., Zhang, W., Lopez-Berestein, G. And Sood, A. K. (2014) Hypoxia Mediated Downregulation of miRNA Biogenesis Promotes Tumour Progression. *Nature Communications*, 5. P. 1-26.
49. Scherer, W. F., Syverton, J. T. And Gey, G. O. (1953) Studies on the propagation in vitro of poliomyelitis viruses. *The Journal of Experimental Medicine*, 97. P. 695-710.
50. Schmittgen, T. D. And Livak, K. J. (2008) Analyzing real-time PCR data by the comparative CT method. *Nature*, 3. P. 1102-1108.
51. Seki, S., Kawaguchi, Y., Chiba, Y., Mikami, Y., Kizawa, H., Oya, T., Mio, F., Miyamoto, Y., Masuda, I., Tsunoda, T., Kamata, M., Kubp, T., Toyama, Y., Kimura, T., Nakamura, Y. and Ikegawa, S. (2013) A functional SNP in CILP, encoding cartilage intermediate layer protein, is associated with susceptibility to lumbar disc disease. *Nature Genetics*, 6. P. 607-612.
52. Slack, F.J. and Stahlhut Espinosa, C.E. (2006) The Role of MicroRNAs in Cancer. *Yale Journal of Biology and Medicine*, 79. P. 131-140.
53. Soule, H. D., Vazquez, J., Long, A., Albert, S. and Brennan, M. (1973) A human cell line from pleural effusion derived from a breast carcinoma. *Journal of the National Cancer Institute*, 51. P. 1409-1416

54. Turabelidze, A., Guo, S. and DiPietro, L. A. (2011) Importance of Housekeeping gene selection for accurate RT-Qpcr in a wound healing model. *Wound repair generation*, 10. P. 460-466.
55. Wijnhoven, B. P. L., Michael, M. Z. And Watson, D. I. (2007) MicroRNAs and cancer. *British Journal of Surgery*, 10. P. 23-30.
56. Wong, M.L., and Medrano, J.F. 2005. Real-time PCR for mRNA quantitation. *BioTechniques*, 39. P. 1-11.
57. Yuan, J.S., Reed, A., Chen, F., and Stewart Jr, C.N. 2006. Statistical analysis of realtime PCR data. *BMC Bioinformatics*, 7. P. 1-1.
58. Yeung, T., Leung, C. S., Wong, K., Samimi, G., Thompson, M. S., Liu, J., Zaid, T. M., Ghosh, S., Birrer, M. J. And Mok, S. C. (2013) TGF-B modulates ovarian cancer invasion by upregulating CAF-derived versican in the tumour microenvironment. *Cancer Research*, 73. P. 5016-5028.
59. Yue, J . and Tigy, G. (2006) MicroRNA trafficking and human cancer. *Cancer Biology and Therapy*, 5. P. 573-578.
60. Zhang, S., Lu, Z., Unruh, A. K., Ivan, C., Beggerly, K. A., Calin, G. A., Li, Z., Bast Jr, R. C. And Le, X. (2015) Clinically Relevant microRNAs in Ovarian Cancer. *Molecular Cancer Research*, 13. P. 393-401.
61. Zyss, D. And Gergely, F. (2009) Centrosome function in cancer: guilty or innocent?. *Trends in cell biology*, 19. P. 334-346.

Chapter 5

General discussion and future directions

5.1 General discussion

5.2 Future work

5.3 Reference List



Chapter 5

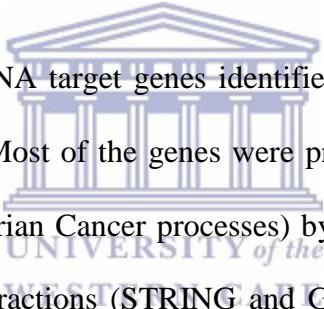
General discussion and future directions

5.1 General discussion

Cancer is a complex disease where cells in a specific part of the body such as the ovary begin to reproduce uncontrollably. It is one of the leading causes of morbidity and mortality in women, worldwide. In 2015, there were 8.8 million cancer-related deaths globally. The number of cancer cases is expected to increase by 70 % over the next decade (World Health Organization, 2016). Cancer mortality cases can be reduced or avoided through early diagnosis and treatment. However, the lack of specific symptoms in the early stages of the disease and the limited understanding of the disease development and progression at a molecular level makes early diagnosis difficult (National Cancer Institute, 2015; Bristow *et al.*, 2014).

