IDENTIFICATION OF miRNA'S AS SPECIFIC BIOMARKERS IN PROSTATE CANCER DIAGNOSTICS : A COMBINED *IN SILICO* AND MOLECULAR APPROACH

Firdous Khan

Thesis presented in fulfilment of the requirements for the

Degree *Philosophiae* Doctor

at the Faculty of Natural Sciences, University of the Western Cape

Advisor: Dr Ashley Pretorius

Co-Advisor: Prof. M. Meyer

May, 2015

ABSTRACT

Cancer is a class of diseases characterized by uncontrollable cell growth. There are over 100 different types of cancer, and each of these cancers are classified by the type of cell that it initially affects. For the purpose of this research we will be focussing on prostate cancer (PC). Prostate cancer is the second most common form of cancer in men around the world and annually approximately 4500 men in South Africa are diagnosed making PC a global epidemic. Prostate cancer is a type of cancer which starts in the prostate it is normally a walnut-sized gland found right below the bladder. PC follows a natural course, starting as a tiny group of cancer cells that can grow into a tumour. In some men if PC is not treated it may spread to surrounding tissue by a process called direct invasion/ spread and could lead to death.

Current diagnostic tests for prostate cancer have low specificity and poor sensitivity. Although many PC's are slow growing there is currently no test to distinguish between these and cancers that will become aggressive and life threatening.. Therefore the need for a less invasive early detection method with the ability to overcome the lack of specificity and sensitivity of current available diagnostic test is required. Biomarkers have recently been identified as a viable option for early detection of disease for example

biological indicators ie. DNA, RNA, proteins and microRNAs (miRNAs).

Since first described in the 1990s, circulating miRNAs have provided an active and rapidly evolving area of research that has the potential to transform cancer diagnostics and prognostics. In particular, miRNAs could provide potentially new biomarkers for PC as diagnostic molecules. Circulating miRNAs are highly stable and are both detectable and quantifiable in a range of accessible bio-fluids, having the potential to be useful as diagnostic, prognostic and predictive biomarkers. In this study we aimed to identify miRNAs as potential biomarkers to detect and distinguish between various types of PC in its earliest stage. The major objectives of the study were to identify miRNAs and their gene targets that play a critical role in disease onset and progression to further understand their mechanism of action in PC using several *in silico* methods, and to validate the potential diagnostic miRNAs using qRT-PCR in several cell lines.

The identification of specific miRNAs and their targets was done using an "in-house" designed pipeline. Bioinformatic analyses was done using a number of databases including STRING, DAVID, DIANA and mFold database, and these combined with programming and statistical analyses was used for the identification of potential miRNAs specific to PC. Our

study identified 40 miRNAs associated with PC using our "in-house" parameters in comparison to the 20-30 miRNAs known to be involved in PC found in public databases e.g. miRBase. A comparison between our parameters and those used in public databases showed a higher degree of specificity for the identification PC-associated miRNAs.

Gene targets for these 40 miRNAs were identified using the publicly available TargetScan database. A total number of 785 targets were identified. The target gene list was streamlined by eliminating duplicates using a Perl script, resulting in a final target list of 551 genes. The final target gene list was subjected to Functional Annotation and cluster analyses using DAVID. Seventy seven clusters were generated, 3 of which were identified as clusters of processes involved in cancer onset as well as progression. These clusters contained a total of 21 gene targets. These gene targets were then correlated to the miRNAs identified by our "in-house" technique. From the 40 miRNAs 13 were found to be of interest based on Gene Ontology, Functional annotation and cluster analysis of their targeted genes, of which 8 miRNAs (MiR1-8) were tested during subsequent sections of this study.

These selected miRNAs were analysed using different bioinformatics tools, and were confirmed to be novel miRNAs associated with PC.

The identified miRNAs were experimentally validated using qRT-PCR to generate expression profiles for PC as well as various other cancers. Prostate lines utilised in this study included PNT2C2 (normal) which was compared to BPH1 (Benign) and LNCaP (Metastatic). In the study the expression profiles of eight potential miRNA biomarkers for the detection of PC was determined using qRT-PCR, and to distinguish PC from other cancers.

QRT-PCR data showed that miRNA-3 and -5 were up-regulated in the BPH1 and LNCaP when compared to PNT2C2. In addition miRNA-8 was also shown to be up-regulated in LNCaP. Based on these results it was shown that a miRNA profile could be established to distinguish between BPH1 and the LNCaP prostate cell lines. The results suggest that one miRNA as a diagnostic marker may be sufficient to differentiate between different cancer cell lines. Furthermore by creating a unique profile for each cancer cell line by using a combination of miRNAs could be a suitable approach as well.

Finally, it was shown that through the use of a single or combination of all eight miRNAs a unique profile for all the cancer cell lines tested in this study can be created. This is an important finding which could have potential diagnostic or prognostic implications in clinical practice.

Keywords: Prostate cancer, miRNA, biomarker discovery, diagnostics, early detection, cancer, specificity, sensitivity, qRT-PCR, Expression, Bioinformatics.

DECLARATION

I declare that "**Identification of miRNA's as specific biomarkers in prostate cancer diagnostics : A combined *in silico* and molecular approach**" is my own work, that it has not been submitted for degree or examination at any other university, and that all the resources I have used or quoted, and all work which was the result of joint effort, have been indicated and acknowledged by complete references.

_____

Signed August 12 2015

Table of Contents

DEDICATION

Dedicated to my Mum and Dad

Zubayda Khan

&

Abubaker Khan

ACKNOWLEDGEMENTS

The price of success is hard work, dedication to the job at hand, and the determination that whether we win or lose, we have applied the best of ourselves to the task at hand.

Vince Lombardi

All praise and glory be to the Almighty ALLAH (saw) my creator without whom non of this would be possible. To my supervisor Dr Ashley Pretorius I am eternally grateful to you for always being a pillar of strength to me. You were not only a mentor to me but a best friend that I will cherish till the end. I am indebted to you for your enthusiasm towards my progress and for believing in me, It has been a privilege to work under your supervision throughout this journey. Prof Mervin Meyer thank you very much for the opportunities you have provided me with during the past few years . Mr Faghri February words cannot describe how thankful I am for the time and effort you spent on me to give me an invaluable learning experience, I could not have asked for a better teacher, mentor and friend, shukran. Bridget Daniels thank you very much for being such an inspiration to me I have learned so many life lessons from you. My dearest friend Moosa Joseph thank you so much the constant motivation and for the gift of friendship I will cherish it always and for ever. To our research group I am thankful that I have had the privilege to work with so many brilliant minds. To my

Husband, Imthiaz Shaik, I am eternally grateful to you for all the love, motivation and patients you have blessed me with during my journey no words can describe the gratitude I feel knowing I have a pillar of support now and always. To my family shukran to all of you for pushing me to realise my dreams, I am eternally grateful. Each and every one of you have had a part in sculpturing and challenging me to become the best that I can be I Thank You all.....

LIST OF ABBREVIATIONS

/min	per minute
%	Percent
3UTR	three prime untranslated region
A549	Human alveolar basal epithelial malignant cells
ADNP	Activity-dependent neuroprotector homeobox
AIPC	PDZ domain containing 2
AR	Androgen Receptor
AR-Negative	Androgen receptor Negative
ARCR1C	Activin A receptor, type IC
ATCC	American Type Culture Collection
ATM	Ataxia telangiectasia mutated
BAK1	BCL2-antagonist/killer 1
BCL1	B-cell CLL/lymphoma 1
BCL2	B-cell CLL/lymphoma 2
BFAR	Bifunctional Apoptosis Regulator
BPH1	Benign Prostatic Hyperplasia 1
BPH1	Benign Prostatic Hyperplasia

BTG2	Basal protein B-cell translocation gene 2
CaOV3	Human Ovarian Cancer cell line
Caski	Human cervical carcinoma
cDNA	complementary Deoxyribonucleic Acid
CFLAR	CASP8 and FADD-like apoptosis regulator
CLN8	Ceroid-lipofuscinosis, neuronal 8
CO2	Carbon Dioxide
CRB3	crumbs homolog 3 (Drosophila)
CRPC	Castrate Resistant Prostate Cancer
CSRNP3	Cysteine-serine-rich nuclear protein 3
DAVID	Database for Annotation, Visualization and Integrated Discovery
DIANA	DNA intellegent analysis
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic Acid
DRE	Digital Rectal Examination
DVL2	dishevelled, dsh homolog 2 (Drosophila)
EIF4EBP1	eukaryotic translation initiation factor 4E binding protein 1

F11R	F11 receptor
FOXC1	Forkhead box C1
g	gravitational force
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
GEA	Gene experession atlas
GEO	Gene expression omnibus
HepG2	Hepatocellular carcinoma
HT29	Human Colon adenocarcinoma cells
ING4	Inhibitor of growth family, member 4
K-RAS	Kirsten rat sarcoma viral oncogene homolog
LIG4	Ligase IV, DNA, ATP-dependent
LNCaP	Lymph Node Prostate adenocarcinoma
LNCaP	AR-sensitive human prostate adenocarcinoma
MIAME	Minimum information about a *microarray* experiment
MAPK	mitogen-activated protein kinase-activated protein kinase
MCF12aa	Human mammary epithelial cells 12a
MCF7	Human Breast Adenocarcinoma 7

mg	milligrams
miRNA	micro Ribonucleic Acid
ml	millilitres
mm	millimetres
mm^2	millimetres squared
MMP7	matrix metallopeptidase 7
MNT	MAX binding protein
mRNA	Messenger ribonucleic Acid
NFκB	nuclear factor kappa beta
NGS	Next generation sequencing
ºC	Degrees Celcius
ORF	Open Reading Frame
PAK7	p21 protein (Cdc42/Rac)-activated kinase 7
PC	Prostate Cancer
PCR	Polymerase Chain Reaction
PNT2C2	Prostate Normal Epithelium cells
PNT2C2	Human prostate Epithelial
PRKCI	Protein kinase C, iota

PSA	Prostate Specific Antigen
PTEN	Phosphatase and tensin homolog
PTPRC	Protein tyrosine phosphatase, receptor type, C
qRT-PCR	Quantitative Real- Time Polymerase Chain Reaction
RAB27A	RAB27A, member RAS oncogene family
RISC	RNA-induced silencing complex
RNA	Ribonucleic Acid
s	seconds
SH3RF1	SH3 domain containing Ring Finger
siRNA	Small interference Ribonucleic Acid
Tm	Melting Temperature
TMX1	Theoredoxin-related transmembrane protein
TNFSF13B	Tumour necrosis factor superfamily 13B
TNFSF15	Tumor necrosis factor (ligand)

	superfamily, member 15
UCSC	Univerity of California Santa Cruz
URL	Uniform Resource locator
UTR's	Untranslated regions
WNT3A	wingless-type MMTV integration site family, member 3A
YWHAZ	Tyrosine 3-monooxygenase/ tryptophan 5- monooxygenase activation protein, zeta polypeptide

LIST OF FIGURES

Chapter 1 **Pg**

LIST OF TABLES **Pg**

CHAPTER1

Introduction

1. Hallmarks of Cancer

1.1 Genetics of cancer

1.2 Genes responsible for genetic malfunction

1.2.1 Oncogenic gain of function mutations from proto-oncogenes

1.2.2 Oncogenesis gain of function from mutated tumour supressor genes

1.2.3 DNA mismatch repair genes

1.2.4 Model of Carcinogenesis in HNPCC and Sporadic tumours with MMR Deficiency

1.3 What is the prostate?

1.3.1 What is Prostate cancer

1.3.2 Genetics of prostate cancer

1.4 Risk factors involved in prostate cancer

1.4.1 Age

1.4.2 Nationality and Ethnicity

1.4.3 Diet and Lifestyle

1.4.4 Genetics

1.5 Staging of Prostate Cancer

1.6 Signs and symptoms early and late warning signs of cancer

1.7 Screening and diagnosis of prostate cancer

1.7.1 Digital Rectal Examination (DRE)

1.7.2 The Prostate Specific Antigen (PSA) test

1.7.3 Transrectal ultrasound and biopsy

1.7.4 Computed tomography (CT) scan

1.7.5 Magnetic Resonance Imaging (MRI)

1.8. Biomarkers

1.8.1 Biomarker discovery using Gene expression profiling technology

1.9. The need for bioinformatics in Biomarker discovery

1.10. The importance of MicroRNAs

1.10.1 Biogenisis of miRNAs

1.10.2 miRNA and cancer stem cells in prostate cancer

1.10.3 The role of miRNAs in the processes of Epithelial-to-Mesenchymal Transition (EMT) in PC

1.10.4 The relationship between circulating miRNAs and tumour progression in prostate cancer

1.11 The specific roles of certain miRNAs in prostate cancer

1.11.1 Tumour suppressor miRNAs

1.11.2 tumour promoter oncomiRNAs

1.12. Conclusions and perspectives

CHAPTER1

Introduction

Cancer can broadly be described as a group of more than 200 different types of diseases that has the ability to affect any cell in any organ of the body (Fayed, 2009). It is as a result of a series of molecular events that ultimately compromise the normal functioning or properties of that specific cell in a specific part of the body (Hartwell, 2012). In the case of normal cells they are subjected to signals that would allow the cell to divide, differentiate into another cell or perish (Hejmadi, 2010). Where as in cancer cells these processes are hindered as a result of mutated genes which lead to uncontrollable cell growth and eventually tumour formation.

The general principle on which modern cancer biology rests, is the change of these normal cells as a result of changes in the networks on the molecular, biochemical and cellular level and on each level there is a number of ways disturbances can occur (Hejmadi, 2010). General disturbances in the normal processing and pathways in the cell with regards to cancer can result in the loss of normal cell growth, of which the normal tissues have a constant controlled balance of new cell growth and the death of old cells (American Cancer Society, 2012). In normal tissues,

the rates of new cell growth and cell death are kept in balance, and when this balance is compromised or altered it results in the uncontrolled growth of cells with the loss of the cells ability to undergo apoptosis or cell suicide mechanisms which are used by the cell to self-destruct and maintain cell homeostasis (Kleinsmith *et al.,* 2011).

Abnormalities which make cancer cells different from normal cells usually result from protein coding genes that govern cell division or genes which are linked to cell division (Thiel, 2010). As time progresses, these genes which play a pivotal role in repairing DNA damage, become more mutated and may become non- functional. It then cascades from mutations within the parent cell and is passed on to the daughter cells. Some mutated cells die but some of them have a selective advantage allowing it to multiply more rapidly than the normal cells, which is descriptive of cancer cells. These cells show enhanced growth that have gained functions usually repressed in normal cells (Thiel, 2010).

Cancer cells have the ability to develop a degree of self-sufficiency which aid in their proliferation and uncontrolled growth, if left to follow its natural course it would spread and lead to an almost definite case of fatality. Data shows that 90% of cancer-related deaths are as a result of

metastasis- tumour spreading (Hejmadi, 2010). The changes that these cells inherit allow the cell and its progeny to grow and divide in the presence of normal cells. In the presence of cancer cells these inherited changes would naturally inhibit the growth of cells which are in close proximity and therefore the ability to do so allow them to spread and invade other organs (Thiel, 2010).

As a result of the increase in cell proliferation more cells keep dividing and the division of the excess amount of the cells, which is not needed by the body then forms a mass of abnormal tissue referred to as a tumour (ADSTR, 2013). Tumours are classified into two forms, which can either be malignant or benign. A malignant tumour can be described as a cancerous tumour, that has the ability to invade and cause damage to nearby tissues. Due to their cancerous nature the cells metastasise by breaking away and entering the bloodstream and lymphatic system, causing havoc by forming tumours elsewhere in the body (MHS, 2010). A benign tumour is not cancerous, and are usually removable and also do not possess the ability to spread to other parts of the body and spread to other tissues (MHS, 2010).

Tumours are classified according to their origin and tissue types, with

origins as organs such as the liver and breast amongst others. Classification on the basis of the type of cancer are divided into 5 major categories, they include **(a)** carcinomas which is cancer starting in the epithelial cells that form a lining internally or externally of organs inside the body. There are many types of carcinomas two of the most common types include **(i)** adenocarcinomas which develops in an organ or gland and **(ii)** squamous cell carcinomas which affect the squamous epithelium (Mandal, 2010). The **(b)** sarcomas are cancers that stems from the connective and supportive tissues which include muscles, bones, cartilage and fat (Fendler, 2011). The **(c)** myelomas are the cancers' that arise in the plasma cells of bone marrow and is classified as a type of blood cancer (Hejmadi, 2010). **(d) Leukaemias** are types of cancer that are clustered under the blood cancers group (Fendler, 2011). **(e)** Lymphomas are lymphatic system cancers that affect the lymph nodes in specific areas of the body (Kleinsmit, *et al.,* 2011).

1. Hallmarks of Cancer

Functional capabilities that categorise cancer cells include self-sufficiency in growth signals, insensitivity to anti-growth signals, the ability to evade programmed cell-death. They also invade tissue and metastasise to other

organs in the body, have replicative potential and the ability to form their own blood vessels for nutritive purposes (Negrini *et al.,* 2010). Genomic instability was also considered as one of the hallmarks of cancer however it is now considered separately because it is not a functional capability of cancerous cells but a property that gives rise to the hallmarks of cancer as a result of acquisition (Negrini *et al.,* 2010). Additional hallmarks are related to the presence of stress in cancer cells such as DNA damage and DNA replication, oxidative stresses, proteotoxic stresses and metabolic stresses (Negrini *et al.,* 2010). Five new hallmarks were also added and is qualitatively different from the original hallmarks in that they do not describe functional capabilities of cancers but rather the state of cancer cells which are characterised by the presence of the aforementioned stresses.

Sporadic and inherited cancers

There are two main categories of cancer (I) sporadic cancers and (ii) hereditory cancers (Figure 1.1). Ninety to ninety fiver percent of cancers are sporadic with the remaining 5-10% being hereditory (Lee and Muller.2010)

Sporadic cancers occur by chance in certain cells as a result of a mutation

(Darrough, 2008). Individuals who develop sporadic cancers did not inherit the cancer from parent cells thus only the tumour cells contain the mutations which cannot be passed on to future generations (Adams, 2011). Hereditary cancer is inherited from an already mutated parent gene inside the ovum or sperm cell, the gene is then passed down at the time of conception (Darrough, 2008) These germline mutations predispose members of a family to the onset of a particular cancer(Adams, 2011).