Another reason for the increased mortality rates in cancer patients is due to the lack of clinically available diagnostic tools with adequate sensitivity and specificity for early detection (National Cancer Institute, 2015). Also, some of the diagnostic tools are invasive, ineffective and expensive. In Ovarian Cancer, if detected and treated while the cancer cells are still localized (Stage IA and B), the 5-year survival rate is above 92 %. However, if the cancer cells have spread to different organs and tissues, the 5-year survival rate is less than 17 % (Cancer Research UK, 2016). Therefore, adequate diagnostic tools are required for early diagnosis of cancer specifically Ovarian Cancer (Konishi, Koshiyama and Matsumura, 2014).

The purpose of the study was to discover biomarkers for Ovarian Cancer to aid in the early diagnosis of the disease and overcome the limitations of the current diagnostic tools. Previously *in silico* methodologies were employed to identify candidate microRNAs that have not previously been described nor showed association to Ovarian Cancer as well as microRNA target genes associated with the initiation and progression of Ovarian Cancer (Lottering, 2015). The study identified six miRNAs and seven miRNA target genes from a large number of miRNAs and target genes that were verified through subsequent steps of elimination using *in silico* methods and thus were confidently selected for further bioinformatics analysis and molecular validation.



In the present study, the microRNA target genes identified were functionally characterized using various *in silico* analysis. Most of the genes were predicted to be involved in cancer-related processes (especially Ovarian Cancer processes) by Gene Ontology (Gene Ontology Consortium), protein-protein interactions (STRING and GeneMANIA), transcription factor analysis (GeneCards and TFactS®) and pathway analysis (Reactome) (Chapter 3). This section of the study successfully implicated the miRNA target genes in cancer, based on their regulatory elements (TFs), interacting proteins and specific pathway involvement analysis. Furthermore, this subsequently linked the miRNAs that target those genes to processes implicated in cancer specifically Ovarian Cancer.

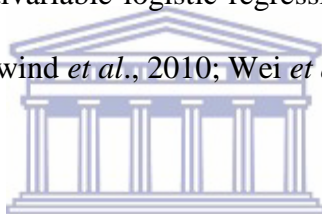
Furthermore, *in silico* prognostic/predictive analysis of the miRNA target genes was done (Chapter 4). Predictive validation showed that most of the target genes showed prognostic significance based on the Kaplan-Meier plots generated and the p-values (p value < 0.05) obtained for analysis. Two out of the seven target genes (CILP and FARP1) were common in

multiple databases used for the prognostic studies. These genes have also been linked to various factors and processes that result in poor prognosis of Ovarian Cancer as described in Section 2.4.

In essence, computational biology in conjunction with molecular biology provides a powerful combination in completely understanding the potential biomarkers identified. Our study shows a unique expression profile for each miRNA across various cancer and non-cancer cell lines investigated. In this study, a panel of cancer cell lines were investigated to determine a unique miRNA expression profile for each miRNA. The study revealed that the miRNAs were differentially expressed across all cancer cell lines compared to the non-cancerous cell line (KMST-6). One miRNA was shown to be a potential diagnostic biomarker in Ovarian Cancer (miR1) based on the expression profile, though it has been suggested that one biomarker is not sufficient to diagnose such a complex disease (Agarwal *et al.*, 2011). This miRNA targets CILP as shown in Table 2.1. In Section 3.5, CILP has been linked to poor prognostic patient outcome in various databases. Also, in a previous study, it was shown that the target gene was preferentially expressed in the ovary compared to other tissues (Lottering, 2015). However, it is expressed at lower levels in tumour tissue. It can be postulated that the miRNA down-regulates CILP in ovarian tumours. However, further analysis is required to determine the relationship between miR1 and CILP. MiR2 and miR4 were down regulated in Ovarian Cancer (CoaV-3) as well as in Cervical Cancer (HeLa). However, to increase the predictive accuracy of the candidate biomarkers, it could be used in combination with the current predictive biomarkers or known biomarkers used within the clinical setting (Huang, Hu and Sood, 2014; Nozaki *et al.*, 2009).