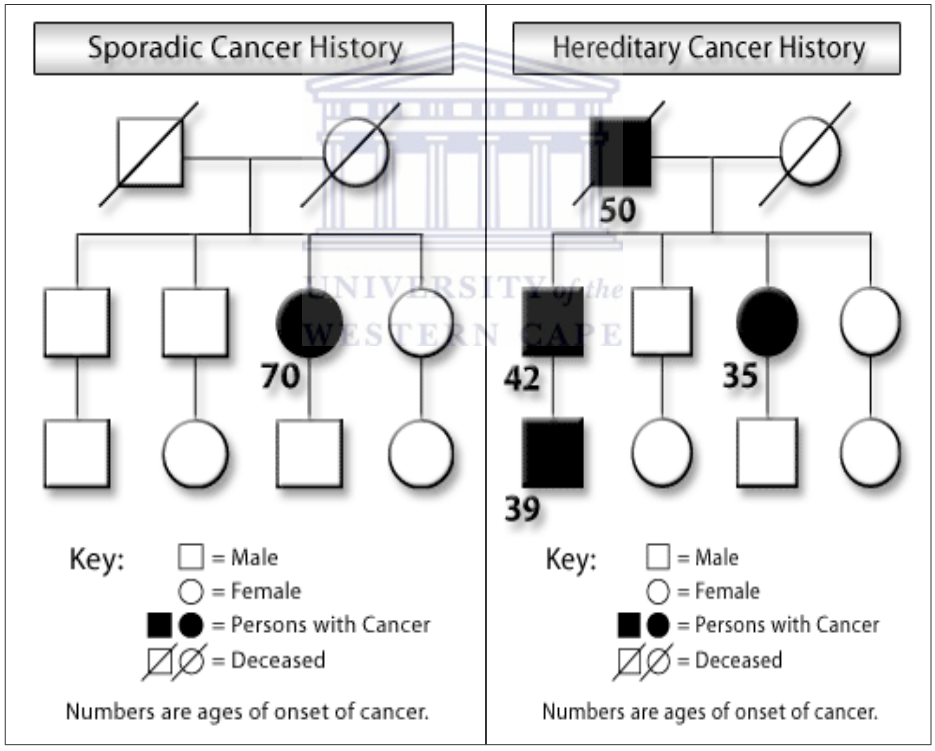


Figure 1.1: Graphical representation of the difference between sporadic and hereditary cancer history. Where the image on the left

shows the sporadic occurrence of cancer in a single individual. The image on the right shows hereditary cancer where mutated genes are passed on to future generations increasing the risk of cancer development in downstream generations (Taken from Darrough 2008)\

1.1 Genetics of cancer

A normal cell needs to undergo several changes in order to become a cancer cell. Of these changes, mutational changes are considered one of the key alterations. Because of these changes, cancer can therefore be classified as a genetic disease on the cellular and inheritance level (Ponder, 1992). It is estimated that about 35000 genes in the human genome have been known to have some sort of association with various cancers (Thiel, 2010).

Mutations in genes result in altered proteins, such changes could occur during cell division or as a result of external agents or a random event. Most cancers result from mutations in somatic cell or in germline cells (Yuluğ, 2012). A germline mutation is any detectable and heritable variation in the lineage of germ cells. Mutations in these cells are transmitted to offspring. Mutations of this nature gives rise to a constitutional mutation in the offspring, that is, a mutation that is present in

virtually every cell. A constitutional mutation can also occur very soon after fertilisation, or continue from a previous constitutional mutation in a parent (Hanahan and Weinberg., 2011). Somatic mutations however are described as a change in the genetic structure that is not inherited from a parent, and also not passed to offspring (Capaccio et al., 2010). The distinction between germline and somatic mutations is important in animals that have a dedicated germ line to produce reproductive cells. These mutations are not only known to lead to cancer but are also seen as precursors for disease as they can also be passed down to subsequent generations (Park and Vogelstein, 2003).

1.2 Genes responsible for genetic malfunction

Malfunctioning genes have been classed into two main groups called oncogenes and tumour suppressor genes. Tumour suppressor genes are genes whose gene products inhibit cell division or cause cell death. Tumour suppressor genes with a subtype of mismatch DNA repair genes. DNA mismatch repair genes play a role in preventing mutations which could lead to cancer (Hejmadi, 2010). Oncogenes play a role in the production of proteins that enhances cell proliferation or inhibits cell death. The

aforementioned genes usually play a role in the maintenance of normal genomic stability and are involved in the regulation of the cell cycle (Devi, 2005).

1.2.1 Oncogenic gain of function mutations from proto-oncogenes

Proto-oncogenes are gatekeeping genes whose products control cell proliferation. They mutate and are converted to oncogenes this may cause a change in the protein structure, causing an increase in protein (enzyme) activity or a loss of regulation and thus and renders the gene in a state of continuous activation (American Cancer Society, 2012). They also encode growth promoting signal molecules and their antiapoptotic receptor proteins with transcription factors genes (Park and Vogelstein, 2003).

There are several mechanisms by which proto-oncogenes are converted to a oncogenes. These include point mutations, chromosomal translocation and amplification. During conversions by point mutations a single base pair is mutated in a proto-oncogene. The result of this mutation leads to a protein whose product is constantly being produced (Thiel, 2010). Chromosomal translocation is divided into two parts that make them different from point

mutations, the conversion is the first part which is the fusion of two genes resulting in a hybrid gene this gene then encodes for a chimeric protein. The second part is a change within a promoter of a growth regulatory gene affecting gene expression (Lodish *et al.,* 2003). Amplification is the last known mechanism where there is DNA replication of a segment of DNA containing a proto-oncogene which then results in the over production of the encoded protein (Lodish *et al.,* 2003). The protein product of point mutations and the fusion of genes as a result of chromosomal translocation are different from the original protein that would have been encoded by a functional proto-oncogene (Lodish *et al.,* 2003). The protein product of chromosomal translocation and amplification results in the same protein product as the original, the only difference is that the normal production levels of the protein are no longer as before, with this overproduction leads to oncogenesis (Lodish *et al.,* 2003).

1.2.2 Oncogenesis gain of function from mutated tumour suppressor genes

Suppressor genes control the rate of cell growth and division. Mutations in these genes result in cells that no longer show contact inhibition (Thiel, 2010). In contrast to oncogenes, tumour suppressors suppress the ability of mutated genes to proliferate. In the event that these suppressor genes are no

longer present or non functional, a cell would then exhibit uncontrolled cell proliferation which could lead to the formation of malignancies (Hohenstein, 2004). The concept of tumour suppressor genes was first described by Alfred Knudson using statistical analysis. The Knudson hypothesis (also known as two-hit hypothesis) (Figure 1.2) is the hypothesis that cancer is the result of accumulated mutations to a cell's DNA. It was first proposed by Carl O. Nordling in 1953, and later formulated by Alfred G. Knudson in 1971. Knudson's work led indirectly to the identification of cancer-related genes. (Lodish *et al.,* 2003)

Knudson performed a study on children suffering from retinoblastoma (Lodish *et al.,* 2003). In a normal cell there are two alleles of tumour suppressor genes. The inactivation of the first allele by point mutations, small insertion/deletions, structural changes or loss of chromosomal function can be considered as the first hit, usually also considered to be unchanged (Nowak *et al.,* 2003). The second hit results in the inactivation of the second allele with a decreased rate of production of protein as a result of the mutations caused by the first hit event coupled with mitotic recombination. Large deletions or chromosomal losses do not act as a first or second hit as such mutations would be lethal to the cell (Nowak *et al.,*

2003). With tumour suppressor genes two alleles must be lost in order to promote tumour development (Hejmadi, 2010).

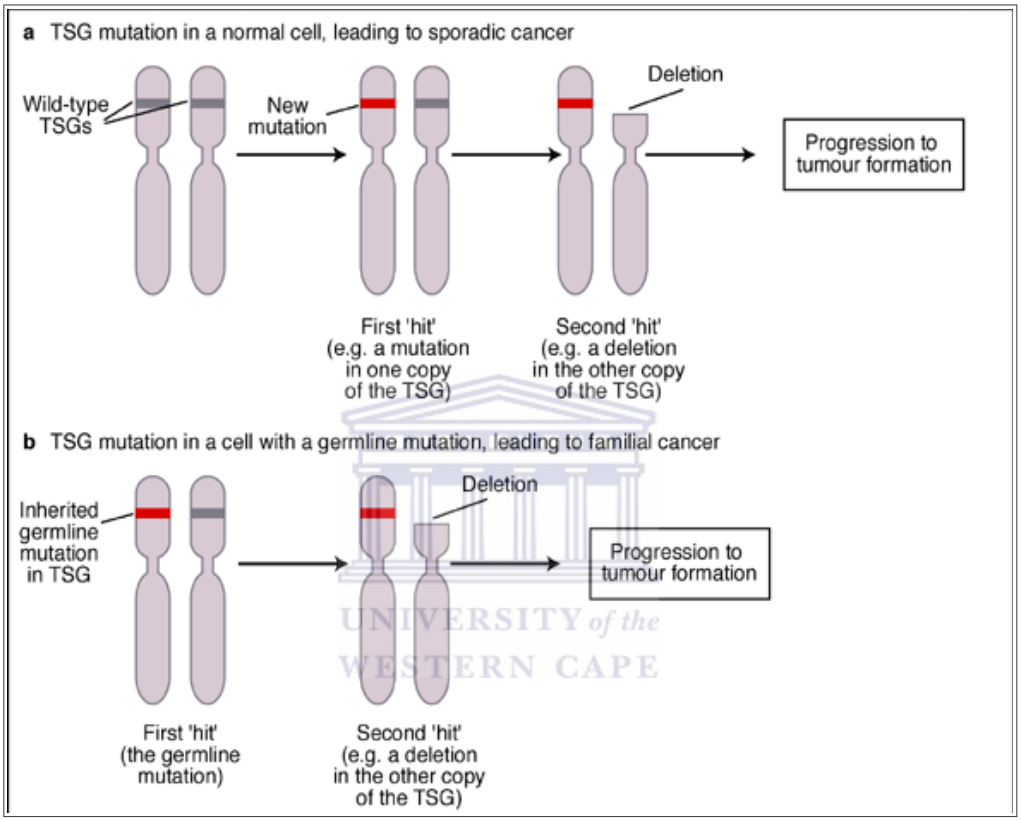


Figure 1.2: Knudson's two-hit hypothesis for tumourigenesis involving a tumour suppressor gene (TSG). One pair of chromosomes is depicted, with one TSG [the normal gene (grey), the mutated gene (red), and deletion of the gene (absence) are shown]. (a) Normal individuals have two normal copies of the TSG, so two independent 'hits' (mutations) are required in the same cell to initiate a cancer. (b) Individuals with a germ line mutation of the TSG already have a first 'hit' in every cell and require only one

subsequent ‘hit’ in a cell to initiate a cancer

In numerous cancers tumour suppressor genes have deletions or point mutations that stop the production or induce the production of a non functional protein (Hejmadi, 2010). Also there are a large number of these genes that have been hypothesised to exist and many of them have been identified in being linked to cancer development (Park and Vogelstein, 2003).

1.2.3 DNA mismatch repair genes

DNA mismatch repair (MMR) genes correct mismatches in the DNA generated during replication, thereby preventing mutations from becoming permanent in dividing cells. Because MMR genes reduce the number of replication-associated errors, defects in MMR genes increase the spontaneous mutation rate (Lodish *et al.,* 2003). Inactivation of MMR in human cells is associated with hereditary and sporadic human cancers (Christmann *et al.,* 2003). Since the MMR gene system is required for cell cycle arrest and/or programmed cell death in response to certain types of DNA damage, MMR genes play a role in the DNA damage response

pathway that eliminates severely damaged cells and prevents both mutagenesis in the short term and tumourigenesis in the long term. (Darrough, 2008).

There are at least three ways in which mismatched nucleotides arise in DNA, these include physical/chemical, insertions/deletions and genetic recombination to DNA and its precursors. Physical or chemical means are represented by actual damage to the DNA and its precursors. Insertions or deletions are represented by mis-incorporation of nucleotides in DNA during the replication process and genetic recombination produces regions of heteroduplexity in the regions of the DNA (Christmann *et al.,* 2003). DNA mismatch repair genes are inactivated in human cancers and because of their cellular function it might be argued that they form a unique subset of tumour suppressor genes (Park and Vogelstein, 2003). DNA mismatch repair systems correct DNA base pairing errors in newly replicated DNA. Mispaired nucleotides present after replication of DNA along with small insertion or deletion mutations that tend to occur at repetitive sequences (Masuda *et al.,* 2011). Mismatch repair system is an excision/resynthesis system and can be divided into 4 phases namely recognition of a mismatch,

recruitment of repair enzymes, excision of the incorrect sequences and then finally re-synthesis by DNA polymerase using a parental strand as a template (Masuda *et al.,* 2011).

A new pathogenetic mechanism leading to cancer has been delineated in recent years when human homologues of DNA mismatch repair (MMR) genes have been identified and shown to be involved in various types of cancer. Germ line mutations of MMR genes cause susceptibility to a hereditary form of colon cancer, hereditary non-polyposis colon cancer (HNPCC) (Levinson and Gutman, 1987; Strand *et al.,* 1993), which represents one of the most common syndromes associated with cancer predisposition in man. The existence of mismatch repair (MMR) enzymes in bacteria has been known for at least two decades and more recently, similar activities have been identified in yeast and higher eukaryotes (Modrich, 1991). The discovery of human MMR genes was greatly facilitated by advances made in genetic studies of hereditary nonpolyposis colorectal cancer (HNPCC). After the first HNPCC susceptibility locus was mapped to chromosome 2p by linkage analysis (Peltomaki *et al.,* 1993), it was found that tumours from HNPCC patients showed instability at

multiple random microsatellite sequences throughout the genome (Aaltonen *et al.,* 1993). A similar phenotype had previously been observed in bacterial and yeast strains with DNA mismatch repair gene mutations (Levinson and Gutman, 1987; Strand *et al.,* 1993). These data together provided a functional clue that resulted in the identification of human homologues of bacterial and yeast MMR genes

1.2.4 Model of Carcinogenesis in HNPCC and Sporadic tumours with MMR Deficiency

Molecular genetic studies support the idea that multiple genetic changes are necessary for tumourigenesis. Statistically, it has been estimated that colorectal tumours require four to seven mutations to develop (Renan, 1993). Colorectal carcinoma cell lines with MMR deficiency are hypermutable with mutation rates that can be several hundredfold compared to normal human cells (Parsons *et al.,* 1993, Bhattacharyya *et al.,* 1994). Importantly, mutations in HNPCC-related MMR genes cause significantly elevated mutation rates not only in repetitive sequences but in non repetitive gene sequences (such as the locus for hypoxanthine guanine phosphoribosyltransferase) as well (Bhattacharyya *et al.,* 1994). Figure 1.3 shows a model of colon cancer development in HNPCC. The basic

principles of this model are applicable to any tumours with inactivation of both copies of a MMR gene as an early event in their pathogenesis. In tumours with MMR deficiency, a classical tumour suppressor mechanism is supported in that two hits are required to cause a phenotypic effect (Parsons *et al.,* 1993; Liu *et al.,* 1995; Hemminki *et al.,* 1994). Most MMR gene mutations are point mutations resulting in truncated protein products (Lui *et al.,* 1996; de la Chapella and Peltomaki, 1995). Loss of critical MMR activity as a consequence of these mutations is presumed to occur early in tumour development, already at the adenoma stage (Shibata *et al.,* 1994). Ensuring genetic destabilization initiates a cascade of further mutations (Loeb, 1994), probably targeting different genes in tumours from different organs. In colon cancer, multiple somatic mutations, primarily single nucleotide deletions and G:C to A:T transitions, have been reported to occur in the APC and p53 genes (Lazar *et al.,* 1994). Furthermore, inactivating frame-shift mutations in a polyadenine tract located in the coding sequence of the transforming growth factor β-receptor II (TGFβ-RII) gene are frequent in colon cancer, but rare in endometrial cancers with microsatellite instability, which suggests different pathogenetic routes in these tumours (Markowitz *et al.,* 1995; Myeroff et al. 1995). In some

instances, even one hit (constitutional heterozygosity for an MMR gene mutation) may be sufficient for MMR deficiency to become manifested (Parsons *et al.,* 1995).