5.2 Future work

Further bioinformatics analysis should be done between the identified miRNA targets and known clinical biomarkers for Ovarian Cancer. Also, bioinformatics should be employed to further understand the identified target genes and identify other relations it might have to cancer, specifically Ovarian Cancer. Using *in silico* methods to discover novel biomarkers should be a continuous process as databases are regularly updated. Also, further prognostic analysis using both *in silico* and molecular approaches will be evaluated to conclusively determine the prognostic significance of the miRNA target genes. The prognostic significance of the miRNAs can also be determined using various molecular processes and statistical analysis including multivariable logistic regression models and multivariable Cox proportional hazards models (Schwind *et al.*, 2010; Wei *et al.*, 2013)



UNIVERSITY of the
WESTERN CAPE

The study recognized the need for additional cell lines to confirm the specificity of the miRNAs identified. Future experiments can be done on urine, saliva and tissue samples to determine their efficiency in biological samples as potential biomarkers in Ovarian Cancer. A molecular approach can be employed to evaluate the gene expression patterns of the candidate target genes identified in cancerous and non-cancerous cell lines. The objective would be to compare the expression profiles of the miRNAs and the target genes to determine deregulation in a specific miRNA and its target gene. Experimental validation will be done on the deregulated miRNA and target gene such as the luciferase assay to determine if the miRNA do target the predicted target gene (s).

Further studies could include the exploration of nano-diagnostics by using nano-devices. Biomarkers present in biological fluids exist in small quantities therefore it might be masked by other proteins. Nanotechnology can ensure the detection of these miRNA biomarkers at a nano-scale by enhancing the throughput and sensitivity of the identification and screening of the potential biomarkers (Hu *et al.*,2011). Therefore, the combination of nanotechnology in the form of gold nanoparticles could be used in the development of a lateral flow device for the early detection/diagnosis of Ovarian Cancer as shown in Figure 6.1. This has the potential to be developed into a diagnostic test that is cost effective, non-invasive and it can be rapidly adapted into clinical practice (Sharma *et al.*, 2015; Sajid, Kwade and Duad, 2014).

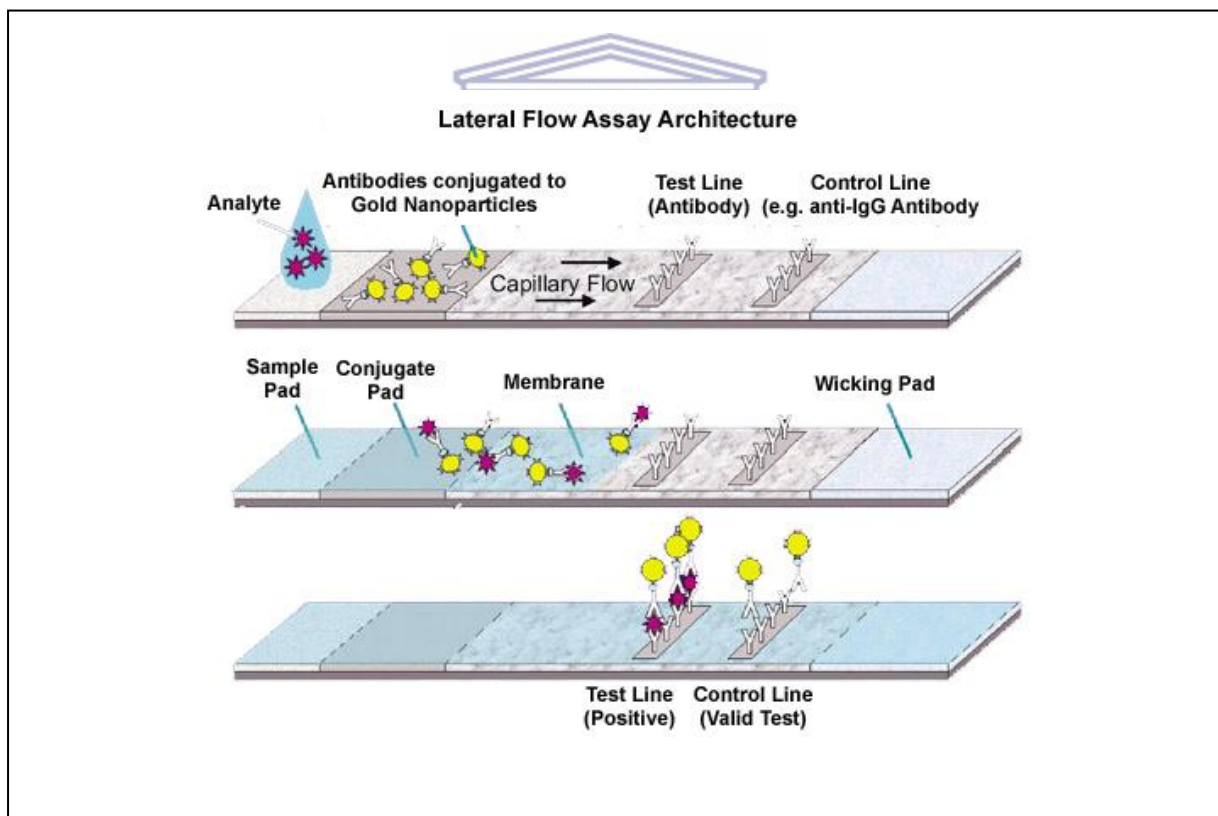
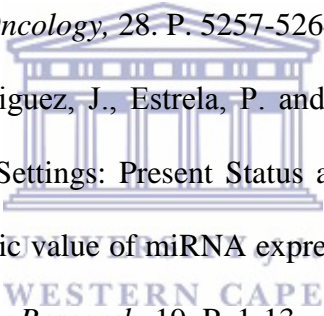