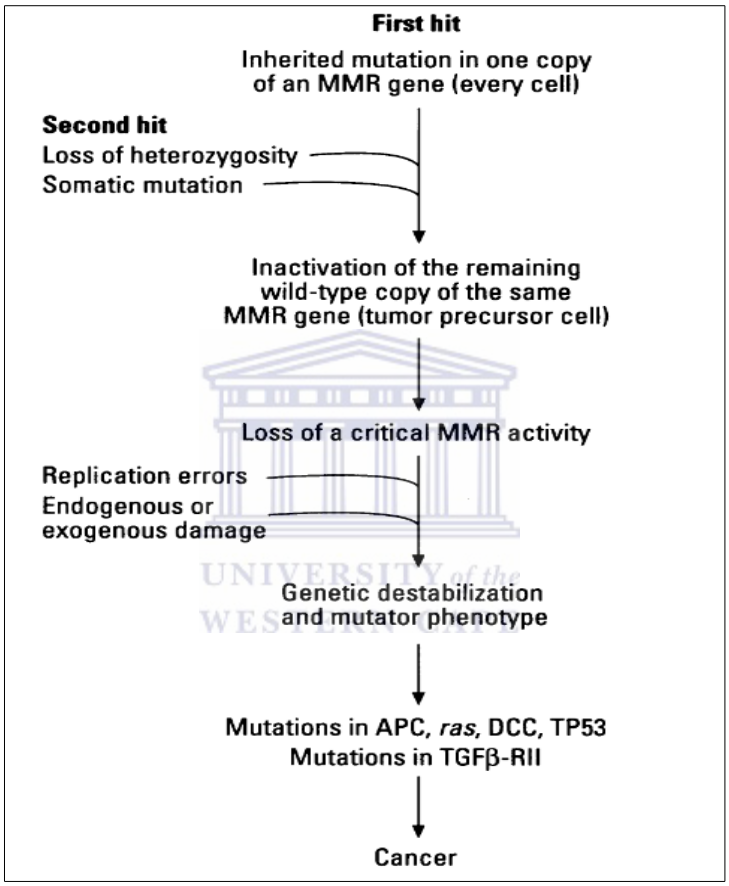


Figure 1.3: Graphical representation of the a model of colon cancer development in HNPCC. The basic principles of this model are applicable to any tumours with inactivation of both copies of a MMR gene as an early event in their pathogenesis. (Taken from Renan, 1993)

1.3 What is the prostate?

The prostate is a walnut sized gland located beneath the bladder surrounding the urethra and lies in the pelvic cavity immediately in front of the rectum (CANSA, 2013). The prostate starts to develop before birth and grows rapidly during puberty as a result of the male hormone androgen which is testosterone that is converted into dihydrotestosterone (DHT) the main signalling hormone for the prostate to grow (American Cancer Society, 2012). It usually stays the same size in adults or grows really slowly as long as these hormones are present (American Cancer Society, 2012). It produces seminal fluid assisting transport of sperm during ejaculation and forms part of the male reproductive system (Herbst, 2013). The prostate is muscular and partly glandular, made up of three lobes: central lobe and 2 adjacent lobes, it possesses ducts which enter into the prostatic portion of the urethra (Figure 1.4) (MHS, 2010).

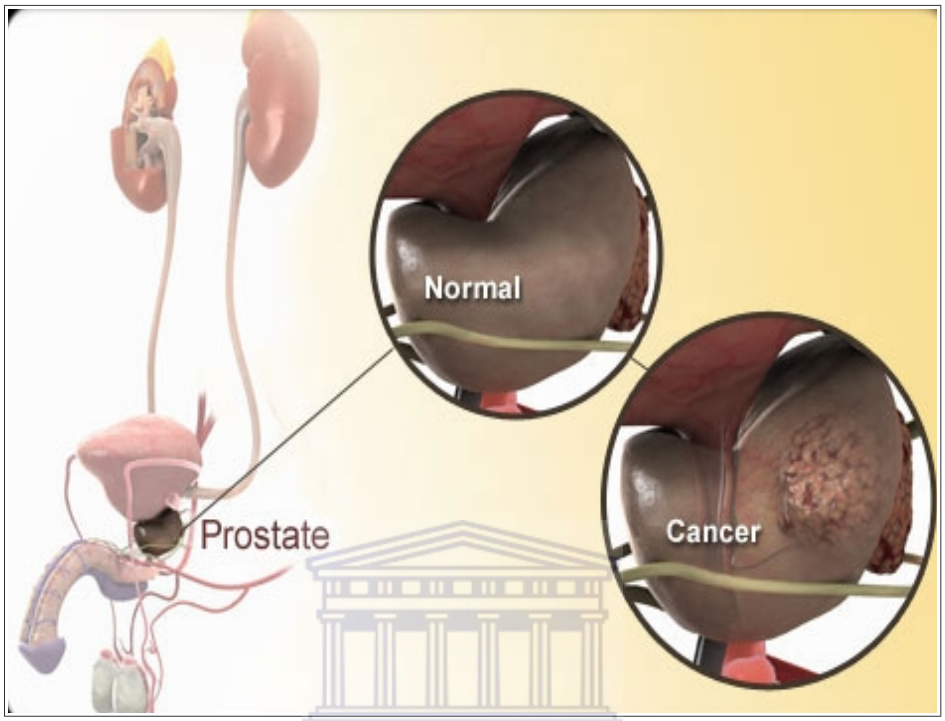


Figure 1.4: Diagram illustrating cancerous tissue in the prostate gland and normal prostate gland Adopted from (MHS, 2010)

1.3.1 What is Prostate cancer

Prostate cancer is caused by changes such as insertions, deletions and genetic recombination in the DNA of a normal prostate cell. DNA makes up the genes, which control how cells behave. Approximately 5 to 10% of prostate cancers are linked to these inherited changes in the DNA

(American Cancer Society, 2012). In essence, it is these changes that bring about abnormal growth in normal prostate tissue (CANSA, 2013). There are different types of prostate cancer which include adenocarcinomas, squamous and small cell carcinomas (CANSA, 2013). Adenocarcinomas are the most common form of prostate cancer and originates from the epithelial cells in the main glandular zone of the prostate (Herbst, 2013). Squamous cell carcinomas are small round cells and usually forms at the nerve cells. These cells usually form a very aggressive type of prostate cancer and because of the decrease in detection ability associated with squamous cells, cancer is generally detected at the very late stage (Herbst, 2013).

1.3.2 Genetics of prostate cancer

Prostate cancer genetics is a complex process as there are many genes and risk factors that aid in its pathogenesis (Hughes *et al.,* 2013). Molecular studies have provided unexpected clues on prostate cancer and its progression (Nelson *et al.,* 2003). Genetic factors are able to influence prostate cancer at many different stages through susceptible genes that could possibly have a role in the hereditary form of prostate cancer, yet this

is not completely understood (Rebbeck, 2005). Candidate genes functioning in body involving the male hormones, for example testosterone, has an effect on prostate growth as well as genes responding to the immune systems response to tumour growth (Rebbeck, 2005). Chromosomal regions containing extensive gene families proved helpful for identifying important cancer susceptibility genes (Zhang *et al.,* 2003). Somatic gene defects found in prostate cancer cells contain somatic mutations, gene deletions, gene amplification, chromosomal rearrangements and DNA methylation and most abnormalities reported were found at chromosomes 7p, 7q, 8q and Xq and losses at 8p, 10q, 13q and 16q as indicated in Figure 1.5 (Nelson *et al.,* 2003). Furthermore, these somatic changes in tumour suppressor genes targets genes in regions of chromosomes 7p, 7q, 8q and Xq (Figure 1.5)(Nelson *et al.,* 2003). Genetic instability and changes in gene expression, also contribute to genetics of prostate cancer (De Marzo *et al*. 2003). Gene expression changes include proliferation genes, androgen receptors and finally apoptosis and stress related genes (DeMarzo *et al*. 2003).

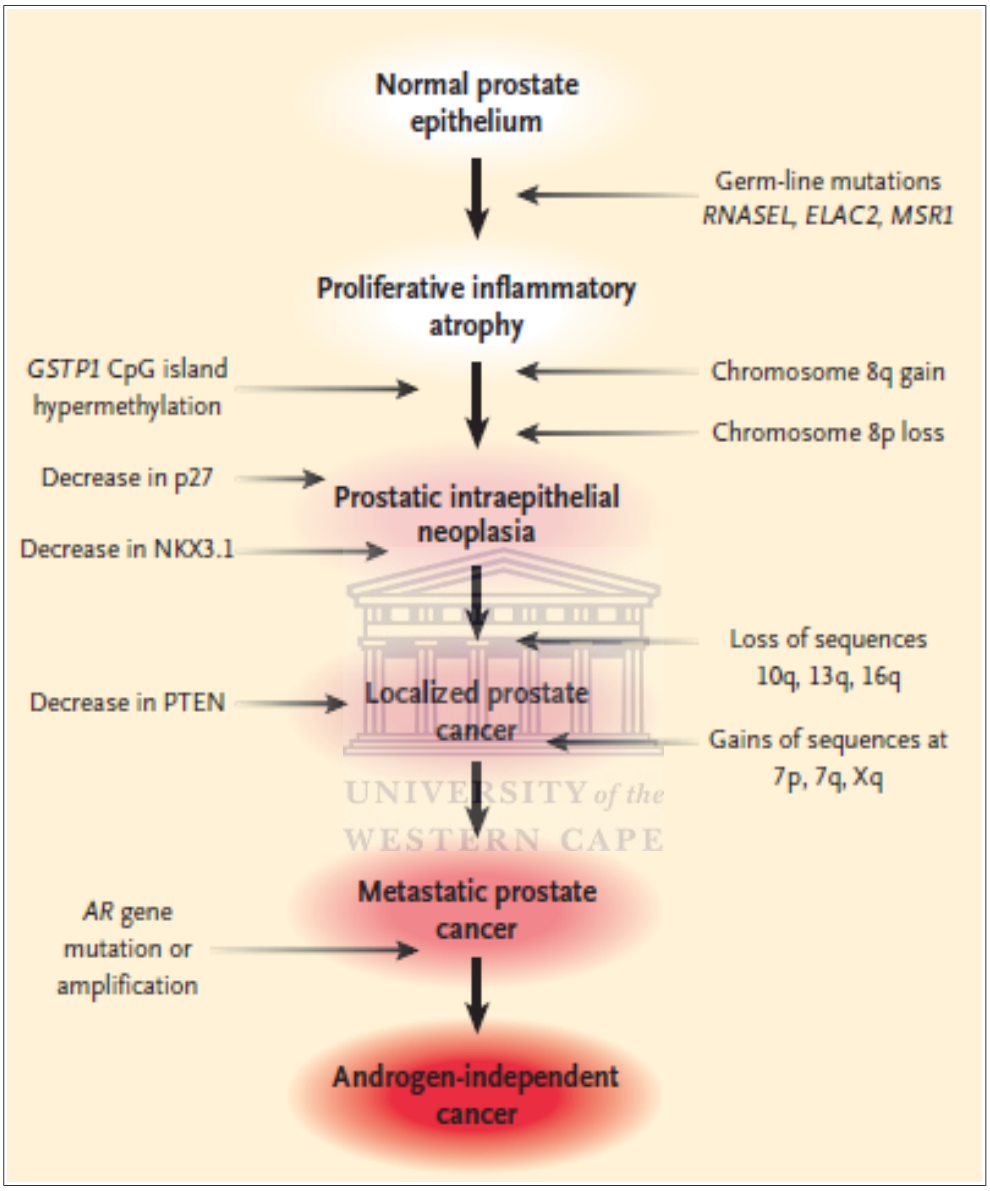


Figure 1.5: Molecular Pathogenesis of Prostate Cancer (Adopted from Nelson *et al* 2003)

1.4 Risk factors involved in prostate cancer

1.4.1 Age

Studies have shown that the incidence of PC increases with age (Herbst, 2013). However, it mostly affects men older than 50 years but it is not uncommon for PC to develop in men 40 years and younger (American Cancer Society, 2012).

1.4.2 Nationality and Ethnicity

The incidence of prostate cancer is more prevalent in certain ethnic groups. According to a study done in the Western Cape by Heyns *et al.,* 2011, it was found that the incidence of prostate cancer in white men were much higher as a result of many clinicopathological factors, when compared to coloured and black men in the population. However, according to the Cancer Association of South Africa the prevalence of prostate cancer in black men is on the increase Fisher *et al.,* 2011. In the USA the prevalence of prostate cancer in African-American men was much higher when compared to other ethnic groups and was diagnosed at much more advanced stages and was twice as likely to lead to death when compared to the rest of the male population. It also occurred less often in Asian-American and Hispanic or Latino men than in non-Hispanic (American

Cancer Society, 2012). These findings could be as a result of the complexity of the gene pool in mixed populations. Due to this genetic complexity and the variability of the disease it could be a defining factor for the increase in incidence rates in populations with mix race descent.

1.4.3 Diet and Lifestyle

Diet and lifestyle are also major contributing risk factors in the development of prostate cancer. Western industrialized countries diets has shown to be a contributing factor to prostate cancer, whereas a diet containing high fat, excess fibre, soy protein, vitamin E, selenium, carotenoids and herbal preparations all have a decreasing effect on prostate cancer risk (American Cancer Society, 2012). Lifestyle factors also play a role such as lack of exercise, increased alcohol intake, tobacco usage, sun exposure and also include environmental factors such as where one works and working with chemicals that could be carcinogenic (MHS, 2010).

1.4.4 Genetics

PC is hereditary, hence having a father or brother with PC usually doubles the risk of being confronted with the disease and risk increases when more than one family member is affected. It could also occur by chance, through acquiring somatic genes in your life time (MHS, 2010). Finally,

inflammation of the prostate, vasectomy and viral infections increase the risk as they have an effect on the properties of the normal cells inside the prostate and cause changes that can result in oncogenesis (MHS, 2010).

1.5 Staging and Grading of Prostate Cancer

Cancer is classified according to stages of progression. There are many classification methods used to determine the stages of cancer. The prognosis of prostate cancer depends on how fast the cancer has spread at the time of diagnosis (CANSA, 2013; American Cancer Society, 2012). These systems of staging are used to describe how far the cancer has spread and to accurately identify the stage of prostate cancer and are inevitably extremely important as it helps with providing the patient with the optimal treatment needed (CANSA, 2013). Figure 1.6 illustrates one method of prostate cancer staging called the Number Staging Method. In Stage I the cancer is found in the prostate this stage of prostate cancer is microscopic and usually goes undetected using the current diagnostic methods (CANSA, 2013). In Stage II the tumour has grown in size but is confined to the prostate and has not extended beyond the prostate. Stage III is where the cancer has slightly spread outside the prostate but barely to nearby tissue, such as the seminal vesicles (CANSA, 2013). Lastly, Stage IV is

where the cancer has metastasised outside the prostate to other tissues, this stage of cancer spreads most commonly to the lymph nodes, bones, liver or lungs (CANSA, 2013).

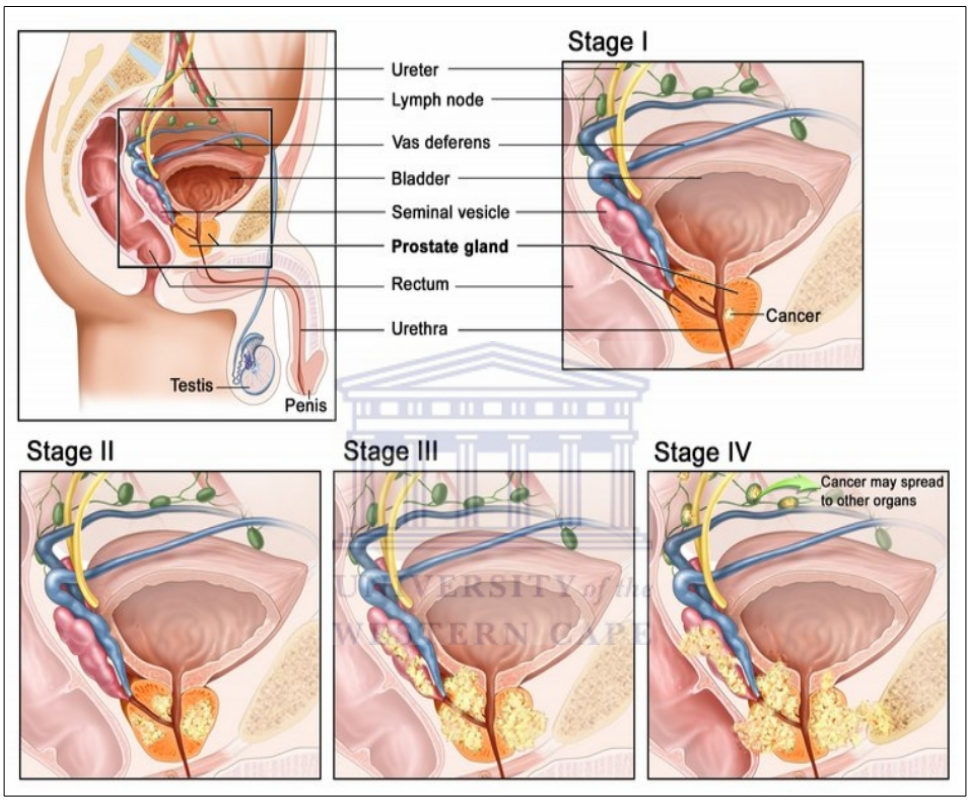


Figure 1.6: Classification of prostate cancer at its different stages using the Number Staging Method. From left to right. Normal prostate, Stage I: Cancer found in the prostate gland but is very small undetected by DRE, Stage II: Cancer has grown inside the prostate but is limited to only the prostate, Stage III: Cancer has spread to most of the prostate and to seminal vesicles outside of the prostate but very minimal and Stage IV: The cancer has taken over all of the prostate and has metastasised to other organs in the body.
(Adopted from www.medicalook.com/Cancer/Prostate_cancer.html)

1.6 Signs and symptoms early and late warning signs of cancer

Prostate cancer is usually asymptomatic which means that there is no early signs of the disease, therefore, prostate cancer screening is vital by going for annual physical examinations and blood tests to deliver the best possibility of identifying this disease while it is still in its early stages (MHS, 2010). Common symptoms of prostate cancer initially starts affecting a man's urinary system, such as weak or interrupted flow of urine (dribbling), an increased sensation to urinate especially at night, the feeling of urine staying behind in the bladder or holding back urine, difficulty urinating, inability to control the flow of urine also a painful or burning sensation when urinating and lastly an increased sense of discomfort in the pelvic area as well as swelling in the legs (MHS, 2010; CANSA, 2013). In the case of when the cancer has become more advanced blood might be found in the urine or semen, inability to ejaculate as well as erectile dysfunction, pain in the hips, back, chest as well as pelvic discomfort transforming into pain, weakness or numbness in the legs and feet, loss of bowel and bladder control, difficulty in urinating then turns to the complete inability to pass urine and lastly weight loss and bone pain (American Cancer Society, 2012; CANSA, 2013). However, it should be noted that these symptoms are not only limited to prostate cancer.

Another Staging method is known as the TNM staging method. The TNM system is one of the most widely used cancer staging systems. This system has been accepted by the Union for International Cancer Control (UICC) and the American Joint Committee on Cancer (AJCC) (CANSA, 2013). Most medical facilities use the TNM system as their main method for cancer reporting. The TNM system is based on the size and/or extent (reach) of the primary tumour (T), the amount of spread to nearby lymph nodes (N), and the presence of metastasis (M) or secondary tumours formed by the spread of cancer cells to other parts of the body. A number is added to each letter to indicate the size and/or extent of the primary tumour and the degree of cancer spread (CANSA 2013). Examples of how the system is implemented is available below

- Primary Tumour (T)
  TX: Primary tumour cannot be evaluated; T0: No evidence of primary tumour; Tis: Carcinoma in situ (CIS; abnormal cells are present but have not spread to neighbouring tissue; although not cancer, CIS may become cancer and is sometimes called pre invasive cancer); T1, T2, T3, T4: Size and/or extent of the primary tumour
- Regional Lymph Nodes (N)
  NX: Regional lymph nodes cannot be evaluated; N0: No regional lymph node involvement
  N1, N2, N3: Degree of regional lymph node involvement (number

and location of lymph nodes)

- Distant Metastasis (M)
  MX: Distant metastasis cannot be evaluated; M0: No distant metastasis; M1: Distant metastasis is present

Additionally grading methods also exist one such method is the The Gleason grading system. This system is used to help evaluate the prognosis of men with prostate cancer using samples from a prostate biopsy. Together with other parameters, it is incorporated into a strategy of prostate cancer staging which predicts prognosis and helps guide therapy. The Gleason grading system for PC tissue is based on how the tissue looks under a microscope. Gleason scores range from 2 to 10 and indicate how likely it is that a tumour will spread. A low Gleason score means the cancer tissue is similar to normal prostate tissue and the tumour is less likely to spread; a high Gleason score means the cancer tissue is very different from normal and the tumour is more likely to spread (Brase *et al.,* 2011).

1.7 Screening and diagnosis of prostate cancer

Screening for prostate cancer aids in the early diagnosis of its malignancy before the patient presents any symptoms, therefore making the management of prostate cancer more efficient and the delivery of a wider range of therapeutic modalities (Stravidis *et al.,* 2010; Hoffman, 2011; Wolf *et al.,* 2010). The basis of urological examinations for prostatic cancer

screening are called “the diagnostic triad” which consists of the Digital Rectal Examination (DRE) , prostate specific antigen (PSA) test and investigation with the aid of the trans rectal ultrasound biopsy (Stravidis *et al.,* 2010; Qaseem *et al.,* 2013)

1.7.1 Digital Rectal Examination (DRE)

The DRE is a procedure where a doctor would insert a lubricated gloved finger into the rectum to examine the prostate for any abnormalities such as bumps or hard areas (Figure 1.7) (American Cancer Society, 2012, CRUK, 2012, MHS, 2012, They, 2012). The limitations to this method is because only the posterior and lateral part of the prostate can be palpated this then leaves 40-50% of cancers beyond reach, meaning a possible false diagnoses. It therefore poses a disadvantage not only giving a false negative diagnosis but also lacking in sensitivity and specificity in males that are asymptomatic (American Cancer Society, 2010; Feightner 2008).

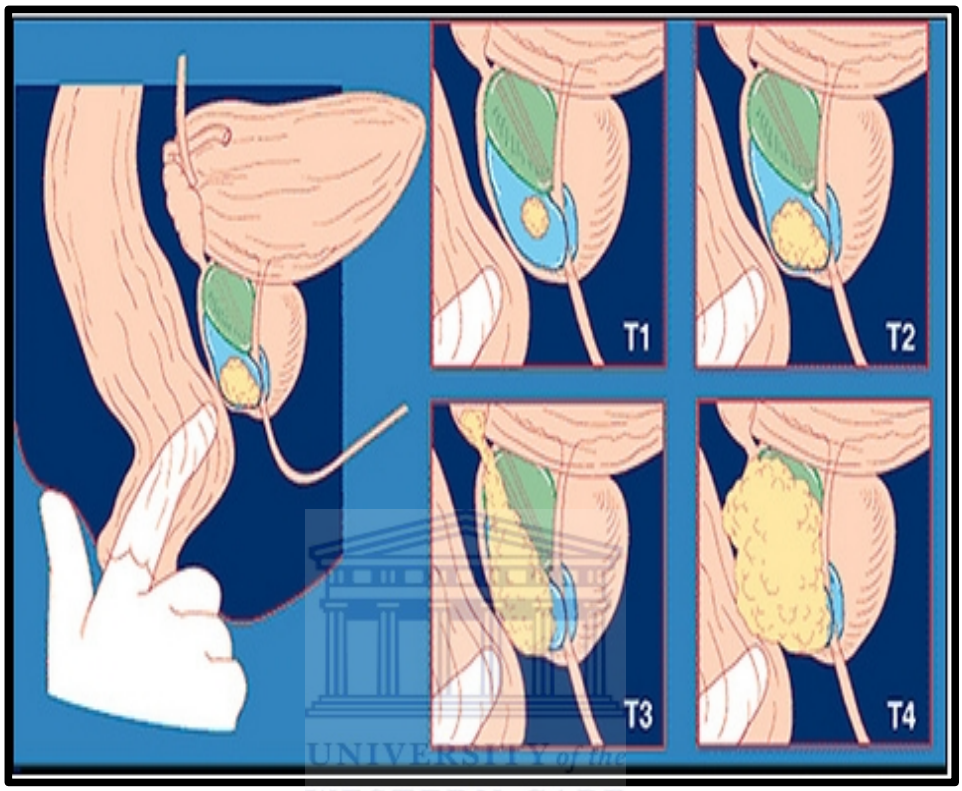


Figure 1.7: Digital rectal exam: side view of the male reproductive and urinary anatomy, including the prostate, rectum, and bladder. Where the clinician inserts the forefinger into the anus to determine if irregularities exist within the prostate. Also depicting 4 stages indicating the various degrees of PC development where T1 is the earliest stage of PC and T4 shows PC during metastasis (taken from American Cancer Society, 2012)

1.7.2 The Prostate Specific Antigen (PSA) test

PSA test is a diagnostic blood test in which the PSA levels ins the blood are measured as depicted in Figure 1.8 (MHS, 2010). It is performed to determine the incidence of any prostate abnormalities which include prostate cancer (CANSA, 2013). PSA is a protein produced by the cells in the prostate gland in normal, cancerous and benign prostate tissue (American Cancer Society, 2010; CANSA, 2013). PSA is mostly found in semen and a small amount enters into the bloodstream. The PSA levels in healthy men are supposed to be less than 4 nano-grams per millilitre (ng/ml), however, in cancer cells the PSA levels are increased therefore causing the PSA levels in the blood to rise (ACS American Cancer Society, 2010; CANSA, 2013). PSA test is the most widely used test for the possible detection and confirmation of prostate cancer (CANSA, 2013). The limitation of this test is that if a high PSA level is found it does not conclude that a patient has prostate cancer but it could also indicate other problems or abnormalities (CANSA, 2013). Factors that may affect PSA levels causing it to be higher could be due to an enlarged prostate usually in elderly men above the age of 50 years, inflammation of the prostate gland (prostatitis), also infection of the bladder, recent ejaculation, bicycle riding,

urological procedures such as the DRE, certain medication, herbal mixtures and obesity (American Cancer Society, 2012; CANSA, 2013).

If results from the DRE and PSA blood test are abnormal it may suggest that the patient has prostate cancer and a physician would then schedule a trans rectal biopsy better known as a prostate biopsy.

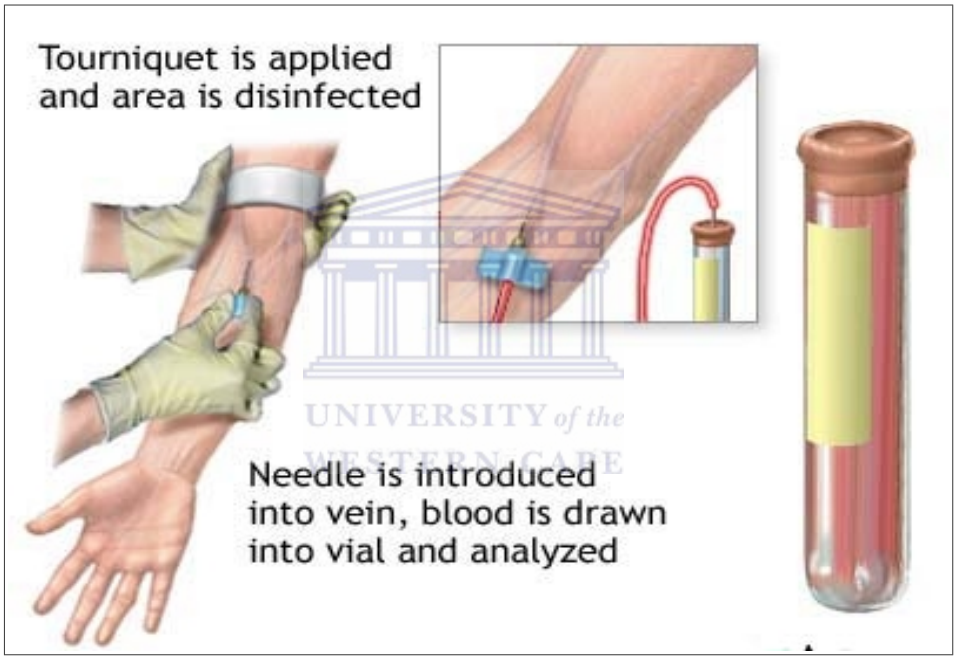


Figure 1.8: Graphical representation of the method used to extract a sample to do a PSA diagnostic screen. Prostate-specific antigen (PSA) is a glycoprotein in the cytoplasm of prostatic epithelial cells. It can be detected in the blood of all adult men. The PSA level is increased in men with prostate cancer but can also be increased somewhat in other disorders of the prostate (taken from CANSA, 2013).

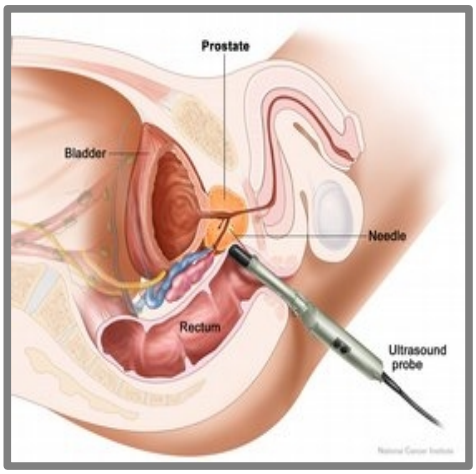
1.7.3 Transrectal ultrasound and biopsy

In the event of abnormal DRE and PSA blood test results a trans rectal ultrasound biopsy better known as a prostate biopsy procedure would be performed as a further screening measure (Simmons *et al.,* 2011; Stavridis *et al.,* 2010; Wolf *et al.,* 2010). A core needle biopsy is the main method to diagnose prostate cancer, specimens are acquired by a urologist who treats issues regarding the male reproductive system. A prostate biopsy involves the removal of prostate tissue for analysis by a histopathologist for cancerous tissue (Figure 1.9B) (CANSA, 2013; American Cancer Society, 2012).

A trans rectal ultrasound is used as a guide to insert thin hollow needles through the wall of the rectum into the prostate gland. When these needles are pulled out, small samples of prostate tissue are removed, this process is repeated eight to 36 times (Figure 1.9 A) (American Cancer Society, 2012; Simmons *et al.,* 2011). The disadvantage to this method is that it is very invasive as it causes discomfort and pain. Furthermore, it could cause infection and there is also the risk of a false negative diagnosis; even though many samples are taken the biopsies might still miss a tumour if none of the needles pass through it (American Cancer Society, 2012).

The trans rectal ultrasound can be used to determine the presence of cancer in the prostate, here a doctor inserts a lubricated probe into the rectum of the male to check for abnormal areas. The probe then sends out sound waves that bounce off the prostate and these results are fed into a computer which then produces an image of the results and is called a sonogram (CANSA, 2013). The limitation to this is that if the cancer is still in its early stages before the tumour can be seen it cannot always be detected, using this method (CANSA, 2013).

A



B

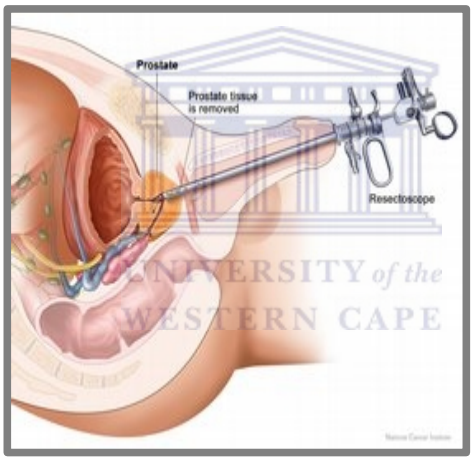


Figure 1.9: Image of current detection method namely the Ultasound and Prostate biopsy. Where figure A represents the trans rectal ultra sound being inserted into the anus to get a clearer look at the status of the prostate. Figure B represents the procedure for the biopsy where a syringe is inserted into the penis to extract a sample from the prostate gland (Taken from American Cancer Society, 2012).

1.7.4 Computed tomography (CT) scan

A computed tomography, commonly known as a CT scan, is a specialised x-ray test producing a detailed outline of cross sectional images of the body (American Cancer Society, 2012). This scan takes several pictures of the specific area of the body under observation and rotates around the subject while it takes these pictures (American Cancer Society, 2012). These results are then fed into a computer where pictures are then combined into images of "slices" of that specific part of the body. In order for these images to be produced a patient will receive an intravenous line through which a contrast dye is injected allowing better outlining of one's body structures on the inside (American Cancer Society, 2012). However, there are limitations to this screening method in that some individuals are allergic to the dye used and get allergic reactions, such as the appearance of hives. If more severe, it can lead to difficulty in breathing, and low blood pressure (American Cancer Society, 2013).

1.7.5 Magnetic Resonance Imaging (MRI)

Magnetic resonance imaging (MRI) can also be used to diagnose prostate cancer. This technique is based on the use of radio waves and strong magnets instead of x-rays, the energy of the radio waves are absorbed into

the body and a specific pattern is created by the type of body tissue and by tumours or cell masses (MHS, 2010). The uniquely created patterns are then translated by a computer to give a very detailed image of the part of the body being examined (American Cancer Society, 2012). This method produces very clear images of the prostate and is useful in showing if the cancer has spread to other parts of the body. Despite all the advantages mentioned, it cannot provide vital information about early stage PC cases. Additionally, men may experience allergic reactions due to the dye, whilst men afraid of confined places might be traumatised after such a scan and the clicking and buzzing noises generated during the scan might be irritating as the scan may take up to an hour (CANSA, 2013). Lastly, due to the magnetic nature of the MRI machine, men with pacemakers or other metal medical implants cannot go for a MRI scan (American Cancer Society, 2012).

An ideal screening test requires it to be minimally invasive, accurate, conveniently available and should be available to the general population at a low cost aiding a better disease outcome because of an accurate and early diagnosis (Stravidis *et al.,* 2010). With all the diagnostic tests, even with

the benefit associated with it, it has a number of potential harms not only physically but psychological as well (Carter *et al.,* 2013)

1.8. Biomarkers and their roles in cancer therapy and diagnosis

A molecular biomarker, according to the National Cancer Institute is classified as a biological molecule present in blood, bodily fluids or tissues that provides an indication of a normal or abnormal process or of a condition or a disease. The common characteristics of a biomarker are their inexpensiveness, reliability and ease of expression remaining unaltered under diseased conditions (Verma *et al.,* 2011). A biomarker should allow the detection of a tumour before a symptomatic mass or other symptoms develop (Verma *et al.,* 2011). These molecules are used as aids to characterize tumour composition for improvement of screening accuracy as well as avoidance of invasive methods for diagnosis. Also should have a high specificity of diagnosis as well as prognosis of the disease for the reduction of false-positives and negatives (Madu & Lu, 2010; Verma *et al.,* 2011).

Biomarkers have the ability to supply information regarding the progression (prognostic), the efficacy of treatment as well as the possible

re-occurrence of the cancer (predictive) (Madu & Lu, 2010; Verma *et al.,* 2011; Haj-Ahmad, 2013). Diagnostic biomarkers are used for the diagnosis of disease with a very high success rate. These biomarkers ideally target specific molecular events based on the differences between cancerous and non-cancerous cells necessary for a specific cancer type to progress and ideally prevent the use of conventional detection methods (Haj-Ahmad, 2013; RAN, 2012; Verma *et al.,* 2011).

The first cancer marker ever reported was the light chain of immunoglobulin in the urine, as identified in 75% of patients with myeloma in 1848 study (Jones, 1848). The test for this marker is still employed by clinicians today, but with use of modern quantification techniques. From 1930 to 1960, scientists identified numerous hormones, enzymes and other proteins, the concentration of which was altered in biological fluids from patients with cancer. The modern era of monitoring malignant disease, however, began in the 1960s with the discovery of the alpha-fetoprotein (Abelev *et al.,* 1963) and carcinoembryonic antigen (CEA) (Gold and Freedman, 1965), which was facilitated by the introduction of immunological techniques such as the radio-immunoassay.

In the 1980s, the era of hybridoma technology enabled development of the ovarian epithelial cancer marker carbohydrate antigen (CA) 125 (Bast *et al.,* 1981). In 1980, prostate-specific antigen (PSA [KLK3]), considered one of the best cancer markers, was also discovered (Papsidero *et al.,* 1980).

Every era, biomarker discovery seems to be associated closely with the emergence of a new and powerful analytical technologies. The past decade has witnessed an impressive growth in the field of large-scale and high-throughput biology such as Next Generation Sequencing (NGS), which has contributed to an era of new technological development. The completion of a number of genome-sequencing projects, the discovery of oncogenes and tumour-suppressor genes, and recent advances in genomic and proteomic technologies, together with powerful bioinformatics tools, will have a direct and major impact on the way the search for cancer biomarkers is conducted. Early discoveries of cancer biomarkers were based mainly on empirical observations, such as the over expression of CEA (Gold and Freedman, 1965). The modern technologies are capable of performing parallel rather than serial analyses, and they can help to identify

distinguishing patterns and multiple markers rather than just a single marker. Such strategies represent a central component and a paradigm shift in the search for novel biomarkers

There are different biomarkers with each having their own respective function with their own molecular, biochemical and physiological properties (RAN, 2012; Verma *et al.,* 2012). These biomarker types include physiologic, anatomic, genetic and proteomic biomarkers (RAN, 2012; Verma *et al.,* 2012). The approach for biomarker discovery employs one of four practices for discovery viz. proteomics based, nucleic acid based, metabolomics and *in silico* approaches which falls under the scope of this project for biomarker discovery (RAN, 2012; Haj – Ahmad, 2013). The advantages of using an *in silico* approach involves the use of computer algorithms to search for biomarkers where curated experiments have been made publicly available allowing one to conduct more than one computerized search for all genomic data sets available to identify a molecule as a potential biomarker (RAN, 2012).