Figure 5.1: An example of a lateral flow device that can be developed for the detection of Ovarian Cancer using the discovered miRNAs (Sajid, Kwade and Duad, 2014).

5.3 Reference List

1. Agarwal, A., Dada, R., Gupta, S., Michener, C., Rein, B.J.D and Safi, J. (2011) Potential Markers for Detection and Monitoring of Ovarian Cancer. *Journal of Oncology* 10, 1 – 17. doi:10.1155/2011/475983.
2. Bristow, R.E., Goodrich, S.T., Miller, R.W., Santoso, J.T., Smith, A., Ueland, F.R. and Zhang, Z. (2014) The effect of ovarian imaging on the clinical interpretation of a multivariate index assay. *American Journal Obstetrics Gynaecology* 211. P. 65-76.
3. doi:10.1016/j.ajog.2014.02.010.
4. Cancer Research UK. (2016) Stages of ovarian cancer. <http://www.cancerresearchuk.org/about-cancer/type/ovarian-cancer/treatment/stages-of-ovarian-cancer>.
5. Hu, Y., Fine, D. H., Tasciotti, E., Bouamrani, A. and Ferrari, M. (2011) Nanodevices in diagnostics. *Nanomedicine and NanoBiotechnology*, 3. P. 11-32.
6. Huang, J., Hu, W. And Sood, A. K. (2014) Prognostic Biomarkers in Ovarian Cancer. *Cancer Biomarkers*, 8. P. 231-251.
7. Konishi, I., Koshiyama, M. and Matsumura, N. (2014) Recent Concepts of Ovarian Carcinogenesis: Type I and Type II. *BioMed Research International*, 2014. P. 1-11. doi:10.1155/2014/934261.
8. Lottering, S (2015) *Identification of miRNAs as biomarkers for early diagnosis of Ovarian Cancer: An in silico approach*. Honours thesis. University of the Western Cape.
9. National Cancer Institute (2015) Ovarian Epithelial, Fallopian Tube, and Primary Peritoneal Cancer Treatment (PDQ®)-Patient Version. National Institute of Health. <https://www.cancer.gov/types/ovarian/patient/ovarian-epithelial-treatment-pdq>.

10. Nozaki, T., Sugiyama, S., Koga, H., Sugamura, K., Ohba, K., Matsuzawa, Y., Sumiday, H., Matsui, K., Jinnouchi, H. and Ogawa, H. (2009) Significance of a Multiple Biomarkers Strategy Including Endothelial Dysfunctional to Improve Risk Stratification for Cardiovascular Events in Patients at High Risk for Coronary Heart Disease. *Journal of the American College of Cardiology*, 54. P. 601-608.
11. Sajid, M., Kwade, A. and Duad, M (2014) Design, formats and applications of lateral flow assay: A literature review. *Journal of Saudi Chemical Society*, 19. P. 689-705.
12. Schwind, S., Maharray, K., Radmacher, M. D., Mrozek, K., Holland, K. B. And Margeson, D. (2010) Prognostic significance of expression of a single miRNA, miR-181a, in cytogenetically normal acute myeloid leukemia: A Cancer and Leukemia Group B Study, *Clinical Oncology*, 28. P. 5257-5264.
13. Sharma, S., Zapatero-Rodriguez, J., Estrela, P. and O'Kennedy, R. (2015) Point-of-care Diagnostics in Low Settings: Present Status and Future Role of Microfluidics. *Biosensors*, 5. P. 577-601.  ic value of miRNA expression signatures in Hepatocellular carcinoma, *Clinical Cancer Research*, 19. P. 1-13.
14. Wei, R., Huang, G., Zhang, M., Li, B., Zhang, H., Shi, M., Chen, X., Huang, L., Zhou, Q., Jia, W., Zheng, S., Yuan, Y. and Wang, H. (2013) Clinical Significance and Prognostic Value of microRNA Expression Signatures in Hepatocellular Carcinoma. *Clinical Cancer Research*, 28. P. 1-13. DOI: 10.1158/1078-0432.CCR-12-2728.