1.8.1 Biomarker discovery using Gene expression profiling technology

Genomic microarrays represent a highly powerful technology for gene-expression studies. Microarray experiments are usually performed with DNA or RNA isolated from tissues, which are labelled with a detectable marker and allowed to hybridize to arrays comprised of gene-specific probes that represent thousands of individual genes (Quackenbush, 2006). The greater the degree of hybridization, the more intense the signal, thus implying a higher relative level of expression.

The massive amount of data per experiment means that the molecular markers and their expression patterns need to be analysed by elaborate computational tools to detect emerging patterns from these huge datasets. Two basic forms of analysis exists which are unsupervised and supervised hierarchical clustering algorithms (Eisen *et al.,* 1998); the latter tools identifies gene-expression patterns that discriminate tumours on the basis of predefined clinical information whilst the former looks at gene expression generally throughout a system. (Golub *et al.,* 1999).

The cancer sub-classification hypothesis states that gene-expression

patterns identified with DNA micro-arrays can predict the clinical behaviour of tumours (Perou *et al.,* 2000). The proof-of-principle for the cancer sub-classification hypothesis has been provided for various malignancies, such as leukaemia, breast cancers and many other tumour types (Alizadeh *et al.,* 2001; Weigelt *et al.,* 2005; Alizadeh *et al.,* 2000; Rosenwald *et al.,* 2002; Pomeroy *et al.,* 2002; lizuka *et al.,* 2004; Chen *et al.,* 2007). For example, results from gene-array technologies have enabled breast cancers to be classified into prognostic categories dependent on the expression of certain genes. The 70-gene-panel micro-array study of survival prediction led to the development of MammaPrint® (Agendia, Amsterdam, The Netherlands) (van de Vijver *et al.,* 2002), which in February 2007 became the first multigene panel test to be approved by the FDA for predicting breast cancer relapse as an *in silico* method.

Similarly, Oncotype DX® (Genomic Health, Redwood City, CA) another genomic micro-array, is based on quantitative real-time PCR and has been commercially available since 2004 (Paik *et al.,* 2004). Oncotype DX® and MammaPrint® uses different analytical platforms and despite their similar clinical indication, they have only a single gene overlap in their panels.

Nevertheless, over the past decade, a tremendous growth in the application of gene-expression profiling has been witnessed. This growth has contributed to the cancer sub-classification theory (Perou *et al.,* 2000), insights into cancer pathogenesis, and the discovery of a large number of diagnostic markers (Pollack, 2007)

1.9. The need for bioinformatics in Biomarker discovery

Bioinformatics leads a researcher to better understand, organise and associate molecules on a greater scale (Settles, 2008). There are many areas and subsequent applications in bioinformatics, areas including genomics, proteomics and phylogenetics with applications relating to the medical field such as diagnosis of disease, drug design and discovery to name a few as well as agriculture and microbial genome applications (Mascrenghe, 2004; Madhani, 2006). Bioinformatics is currently employed for the improvement of the drug discovery pipeline. It is an inexpensive method and less labour intensive and time consuming when compared to traditional molecular methods. Therefore with the aid of bioinformatics as a research tool it is possible to generate a list of priority genes/proteins/miRNA in a short space of time and through an enrichment process of these lists eliminating those which appear not to be useful targets

or a priority based on search criteria and the statistical analysis performed (Madhani, 2006).

New resources will most likely identify novel protein, genetic and low-molecular-weight cancer markers, which may impact on cancer care. Furthermore, with advances in genomic and proteomic technologies, human diseases may be classified on the basis of molecular rather than morphological analysis. Moreover, bioinformatics will serve to link scientific data with clinical information. Despite the optimism, the National Academy of Clinical Biochemistry do not encourage the widespread use of tumour markers unless they affect patient outcome measures (Bast *et al.,* 2001). There is, however, a general agreement that a combination of multiple biomarkers may increase diagnostic sensitivity and specificity over use of individual markers. This is particularly important in relation to the recent development of powerful bioinformatics algorithms, which can interpret multiple parameters much more efficiently than can more-traditional approaches (Finne *et al.,* 2000; Stephen *et al.,* 2002). The most accurate, individualized, predictive assessment for patients might be attained through the use of artificial neural networks (Stephen *et al.,* 2002).

There is no doubt that if these new technological advances prove to be successful in identifying cancer biomarkers for early cancer detection, the clinical benefits are likely to be enormous.

1.10. The importance of MicroRNAs as biomarkers in cancer

miRNA research has received an increasing amount of interest (Buckingham, 2003). miRNAs are made up of phylogenetically conserved endogenous small non coding single RNA genes (Buckingham, 2003; Nilsen, 2007; Zhao and Srivastava, 2007; Tanzer *et al.,* 2008; Sotiropoulou *et al.,* 2009; Sato *et al.,* 2011; Martinez-Sanchez and Murphy, 2013). In humans they make up about 1-3% of the human genome, and approximately 80% of these miRNA genes are located in the introns of both protein and non-coding genes (Buckingham, 2003; Nilsen, 2007; Zhao and Srivastava, 2007; Tanzer *et al.,* 2008; Sotiropoulou *et al.,* 2009; Sato *et al.,* 2011; Martinez-Sanchez and Murphy, 2013). MiRNAs were found to influence the developmental timing, organ development, cell differentiation, proliferation and most significantly apoptosis of cells in several organisms (Davis and Hata, 2009; Sotiropoulou *et al.,* 2009; Pritchard *et al.,* 2012;).

MiRNAs are small, non-coding subset of RNAs which consist of about 18-25 nucleotides and bind to the 3' untranslated region of messenger RNAs (mRNAs) (Wach *et al.,* 2011). By this action, they cause post-transcriptional inhibition or degradation of target mRNA, depending on the degree of complementary base pairing (Doench and Sharp*.,* 2004; Vella *et al.,* 2004 and Xu *et al.,* 2011). miRNAs were first discovered in 1993 while studying *Caenorhabditis elegans* (Lee *et al.,* 1993). The first miRNA discovered was lin-4. It is a small, non-coding RNA molecule that was found to play a role in controlling the timing of larval development in *C. elegans* through a negative effect on lin-14 expression (Lee *et al.,* 1993; Wightman *et al.,* 1993; Vandenboom *et al.,* 2008). Seven years later, in 2000, let-7, the second miRNA was discovered, let-7 miRNA is often repressed in human cancer cells, suggesting that they may target oncogenes and be functionally equivalent to a tumour suppressor (Lee *et al.,* 1993). In fact, recent studies indicate that let-7 miRNA suppresses expression of Ras and other oncogenic proteins (Doench and Sharo*.,* 2004; Vella *et al.,* 2004 and Xu *et al.,* 2011). This suggests that re-expression of miRNA that down regulate oncogenic proteins could be a useful therapeutic approach for

human cancer (Reinhart *et al.,* 2000; Griffiths-Jones *et al.,* 2006; Vandenboom *et al.,* 2008). During the past 12 years, significant advances have been made in miRNA research leading to the discovery of over 4,500 miRNAs in vertebrates, flies, worms, plants, and viruses (Ambros, 2003; Griffiths-Jones *et al.,* 2006; Vandenboom *et al.,* 2008) out of which more than 1,000 miRNAs are fully characterized with and associated with a particular disease and this number is expected to grow in the coming years.

MiRNAs are constantly implicated in the regulation of an increasing number of physiological processes. It is also believed now that they play an important role in the regulation of many cellular functions ranging from maintenance to differentiation and tissue development, metabolism and cell cycle regulation (Bushati and Cohen, 2007; Stefani and Slack, 2008; Gangaraju and Lin, 2009; Coppola *et al.,* 2010). All of these facts leads to the conclusion that aberrant expression of miRNAs will have an impact on various biological processes in which they are implicated, which will result in a variety of pathological events such as inflammation (Baltimore *et al.,* 2008; Cullen, 2009; Coppola *et al.,* 2010), cardiovascular diseases (Latronico *et al.,* 2009), neurodegenerative diseases (Bushati and Cohen,

2008) and most importantly, cancer (Coppola *et al.,* 2010).

The role of miRNAs in cellular growth, differentiation and apoptosis of cancer cells through their interactions with their target mRNA has been well studied (Kloosterman and Plasterk, 2006; Baranwal and Alahari, 2010; Pang *et al.,* 2010). miRNAs may be oncogenic or tumour suppressor molecules (Pang *et al.,* 2010; Vrba *et al.,* 2010), as oncogenic molecules being up-regulated whilst as tumour suppressor molecules being down-regulated in cancers. Generally, the importance of miRNAs in cancer is emphasized by the fact that around 50% of all miRNA genes are positioned in the so called 'fragile sites', the cancer associated genomic regions which are repeatedly changed in cancer. A lot of information has already been identified about aberrant miRNAs expression in cancers; the understanding of the functional importance of these aberrations has not been molecularly exploited (Aqeilan, 2010).

1.10.1 Biogenesis of miRNAs

The biogenesis of miRNA involves many critical steps; the initial phase is the transcription by RNA polymerase II that leads to the formation of

primary miRNA (pri-miRNA) which comprises of hundreds to thousands of nucleotides (Lee *et al.,* 2004; Vandenboom *et al.,* 2008). The second phase is catalysed by a ribonuclease called ribonuclease III (RNase III), Drosha. This step leads to the split of pri-miRNA and results in what is called a precursor miRNA (pre-miRNAs), which usually comprises of around 70 nucleotides, and this phase is accomplished with the help of DGCR8, a protein that helps to ensure perfect and efficient processing of pri-miRNA into pre-miRNAs (Lee *et al.,* 2003; Denli *et al.,* 2004; Vandenboom *et al.,* 2008). Thereafter, a nuclear export factor, called exportin 5, binds to the pre-miRNAs and transports it into the cytoplasm where the next processing phase take place (Yi *et al.,* 2003; Bohnsack *et al.,* 2004, Vandenboom *et al.,* 2008). Here another RNase III Pol, Dicer, interacts with the pre-miRNAs, and the outcome of this process is the formation of a RNA duplex of around 22 nucleotides, which is the mature miRNA consisting of a double-stranded duplex. Dicer usually operates with the help of the trans-activator RNA (transacting response)-binding protein (TRBP) (Bernstein *et al.,* 2001; Ketting *et al.,* 2001, Vandenboom *et al.,* 2008). Subsequently, all these steps allow the mature miRNA interact with the RNA-induced silencing complex (RISC) (Hammond *et al.,* 2000; Schwarz *et al.,* 2002;

Vandenboom *et al.,* 2008), and ultimately allow miRNA to control post transcriptional regulation of functional mRNAs (Figure 1.10). miRNAs have the ability to not only target several hundred genes but also to negatively regulate original mRNA instruction by translational repression or decay depending on the binding conformations leading to no or very little protein production (Nazarov *et al.,* 2013; Wang, 2008; Erson & Petty, 2008; Nilsen, 2007; Pritchard *et al.,* 2012; Fabian *et al.,* 2010; Cuperus *et al.,* 2011 ; Davis & Hata, 2009).

The significance of cancer causing miRNAs (oncomiRs) was established because of its involvement in many cellular processes such as cell differentiation, proliferation and more importantly apoptosis relevant to this research. miRNA expression was found to show a relationship with many cancers and genes thought to function as both tumour suppressors and oncogenes (Esqauela-Kerscher & Slack, 2006). miRNA over expression in tumours are considered oncogenic as they inhibit tumour suppressor genes, cell cycle, differentiation and apoptotic control genes whilst for miRNA under-expression these miRNAs are considered as tumour suppressors and have the opposite effect to oncogenic miRNAs

(Sapre & Selth, 2013 and Lopez-Camarillo *et al.,* 2013).

MiRNA-mRNA mediated interactions show promising implementation for biomarkers, and have provided the potential to a more in depth understanding of diseases and the observation of certain processes in the body and can help combat all the current issues with current diagnostic tests such as specificity, sensitivity and minimal invasiveness (VandenBoom *et al.,* 2008; Sapre & Selth, 2013; Lopez-Camarillo *et al.,* 2013; Wang, 2008; Esquela-Kerscher & Slack, 2006).

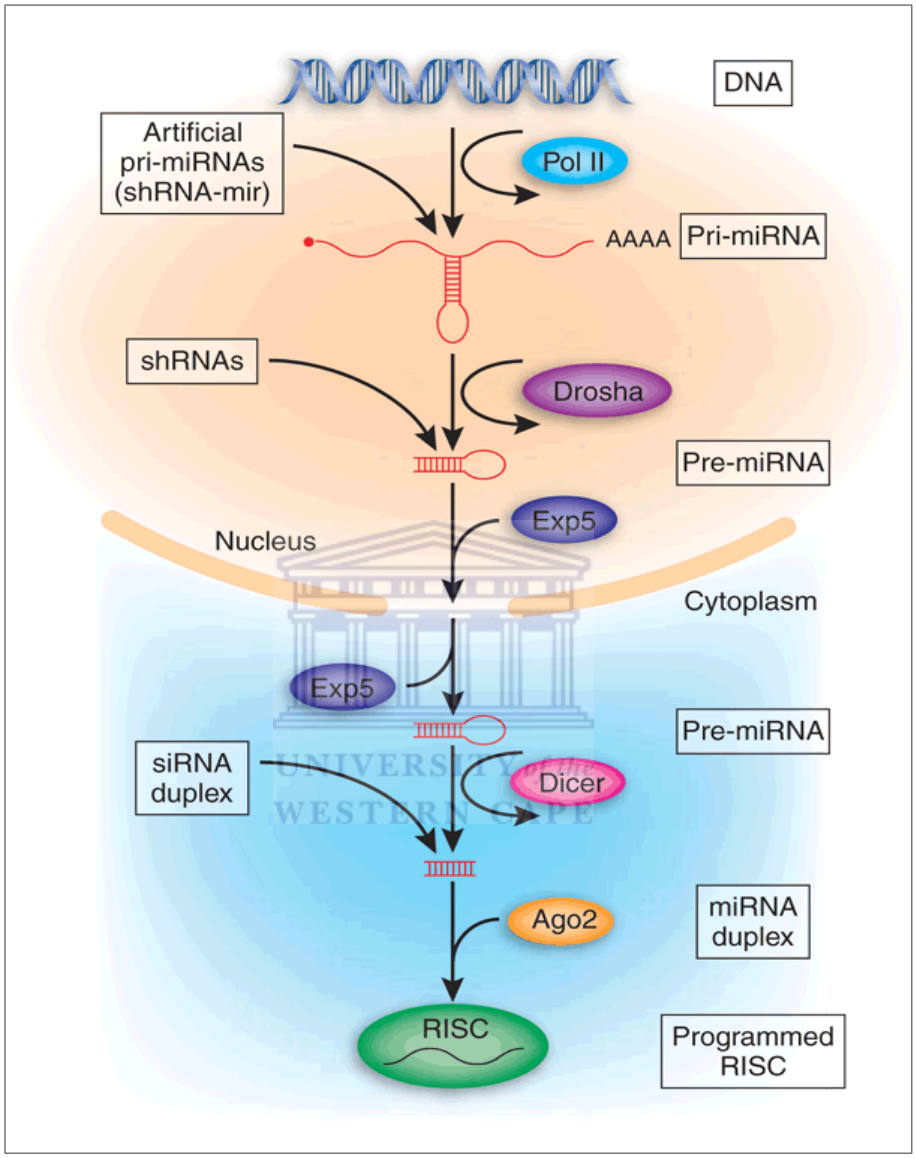


Figure 1.10: The miRNA biogenesis pathway in vertebrate cells. Artificial siRNAs can enter this pathway as synthetic siRNA duplexes, as shRNAs transcribed by Pol II or as artificial pri-miRNAs (shRNA-mir). For simplicity, not all factors involved in miRNA biogenesis are shown. Ago2,

Argonaute-2; Exp5, Exportin-5 (taken from Yi *et al.,* 2003).

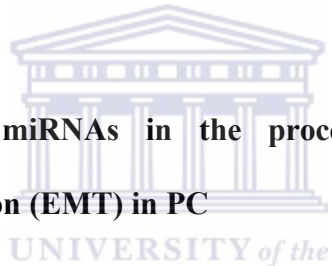
1.10.2 miRNA and cancer stem cells in prostate cancer

Cancer stem cells (CSCs), also called tumour-initiating cells, are a group of cells which play an important role in the progression of cancer and its metastasis (Visvader & Linderman, 2008) . The CSCs hypothesis assumes that cancers are basically derived from a small fraction of cancer cells that have exclusive ability to self-renew and initiate/maintain the tumour (Papagiannakopoulos & Kosik, 2008; DeSano & Xhu, 2009). This specific ability of CSCs allowing the initiation of the development of cancer was documented first in human leukaemia (Dick, 1996; Jin *et al.,* 2006) . For the isolation and identification of CSCs, many markers have been used. Previously, CD44 (an adhesion molecule) was used to identify CSCs, individually or in some instances in combination with other markers. It was shown that Prostate CSCs that have increased clonogenic potential (Collins *et al.,* 2005), tumour initiating ability and metastatic capabilities have an increased $CD44^+$ cell population (Patrawala *et al.,* 2006).

Emerging evidence suggests that miRNAs may function as the regulators of

CSC characteristics (Croce & Calin, 2005; Yu *et al.,* 2007; Shimono *et al.,* 2009; Melton *et al.,* 2010). Studies have clearly shown that the expression levels of certain miRNAs in stem cells are different from other normal tissues (Suh *et al.,* 2004; DeSano & Xu, 2009), suggesting that miRNAs are important regulators of CSC function. One example is miR-34a, a p53 target (Bommer *et al.,* 2007; Chang *et al.,* 2007; He *et al.,* 2007; Raver-Shapira *et al.,* 2007; Tarasov *et al.,* 2007), which has strong anti-tumour and anti-metastatic effects (Liu *et al.,* 2011). Studies have shown that miR-34a is down regulated in $CD44^+$ PC cells, and its over- expression in $CD44^+$ PC cells resulted in the inhibition of clonogenic growth and inhibition of metastatic behaviour and tumour regeneration (Liu *et al.,* 2011). The miR-34a was also established as an important negative regulator of $CD44^+$ PC cells, and it is assumed that this decreased expression of miR-34a in $CD44^+$ PC cells (including CSCs) plays an important role in PC development and metastasis (Liu *et al.,* 2011). Moreover, the expression of miR-34a antagomirs in $CD44^-$ prostate cancer cells promoted tumour development and metastasis (Liu *et al.,* 2011).

Taking into consideration the prevalent expression of CD44 in CSCs and the role of CD44 in mediating CSC migration and homing, suggesting its role in metastasis in various malignancies including PC and thus targeting of miR-34a pathways could become innovative treatment strategies for PC (Liu *et al.,* 2011). Further studies are under way in order to establish the molecular interplay between miRNAs and CSCs in PC and other human malignancies (Liu *et al.,* 2011).



1.10.3 The role of miRNAs in the processes of Epithelial-to-Mesenchymal Transition (EMT) in PC

The process of EMT is defined as one where epithelial cells acquire phenotypic characteristics of mesenchymal cells whereby the epithelial cells change their epithelial cobblestone phenotype to a mesenchymal elongated fibroblastic phenotype, which contributes to increased cell motility and invasion that is required for the metastatic process (Kong *et al.,* 2009; Kong *et al.,* 2010; Peng *et al.,* 2011). During EMT, cancer cells lose epithelial cell-cell junctions associated with a decrease in the expression of epithelial proteins such as E-cadherin and junction

plakoglobin, and the increase in expression of mesenchymal markers such as vimentin, fibronectin and α-smooth muscle actin.

These changes are also associated with augmented activity of matrix metalloproteinases (MMPs), such as MMP2, MMP3, MMP9, which leads to an invasive phenotype (Thiery *et al.,* 2009). All of the above mentioned processes lead to increased invasion and migration of tumours in many cancers including PC (Leshem *et al.,* 2011). Because the miRNAs are part of the cellular signaling circuit that controls EMT (Kalluri & Weinberg, 2009), it has been suggested that many miRNAs families, including the miR-200 and miR-205 families, play important roles in controlling EMT (Bracken *et al.,* 2009; Zeisberg & Neilson, 2009; Ahmad *et al.,* 2011).

The miRNAs 143 and 145 are two miRNAs that are assumed to play a role in EMT In PC. MiR-143 and miR-145 are dysregulated in the primary cancer when compared with normal prostate tissue (Ozen *et al.,* 2008; Clape *et al.,* 2009; Fu *et al.,* 2010; Schaefer *et al.,* 2010). The up-regulation of miR-143 in prostate cancer cells represses mesenchymal markers

(vimentin and fibronectin) and increases the epithelial marker E-cadherin (Peng *et al.,* 2011), while the up-regulation of miR-145 leads to the same effects except for in this case an increase in vimentin is seen (Peng *et al.,* 2011). Previous studies show that miRs-143 and -145 may be suppressors of EMT, and based on the fact that EMT plays an important role in invasion and migration consistent with mesenchymal characteristics allowing detachment and movement of cells from the primary tumour (Micalizzi *et al.,* 2010). Therefore, it can be assumed that miRs-143 and -145 inhibit invasion and migration in prostate cancer.

The role of most miRNAs in regulating EMT is still not clear with just a few miRNAs being evaluated for their role in EMT of PC (Kong *et al.,* 2009; Kong *et al.,* 2010; Peng *et al.,* 2011) (Table 1.1). With the crucial role of miRNAs in EMT, it is expected that the regulation of EMT by miRNAs in PC will be the focus of many research areas in the future.

Table 1.1: Regulation of EMT in PC by miRNAs

MiRNA	Effect in PC	Reference
MiR-200b	Down regulation of ZEB1, ZEB2 and Snail 2	Kong, 2009
MiR-200 family	Reverses EMT and Down regulates Notch-1 and Lin28B	Kong, 2010
MiR-143 and miR-145	Suppression of mesenchymal and up-regulation of epithelial markers	Peng, 2011
Let-7 family	Expression of Epithelial markers	Kong, 2010

1.10.4 The relationship between circulating miRNAs and tumour progression in prostate cancer

Circulating miRNAs have been suggested as encouraging biomarkers for the non-invasive diagnosis in many tumours (Brase *et al.,* 2011). It has been proposed that miRNAs profiles in tumour cells have prognostic value for some cancer patients, and a similar correlation with serum miRNAs profiles should emerge as a viable approach (Brase *et al.,* 2011). Brase *et al.,*2011 found that miR-375 was the top marker in a screening study (metastatic vs. localized prostate cancer) and that its expression is higher in prostate cancer tissue compared to normal epithelium (Brase *et al.,* 2011). The study also found that miR-375 was strongly related to the lymph-node status of the PC patients, no significant difference was observed in the serum levels of patients with Gleason score 7 and 8 tumours.

Based on these findings it seems that the circulating level of miR-375 is associated mainly with systemic disease (lymph node involvement and metastasis) rather than the grading of primary prostate cancer (Brase *et al.,* 2011). However, the definite role of miR-375 in prostate cancer is still not very clear.

miR-141 and miR-200b (both belonging to the same family of miRNAs) were reported to be the most frequently over regulated miRNAs in prostate epithelial cells in comparison with prostate stromal cells (Brase *et al.,* 2011). However, the circulating levels of miR-141 were found to be much higher in the serum of patients with high-risk tumours when compared to intermediate-risk samples. Also, circulating levels of miR-141 have been found to differentiate between patients with metastatic PC and healthy controls (Brase *et al.,* 2011).

Based on these limited studies, one could conclude that circulating miRNAs may offer a good perspective as non-invasive biomarkers for

tumour progression, including prostate cancer, but further research has to be done in this field in order to reach a better understanding of the role of miRNAs in the serum or plasma (Brase *et al.,* 2011).

1.11 The specific roles of certain miRNAs in prostate cancer

As mentioned above, miRNAs are known to play important roles in the progression of different cancers, including PC. Some miRNAs can function as tumour suppressors wherein their elevated levels are indicators of good prognosis (Brase *et al.,* 2011). On the contrary, other miRNAs are promoters of carcinogenesis and their expression levels are elevated in advanced stage of some cancers, which clearly suggests that these miRNAs may offer attractive targets for therapeutics and diagnostics. The figure below shows the categorisation of the various known miRNAs as oncogenes and tumour suppressors (Figure 1.11)

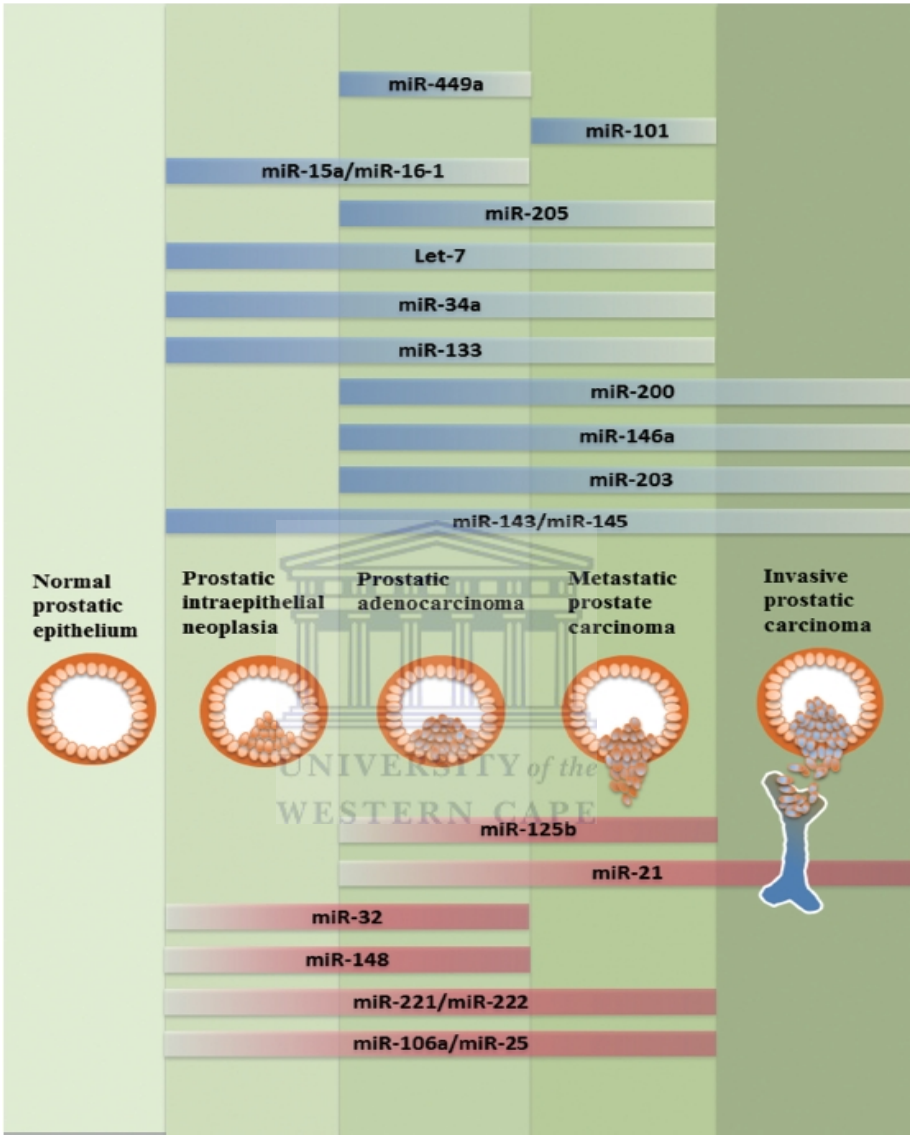


Figure 1.11 : Schematic representation of tumour suppressive miRNAs (blue) and oncogenic miRNAs (red) in the prostatic tumourigenesis progression from PIN to metastatic castrate-resistant prostate cancer CRPC (taken from Brase *et al.,* 2011)

1.11.1 Tumour suppressor miRNAs

1.11.1.1 The role of miRNA-15a and miRNA-16

Both of these miRNAs work as tumour suppressors mediated through deregulation of multiple oncogenes, which include: BCL2, MCL1, CCND1, and WNT3A (Aqeilan *et al.,* 2010). The aforementioned oncogenes can promote cell proliferation, cell survival and invasion (Aqeilan *et al.,* 2010). The down regulation of these miRNAs has been reported in many malignancies including: Chronic Lymphocytic Leukaemia (CLL), Pituitary adenoma, and Prostate carcinoma (Aqeilan *et al.,* 2010). The miR-15a and miR-16 are both located at 13q14.3, and the deletions at this location have been reported in many malignancies including: CLL, Multiple Myeloma (MM), Mantle cell lymphoma, and Prostate carcinoma (Aqeilan *et al.,* 2010).

In a recent study, the expression of miR-15a, miR-16-1 in PC samples showed consistent down regulation of these genes in around 80% of cancer samples compared with that of normal samples (Aqeilan *et al.,* 2010). Studies have also shown that miR-15a, miR-16-1 are down regulated in pituitary adenomas in comparison with normal pituitary, which basically enhances the assumption that they work as tumour suppressors and that

their knock down by allelic loss may contribute to tumourigenesis.

BCL2 is an oncoprotein that performs an important role in the genetic program of the eukaryotic cell. It prevents cell death and its over-expression was found to be related with many cancers such as: leukaemia, lymphoma, and carcinomas in general (Aqeilan *et al.,* 2010). It was reported that miR-15a, miR-16-1 sequences and BCL2 mRNA sequences shares a high degree of complementarity , and thus the previous information collectively suggests that miR-15a, miR-16-1 could suppress BCL2 by post transcriptional repression (Aqeilan *et al.,* 2010). It has been reported that miR-15a, miR-16-1 cluster targets not only BCL2 but also CCD1 (encoding cyclin D1) and WNT3A mRNAs, which promote many prostate carcinogenic features including; survival, proliferation, and invasion (Aqeilan *et al.,* 2010). The *in vivo* knock down of miR-15a, miR-16-1 resulted in hyperplasia associated with CCD1 and WNT3A up regulation. The aforementioned evidence suggest that loss of miR-15a and miR-16-1 may be a significant pathogenic event during the development of PC (Aqeilan *et al.,* 2010).

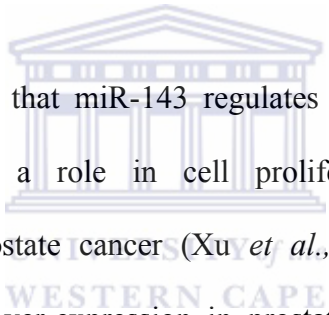
The miR-15 and miR-16 are usually down-modulated in the tumour sustaining stroma, an observation that can be explained by the effect of

cancer cells on the stroma (Musumeci *et al.,* 2011). The miR-15 and miR-16 have a tumour suppressor activity on both the cancer cell level and at the stromal microenvironment (Musumeci *et al.,* 2011). Recently, it was also proposed that miR-15 and miR-16 direct the expression of VEGF and IL-6, two factors that stimulate tumour angiogenesis and bone metastasis, respectively (Musumeci *et al.,* 2011). Moreover, it was shown that re-expression of miR-15 and miR-16 in cancer-associated fibroblasts (CAFs) will cause attenuation of the stromal support capability, and this will result in a decrease in cell proliferation and migration in primary and metastatic tumours (Musumeci *et al.,* 2011). These observations lead to the conclusion that in the context of prostate cancer, miR-15 and miR-16 are tumour suppressors, at least, on two levels such as at the levels of tumour cell and stromal cells.

1.11.1.2The role of miR-143 and miR-145

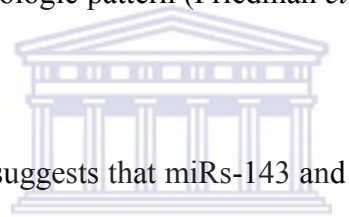
Studies have shown that miR-143 is considerably decreased in PC, and its expression is further decreased during cancer progression (Clape *et al.,* 2009). K-RAS (V-Ki-ras2 Kirsten rat sarcoma), a key molecule of EGFR/RAS/mitogen-activated protein kinase (MAPK) signalling pathway, is a viral oncogene homolog that was incriminated in cell proliferation and

migration in response to growth factors. The MAPK pathway also works at another level through its effect on Androgen receptors (AR), where it increases AR in response to low androgen, and this is considered a main process in Androgen derivation therapy relapse. K-RAS is a potential target of miR-143 (Friedman *et al.,* 2009), thereby lower levels of miR-143 in prostate cancer cells may be incriminated in carcinogenesis due to the lack of its inhibitory effect on K-RAS and MAPK pathways.



Xu et al. 2011 showed that miR-143 regulates K-RAS, p-ERK1/2, and cyclin D1 and plays a role in cell proliferation, migration, and chemosensitivity in prostate cancer (Xu *et al.,* 2011). The study also showed that miR-143 over-expression in prostate cancer cells represses proliferation and migration, thereby augmenting sensitivity to docetaxel by affecting EGFR/RAS/MAPK pathway. The expressions of miRNAs-143 and -145 (another miRNA which is assumed to be a tumour suppressor) were found to be down-regulated considerably in metastatic samples (Friedman *et al.,* 2009). Exploring the correlation of the levels of miRNAs-143 and -145 with clinico-pathological features of PC showed that down-regulation of miRNAs-143 and -145 were negatively associated with bone

metastasis, the Gleason score and the levels of free PSA in primary PC patients (Friedman *et al.,* 2009). Over-expression of miR-143 and -145 by retroviral transfection decreased the ability of migration and invasion *in vitro*, and tumour development and bone invasion *in vivo* of PC-3 cells (a human PC cell line originated from a bone metastatic PC specimen) (Friedman *et al.,* 2009). Their up-regulation also enhanced E-cadherin expression and decreased fibronectin expression in PC-3 cells with features of a less invasive morphologic pattern (Friedman *et al.,* 2009).



The information above suggests that miRs-143 and -145 are related to bone metastasis of PC and may play a biological role in this process (Friedman *et al.,* 2009). One could postulate that the possibility of using them in the clinical setting as biomarkers to individualize different stages of human PC and could predict the development of bone metastasis in patients well ahead of time (Friedman *et al.,* 2009). Even though up-regulation of miRNAs-143 and -145 are found to suppress the aggressiveness and EMT of PC-3 cells with regard to bone metastasis, it did not have the same effects on LNCaP cells that was derived from metastatic lymph node (Friedman *et al.,* 2009). Deregulation of miRs-143 and -145 was not found in lymph node

metastasis in comparison to primary PC specimens. This information suggests that miR-143 and -145 may have a cell type-specific activity and could suppress only bone metastasis without inhibiting lymph node metastasis, and that the loss of these miRNAs could selectively promote metastasis, which could be due to deregulated expression of other miRNAs such as miR-221 (Friedman *et al.,* 2009).

1.11.1.3 The role of miRNA-200

Recent studies have shown that the miR-200 family controls epithelial-mesenchymal transition (EMT) by targeting zinc-finger E-box binding homebox 1 (ZEB1) and ZEB2 (Kong *et al.,* 2009; Ahmad et al. 2011; Ahmad et al 2011) . There is enough evidence to suggest that the processes of EMT can be elicited by various growth factors, such as transforming growth factor β and platelet-derived growth factor-D (PDGF-D), which is expressed in PC tissue (Kong *et al.,* 2009). It was shown that over-expression of PDGF-D in PC3 cells (prostate cancer cell lines with high metastatic potential) leads to the acquisition of the EMT phenotype (Kong *et al.,* 2009).

It was also proved that significant down-regulation of the miR-200 family in PC3 PDGF-D cells and in PC3 cells exposed to purified active PDGF-D protein, resulted in the up-regulation of ZEB1, ZEB2, and Snail2 expression (a transcription factor which belongs to the snail protein family and plays critical roles in the formation of tissues during embryonic development) (Kong *et al.,* 2009). Interestingly, re-expression of miR-200b in PC3 PDGF-D cells led to reversal of the EMT phenotype accompanied with the down-regulation of ZEB1, ZEB2, and Snail2 expression, and all of these events were associated with greater expression levels of epithelial markers (Kong *et al.,* 2009). Moreover, it was proved that transfection of PC3 PDGF-D cells with miR-200b considerably decreased the expression of ZEB1, ZEB2, and Snail2 at both the mRNA and protein levels, with simultaneous greater expression levels of epithelial markers such as E-cadherin, stratifin, CRB3, EPCM, F11R, and connexin 26, all of these events collectively led to the inhibition of cell migration and invasion (Kong *et al.,* 2009). In a breast cancer model, studies have shown that *in vivo* manipulation of miR-200b leads to significantly reduced pulmonary metastases of breast cancer cells (Ahmad *et al.,* 2011) which further supports the role of the miR-200 family in metastases of human cancers.

It is well known that NF-κB plays an essential role in facilitating the processes of EMT induced by different factors through up-regulation of ZEB1 and ZEB2, which in turn suppress the expression of miR-200 family members by binding to the E-box sequence of the miR-200 promoter (Kong *et al.,* 2009). Whereas miR-200 can down-regulate the expression of ZEB1 and ZEB2 by interacting with the 3'-UTR of ZEB1 and ZEB2 mRNA (Kong *et al.,* 2009). All of these findings suggest a double-negative feedback loop between miR-200 and ZEB1/ZEB2 that permits the preservation of the EMT phenotype, even after withdrawal of the initial inducing signal, which might become a critical target for the reversal of EMT (Ahmad *et al.,* 2011). From these facts, it can be concluded that PDGF-D-stimulated attainment of the EMT phenotype in PC3 cells is, partly, an outcome of suppression of miR-200 and that new strategies in which miR-200 would be up-regulated will become an auspicious approach for the treatment of invasive prostate cancer (Kong *et al.,* 2009).

1.11.1.4 The importance of miRNA-488

Another miRNA that has been associated with PC is miR-488. The miR-488 is encoded by the AsTN1 gene. Two mature molecules result from the

processing of its precursors: miR-488* and miR-488 with both forms expressed predominantly in human brain tissue (Sikand *et al.,* 2010). Studies have shown that miR-488* was not expressed in several prostate cancer cell lines, although the aetiology is still unclear and thus it is currently under further investigation (Sikand *et al.,* 2010).

Androgen receptor (AR) is a direct target of miR-488, as miR-488 has a binding site at the 3'UTR of the AR gene where it binds and suppresses its expression (Sikand *et al.,* 2010). It was shown that cells transfected with miR-488 result in reduced expression of AR in both Androgen-dependent (LNCaP) and Androgen-independent (C4-2B) PC cells. In both the cell lines, treatment with miR-488 mimics was found to retard the growth of these cells, but this was not the fact with AR-negative DU145 cells (Sikand *et al.,* 2010), suggesting the regulatory role of miR-488 on AR expression. These results suggest that miR-488 could function as a tumour growth suppressor, which is mediated by deregulation of AR expression (Sikand *et al.,* 2010).

Although it is still too premature to make conclusions, the results of these studies clearly showed that miR-488* transfection into LNCaP and C4-2B cells led to the repression of AR expression, thereby suggesting that development of ways to increase the levels of endogenous miR-488* could have a great impact on designing novel treatment strategies for PC.

1.11.2 Tumour promoter oncomiRNAs

The miRNA expression profiling analyses have shown that many miRNAs are up-regulated in prostate cancer (Pang *et al.,* 2010). These oncogenic miRNAs suppress the apoptosis-related genes, so their over-expression leads to increased tumour growth and metastasis (Pang *et al.,* 2010).

1.11.2.1 The role of miRNA -221 and miRNA-222

The miR-221 and miR-222 are both considered as oncogenic and were found to be associated with the development and metastasis of prostate cancer (Pang *et al.,* 2010). One of the methods through which these miRNAs elicit their effect is by binding to one of their target mRNA, p27kip1 and cause suppression, which results in tumour growth (Galardi *et al.,* 2007; Sun *et al.,* 2009) . Another suggested action for these molecules is their role in the development or maintenance of castration-resistant

prostate cancer (CRPC) phenotype through a mechanism that is not yet clearly understood, although it may be through influencing response of AR-mediated signalling in prostate cancer cells (Sun *et al.,* 2009).

Sudies have shown that miR-221 levels are up-regulated in both Androgen Dependant Prostate Cancer (ADPC) and Androgen Independent Prostate Cancer (AIPC) compared to normal controls (Zheng *et al.,* 2011). One pathologic process that plays an important role in the carcinogenesis and hormone therapy failure in PC is neuroendocrine differentiation (NE), a process that is associated with tumour progression and poor prognosis. Studies have shown that miR-221 is capable of inducing NE differentiation in LNCaP cells in an androgen deprived environment, which may lead to Androgen Independence (AI) (Zheng *et al.,* 2011). It was revealed that miR-221 stimulates the growth of LNCaP and LNCaP-AI cells, and it is consistent with findings that the ectopic introduction of miR-221 in low expressing LNCaP cells, bolstered their growth potential by inducing a G1-S shift in the cell cycle (Zheng *et al.,* 2011).

One paradox to mention about miR-221 is that although the expression of miR-221 is higher in LNCaP-AI cells (which are more invasive) compared to LNCaP (which are less invasive), suggesting that miR-221 promotes invasion of PC cells (Zheng *et al.,* 2011), whereas the up-regulation of miR-221 in LNCaP cells did not increase its ability of increased cell migration, whilst the invasion capacity of LNCaP-AI cells was deregulated by knock-down of miR-221 expression.

The aforementioned facts can be explained by assuming that other pathways are involved in regulating the migration capability of cells during the progression of ADPC into AIPC. It was further concluded that miR-221 could influence PC via regulation of DVL2 (Dishevelled 2), which in summary is an important intracellular mediator of the WNT signalling pathway. An important target gene for WNT is MMP-7 (a regulator of cellular adhesion molecules which controls cellular adhesion, invasion, and migration), and the activation of MMP-7 was found to greatly strengthen the capability to destroy the extracellular matrix, especially in cancer cells (Zheng *et al.,* 2011), suggesting that miR-221 may play an important role in the regulation of invasion and metastasis.

1.11.2.2The role of miR-21

Another onco-miRNA is miR-21; it is usually up-regulated in prostate cancer and plays a role in tumour growth, invasion, and metastasis (Si *et al.,* 2007; Li *et al.,* 2009). Recently, miR-21 was individualized as an oncogene which is up-regulated in various cancers (glioma, breast cancer, colorectal cancer, stomach/gastric cancer, hepatocellular carcinoma, pancreatic cancer, lung cancer, cholangiocarcinoma, leukemia, and prostate cancer. (Li *et al.,* 2009). Anti-sense studies of miRNA-21 in glioblastoma cell lines revealed that it directs cell growth by inhibiting apoptosis while it does not influence cell proliferation (Li *et al.,* 2009). Previous studies have revealed an increase in apoptosis in DU145 and PC-3 cells after blocking miR-21 function, while LNCaP cells, which have low level of miR-21, showed no changes in apoptosis in response to miR-21 blockade (Si *et al.,* 2007; Li *et al.,* 2009). The information mentioned above insinuated that miR-21 plays an important role in the resistance to apoptosis observed in DU145 and PC-3 cells (Li *et al.,* 2009). These results also suggest that miR-21 might have a role to play in AR-negative cells but this needs to be further investigated.

Many genes have been identified as targets of miR-21 in carcinogenesis including; tropomyosine (TPM1), Programme cell death protein 4 (PDCD4), and MARCKS (a gene which encodes myristoylated alanine-rich C-kinase substrate) (Li *et al.,* 2009). It has been shown that MARCKS has high frequency frame-shift mutations during carcinogenesis in hereditary nonpolyposis colorectal cancer (HNPCC) (Li *et al.,* 2009). It was also revealed that the expression of MARCKS is down-regulated in Hepatocellular carcinoma (HCC) tissues in comparison with cirrhotic, benign liver tissues (Li *et al.,* 2009). MARCKS plays a part in TPA-mediated cellular migration in neuroblastoma, probably through its effect as a downstream target of Protein kinase C epsilon (Li *et al.,* 2009).

All these results suggest that MARCKS plays a role in tumourigenesis, and the effect of miR-21 on PC cell motility and invasion may in part be due to its regulation of MARCKS gene. Recent studies have also shown that in the presence of androgen, AR can bind to miPPR-21, a miR-21 promoter, and this results in the over-expression of miR-21 at its transcription level, leading to castration resistance (Ribas *et al.,* 2009). In support of what has

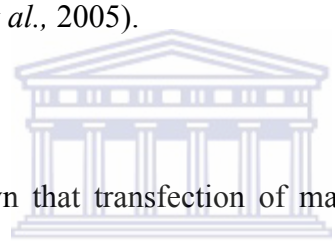
been mentioned above, anti-miR-21 may augment the sensitivity of prostate cancer cells to apoptosis (Li *et al.,* 2009), and also negatively affect the motility and invasive characteristic of cancer cells (Li *et al.,* 2009).

In summary, it appears that miR-21 plays an essential role in apoptosis and metastasis of PC (Li *et al.,* 2009). Since no effective therapy is available to cure PC, in part due to the resistance of androgen-independent advanced prostate cancer cells to apoptotic death, gene therapy that targets miR-21 may be a potential alternative therapy for androgen-independent PC that have an up-regulated expression of miR-21. Moreover, it is tempting to speculate that natural agents (nutraceuticals) could serve as a therapeutic strategy for PC because many of these agents could inhibit the expression of miR-21 as suggested by recent results in cancer of the pancreas (Boa *et al.,* 2011).

1.11.2.3 The role of miRNA-125b

Another onco-miRNA is miR-125b which is considered very important for cell proliferation (Lee *et al.,* 2005) and it is over-expressed in prostate cancer (Pang *et al.,* 2010). It has been reported that the reduction of miRNA-125b was found to be associated with the regulation of cellular proliferation of cancer cells, and this effect was attenuated by co-

transfection of mature miRNA (Lee *et al.,* 2005). The biological role of miR-125b in PC-3 cells was studied and it was revealed that the depletion of miR-125b by numerous transfections of si-125b2 was followed by a substantial proliferation defect in PC-3 cells (Lee *et al.,* 2005). This growth defect was not associated with aberrant accumulation of cells in one stage of cell cycle or by apoptosis. Also, the depletion of miR-125b by 2'-O-methyl oligonucleotide showed the same phenotype with a similar effect on cell proliferation (Lee *et al.,* 2005).



Studies have also shown that transfection of mature synthetic miR-125b causes PC cell growth (Lee *et al.,* 2005; Ali *et al.,* 2010), and this was in part due to its effect on the 3'UTR of BAK1 (a pro-apoptotic member of the BCL-2 gene family that is involved in initiating apoptosis) transcript (Shi *et al.,* 2007). Nonetheless, down-regulation of BAK-1 only, could not attenuate miR-125b's growth stimulatory effect, suggesting that there are other targets for miR-125b in prostate cancer cells (Lee *et al.,* 2005; Shi *et al.,* 2007). Some recent reports individualized EIF4EBP1 (Eukaryotic translation initiation factor 4E-binding protein 1, a gene that encodes one member of a family of translation repressors proteins) as another specific

target for miR-125b in PC (Ozen *et al.,* 2008), but it is still not conclusive, suggesting that future in-depth investigation is warranted.

1.12. Conclusions and perspectives

Prostate cancer (PC) is considered to be the most diagnosed cancer (Peng *et al.,* 2011) and the second leading cause of cancer death in men older than 40 years of age in the USA (Hao *et al.,* 2011). Similarly in South Africa PC is rated the most common cancer in Black African men and the second most common in white South African males with more than 4300 newly diagnosed cases annuallly and nearly 2000 deaths associated with the disease (ACS, 2012). The major problem of PC is the development and acquisition of castrate resistant prostate cancer (CRPC) phenotype which eventually leads to the development of skeletal metastasis (mCRPC), at which point it becomes an incurable disease (Peng *et al.,* 2011). Therefore, investigations are underway to find the molecular basis of mCRPC so that novel therapeutic strategies could be devised. To that end many novel molecules are being tested and interrogated, among which microRNAs (miRNAs) are becoming an attractive area of research.

The roles of miRNAs in PC are becoming clearer by understanding the interactions between miRNAs and their targets and the resulting impact on carcinogenesis of the prostate (Shi *et al.,* 2007; Lu *et al.*, 2008; Pang *et al.,* 2010). It is believed that several miRNAs and their targets are aberrantly expressed in PC which, in turn, alter the cellular growth, invasion, and metastatic potential of prostate cancer cells. The abnormal expressions of certain miRNAs are now considered valuable biomarkers for diagnosis, prognosis and classification of PC (Mattie *et al.,* 2006; Porkka *et al.,* 2007; Pang *et al.,* 2010). All of the above information underscores the importance of the biology of miRNAs in PC. Their specific abnormalities, and how one could regulate their expressions will likely become novel avenues by which newer therapeutic strategies could be developed for the treatment of PC.

Research investigations focused on miRNAs have suggested a strong prognostic and therapeutic importance of miRNAs in PC. Such studies have established an intimate relationship between prostate cancer and miRNA with emerging data clearly suggesting that miRNA is a very promising field although further in-depth mechanistic studies are required to ascertain the

role of specific miRNA(s) and relevant target(s) in the development and progression of PC especially the emergence of metastatic castrate resistant prostate cancer. Once the additional scientific knowledge is gained it would be easier to focus on the development of strategies for up-regulation and down-regulation of specific miRNA as a novel and targeted therapeutic approach for the treatment of PC. For example, recent studies with miR-21 in pancreatic and colon cancer models have shown that targeted regulation of this miRNA can lead to effective anticancer therapy (Luscombe *et al.,* 2001; Ali *et al.,* 2010; Bao *et al.,* 2011; Ali *et al.,* 2011; Ali *et al.,* 2012; Yu *et al.,* 2012).

Based on these studies we hypothesise that the identification of key miRNAs and their targets that are intimately involved with the development and progression of PC and parallel development of strategies for deregulation of miRNAs would allow inhibition of tumour growth, invasion, angiogenesis and metastasis in PC, which is expected to be clinically useful for the management of patients diagnosed with PC.

The main aim of this study is therefore to identify specific miRNAs as

biomarkers for the early detection of prostate cancer. Individual objectives include:

- To identify miRNA and gene targets that play a critical role in disease to further understand their mechanism of action in PC using several *in silico* methods.
- To validate the potential diagnostic miRNAs in cell lines using molecular techniques.

CHAPTER 2

Identification of miRNAs as key indicators for the presence of Prostate Cancer (PC) using in silico methods

2.1 Introduction

2.1.1 Problem identification

2.1.2 Current diagnosis method

2.1.3 OMICs based technologies for biomarker identification

2.1.4 Expression Biomarkers

2.1.5 MiRNA biomarkers

2.2 Methods

2.2.1 Database search and data retrieval

2.2.2 Target gene prediction using the 40 identified prostate miRNAs

2.2.3 Prioritising the miRNA and target gene lists

2.2.4 Functional Annotation of predicted gene targets using DAVID

2.2.5 Co-expression analyses using the STRING databases

2.2.6 Expression profiling for genes targeted by miRNAs

2.2.7 Pathway Analyses using mirPath in DIANA tools

2.3 Results

2.3.1 Data retrieval

2.3.2 Target predictions

2.3.3 Prioritization of target gene lists

2.3.5 Co-expression Analysis Using STRING

2.2.6 Expression profiling of 21 target genes using GEA

2.2.7 Pathway Analysis using DIANA-mirPath

2.4 Discussion

CHAPTER 2

Identification of miRNAs as key indicators for the presence of Prostate Cancer (PC) using *in silico* methods

2.1 Introduction

2.1.1 Problem identification

Cancer is a class of diseases characterized by uncontrollable cell growth. There are over 100 different types of cancer, that are classified by the type of cell that it initially affects (Meltzer, 2005; Croce, 2005). For the purpose of this research we will be focussing on prostate cancer (PC). PC is the second most common form of cancer in men around the world (Figure 2.1 and Figure 2.2) and approximately 4500 men in South Africa are annually diagnosed with the disease making PC a global epidemic (American Cancer Society, 2012) (Figure 2.1). PC follows a natural course, starting as a tiny group of cancer cells that can grow into a tumour. In some men if PC is not treated it may spread to surrounding tissue by a process called direct invasion or spread and could lead to death (American Cancer Society, 2009).

2.1.2 Current diagnostic method

At present, diagnosis of PC derives from serum prostate-specific antigen (PSA) (Maricone, 2011) measurement, digital rectal examination histopathological evaluation of prostate needle biopsies (Essink-Bot *et al.,* 1998) and ultra sound examination of tumours (Bonekamp *et al.,* 2011). Organ-confined prostate cancer is treated with surgery (radical prostatectomy) or radiation therapy, and an increase in PSA level usually represents the first sign of recurrence of disease following treatment (Damber and Aus, 2008). Androgen deprivation is the current therapy of choice for patients with metastatic disease. The PSA test has been examined in several observational settings for initial diagnosis of disease, as a tool to monitor for recurrence after initial therapy, and for prognosis of outcomes after therapy. There is no specific PSA value which can provide assurance that there is no risk of developing PC. Parameter estimates for this test include sensitivity in the range of 70% (Partin and Oesterling, 1994).

Due to the high genetic variability in the course of PC, development or progression, novel biomarkers are strongly needed at the time of diagnosis to facilitate treatment planning. However, current prognostic tools (nomograms) exclusively rely on pathological and clinical parameters

(Damber and Aus, 2008). Therefore, understanding the molecular alterations that distinguish the progressive from a non-progressive disease will allow not only the identification of novel biomarkers to be included in clinical nomograms for improving their predictive power, but will also provide mechanistic information for the discovery of new therapeutic targets and the design of tailored therapeutic interventions.

Several diagnostic methods are currently being employed, however they are invasive and lack specificity and sensitivity (American Cancer Society, 2012; Djulbegovic *et al.,* 2010). Therefore, the need for a less invasive, early detection method with the ability to overcome the lack of specificity and sensitivity is required. Biomarkers have recently been identified as a viable option for the early detection of disease which acts as biological indicators for disease i.e. DNA, RNA, proteins and microRNAs (miRNA). Eradication of advanced PC still represents an unsolved clinical problem, making the development of alternative treatment approaches highly desirable.

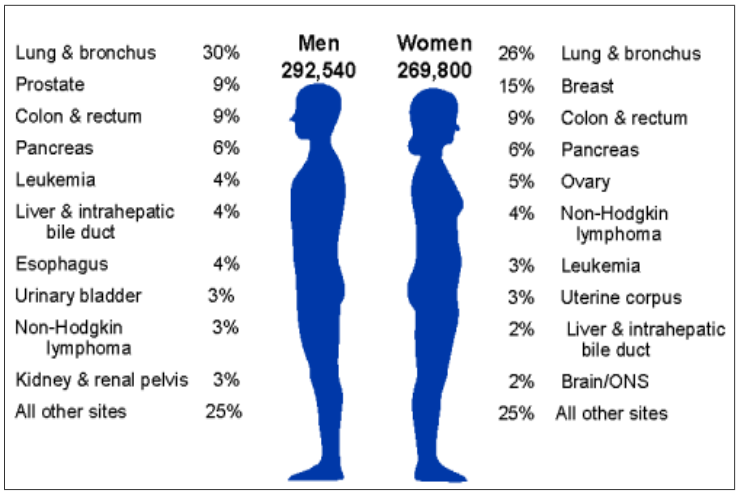
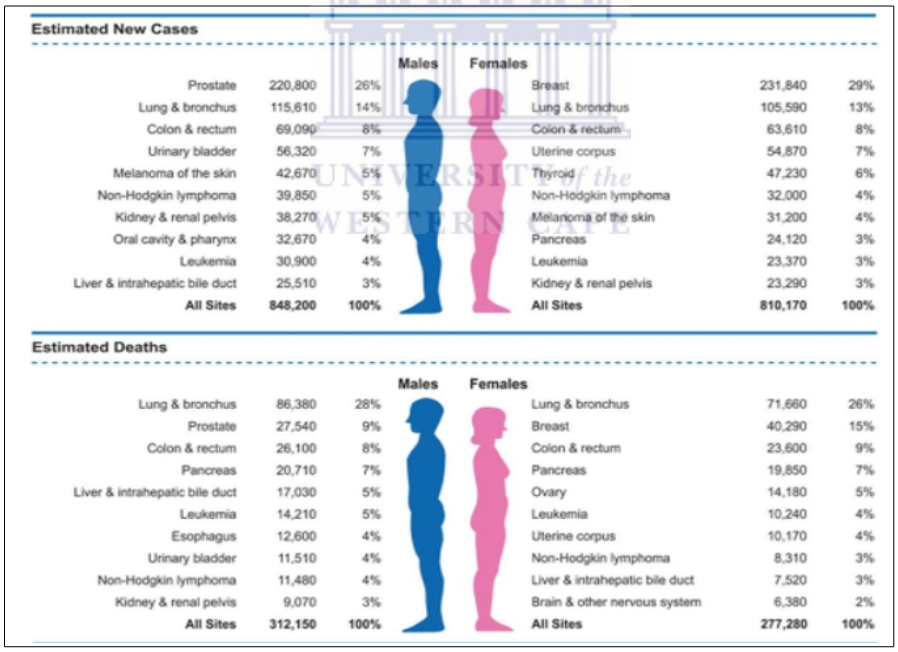


Figure 2.1: Analysis of cancer deaths in the United States of America for the period 2004-2009 (adapted from American Cancer Society, 2009)



Estimated New Cases

Males			Females		
Prostate	220,800	26%	Breast	231,840	29%
Lung & bronchus	115,610	14%	Lung & bronchus	105,590	13%
Colon & rectum	69,090	8%	Colon & rectum	63,610	8%
Urinary bladder	56,320	7%	Uterine corpus	54,870	7%
Melanoma of the skin	42,670	5%	Thyroid	47,230	6%
Non-Hodgkin lymphoma	39,850	5%	Non-Hodgkin lymphoma	32,000	4%
Kidney & renal pelvis	38,270	5%	Melanoma of the skin	31,200	4%
Oral cavity & pharynx	32,670	4%	Pancreas	24,120	3%
Leukemia	30,900	4%	Leukemia	23,370	3%
Liver & intrahepatic bile duct	25,510	3%	Kidney & renal pelvis	23,290	3%
All Sites	**848,200**	**100%**	**All Sites**	**810,170**	**100%**

Estimated Deaths

Males			Females		
Lung & bronchus	86,380	28%	Lung & bronchus	71,660	26%
Prostate	27,540	9%	Breast	40,290	15%
Colon & rectum	26,100	8%	Colon & rectum	23,600	9%
Pancreas	20,710	7%	Pancreas	19,850	7%
Liver & intrahepatic bile duct	17,030	5%	Ovary	14,180	5%
Leukemia	14,210	5%	Leukemia	10,240	4%
Esophagus	12,600	4%	Uterine corpus	10,170	4%
Urinary bladder	11,510	4%	Non-Hodgkin lymphoma	8,310	3%
Non-Hodgkin lymphoma	11,480	4%	Liver & intrahepatic bile duct	7,520	3%
Kidney & renal pelvis	9,070	3%	Brain & other nervous system	6,380	2%
All Sites	**312,150**	**100%**	**All Sites**	**277,280**	**100%**

Figure 2.2: Estimation of new cancer cases and deaths in the United States for the period 2009-2014 (adapted from American Cancer Society, 2014)

2.1.3 OMICs based technologies for biomarker identification

The rapid evolution of high-throughput technologies designed for the screening of biomedical samples with the latest breakthroughs in whole genome sequencing and microRNA (miRNA) profiling has given birth to a number of biological disciplines devoted to the generation and study of multiple OMICs generated data for the identification of potential biomarkers (Dayeti *et al.,* 2012). A number of OMICS fields have been dedicated to identifying various types of biomarkers for the detection of diseases. These biomarkers include genetic, expression, protein, metabolomic and miRNA biomarkers (Lui *et al.,* 2008; Dayati *et al.,* 2012). This study will focus on various strategies that have been used to identify miRNA biomarkers. Figure 2.3 summarizes the latest technologies used as well as diversification of the biomarker types and underlying data types depending on the nature of changes detected by the respective technology.

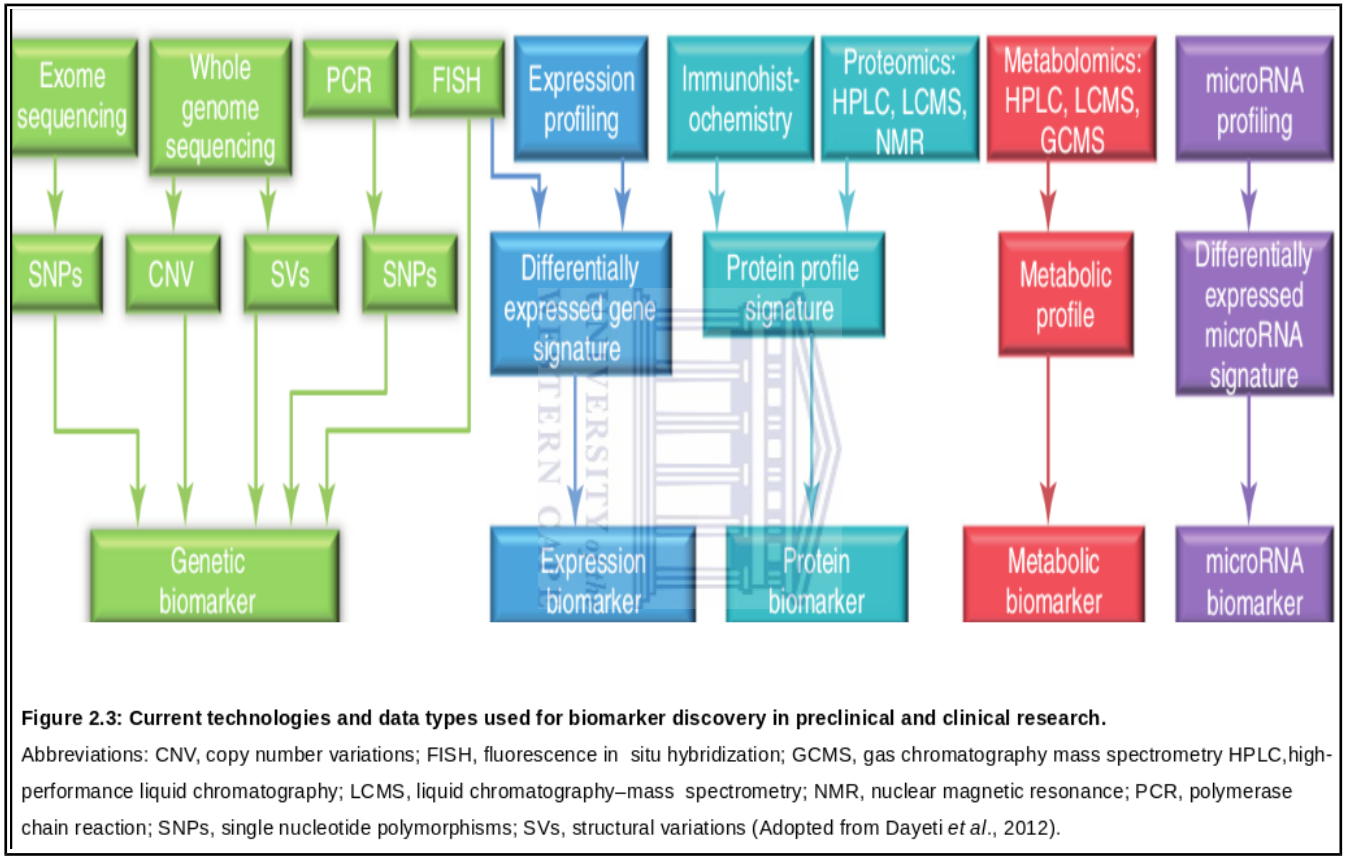


Figure 2.3: Current technologies and data types used for biomarker discovery in preclinical and clinical research.

Abbreviations: CNV, copy number variations; FISH, fluorescence in situ hybridization; GCMS, gas chromatography mass spectrometry HPLC,high-performance liquid chromatography; LCMS, liquid chromatography–mass spectrometry; NMR, nuclear magnetic resonance; PCR, polymerase chain reaction; SNPs, single nucleotide polymorphisms; SVs, structural variations (Adopted from Dayeti *et al*., 2012).

2.1.4 Expression Biomarkers

Expression profiling studies using microarrays and other methods have shown that miRNAs are dysregulated in a wide variety of human cancers. In some instances, the expression of selected miRNAs or specific miRNA signatures was found to correlate with diverse clinico-pathological features of disease and were able to predict patient clinical outcome and/or response to treatment (Garzon *et al.,* 2006). Such findings have highlighted the potential of miRNAs as new diagnostic or prognostic/predictive biomarkers. Moreover, the role of miRNAs functioning as oncogenes and tumour suppressors, as emerged from functional studies in experimental models (Garzon *et al.,* 2006), has generated great interest in their possible use as novel targets or tools for anticancer therapies.

2.1.5 MiRNA biomarkers

MiRNAs are a class of naturally occurring, small non-coding RNA molecules, about 18–25 nucleotides in length. MiRNAs are partially complementary to one or more messenger RNA (mRNA) molecules, and their main function, to down regulate gene expression, is well documented, that include translational repression, mRNA cleavage, and de-adenylation (Flopowics, 2005; He *et al.,* 2005; Mraz *et al.,* 2009). The involvement of

miRNAs in key cellular processes such as proliferation and cell death and their well known negative control over the expression of numerous oncoproteins make them prime candidates as cancer biomarkers (Flopowics, 2005; He *et al.,* 2005). It has also been reported that cancer-specific miRNAs are detected in the blood at the earliest stages of tumour development and increase in their concentration as the tumour progresses over time, thus making them an indicator of the tumour growth (Krutovskikh and Herceg, 2010). Moreover, unlike other types of biomarkers, miRNAs are remarkably stable in circulation, formalin-fixed paraffin embedded tissue, serum and urine thus making them potentially robust oncology biomarkers (He *et al.,* 2005). Functional miRNA species have mostly been validated *in vitro* using luciferase reporter activity (Krutovskikh and Herceg, 2010).

Micro-array profiling is a powerful high-throughput technology capable of monitoring the expression of thousands of small non-coding RNAs at once within tens of samples processed in parallel in a single experiment (Lui *et al.,* 2008). Serial Analysis of Gene Expression (SAGE), Stem-loop quantitative Real Time Polymerase Chain Reaction (qRT-PCR) for mature miRNAs, qRT-PCR for precursor miRNAs and bead-based technologies are

also frequently used for miRNA profiling (Chang-Gong *et al.,* 2008). However, no miRNA biomarker currently exists in cancer clinical practice (Dayeti *et al.,* 2012). It is noteworthy that the field of genomics generating genetic biomarkers witnessed a remarkable rise in clinical acceptance after the sequencing of all human genes through the human genome project (Dayeti *et al.,* 2012).

Similar effort is needed to discover and characterize all the miRNAs in human cells in order to transform the potential of miRNAs as cancer biomarker into clinical success. Further understanding on how miRNAs compete with proteins to bind and control the expression of mRNA as well as the functional interaction networks through which miRNAs exert their tissue specific role, is needed for future clinical translation of these molecules (Krutovskikh and Herceg, 2010).

Although transcriptomics technology (micro-array analysis), is one of the oldest and widely used high throughput technologies, most candidate biomarkers are reported in genomic and by proteomics research (Figure 2.4) (Dayeti *et al.,* 2012). Stability of the signal from genomic analysis as

well as higher stability of the protein versus mRNA might be the reason for those biomarkers outweighing the transcriptomics derived biomarkers. Comparing the number of approved biomarkers to those mentioned in the public domain (e.g. GIOBIOM, (GV BIO Online Biomarker Database)) available at http://www.gvkbio.com, reveals that the majority of candidate biomarkers either failed or have not reached the clinical setting yet (Dayeti *et al.,* 2012).

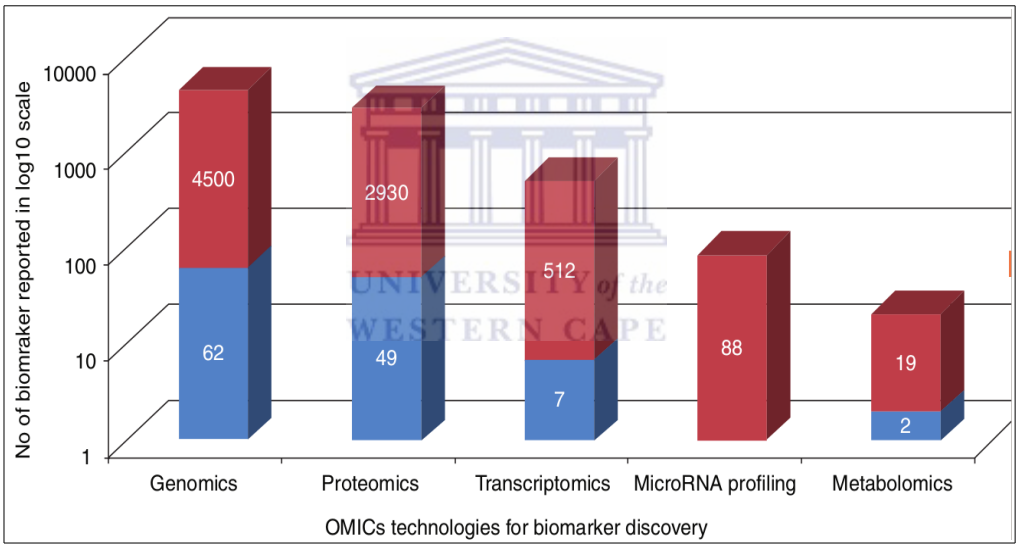


Figure 2.4: Current contribution of OMICs technologies in oncology biomarker discovery. In red: total number of candidate biomarkers reported in the public domain. In blue: number of FDA approved biomarkers in current clinical practice for oncology (Adopted from Dayeti *et al.,* 2012).

Even in the event of a strong signal being derived from high-throughput technologies, its conversion to clinical practice meets a number of associated challenges, the first of which is its functional interpretation. Interpretation of the high-throughput data in the context of molecular patho-physiology of underlying disease and specific treatment is the current rate-limiting step in the biomarker identification and validation pipeline (Dayeti *et al.,* 2012). If properly identified, extracted and interpreted, OMICs data sets can provide valuable biological insights. Functional analysis of OMICs data requires knowledge on the molecular interactions and pathways underlying pathophysiology of diseases and the mode of action of treatment regimes.

Accumulated biological knowledge across different levels, needs to be collected, annotated, translated into a computer-readable format, and stored in a semantically enhanced knowledge base. A knowledge base of this magnitude could then be used for knowledge-based analysis of OMICs data sets through integrative approaches that aims at finding key biological processes, pathways, interaction modules or causative network signatures within disease onset and progression. In addition identifying molecules

involved in these pathways and processes could lead to potential candidate biomarkers.

In this study we aim to provide an integrative method to successfully identify miRNAs as specific biomarkers associated with the early detection of PC through data mining and various *in silico methods (*programming and statistical analyses).

2.2 Methods

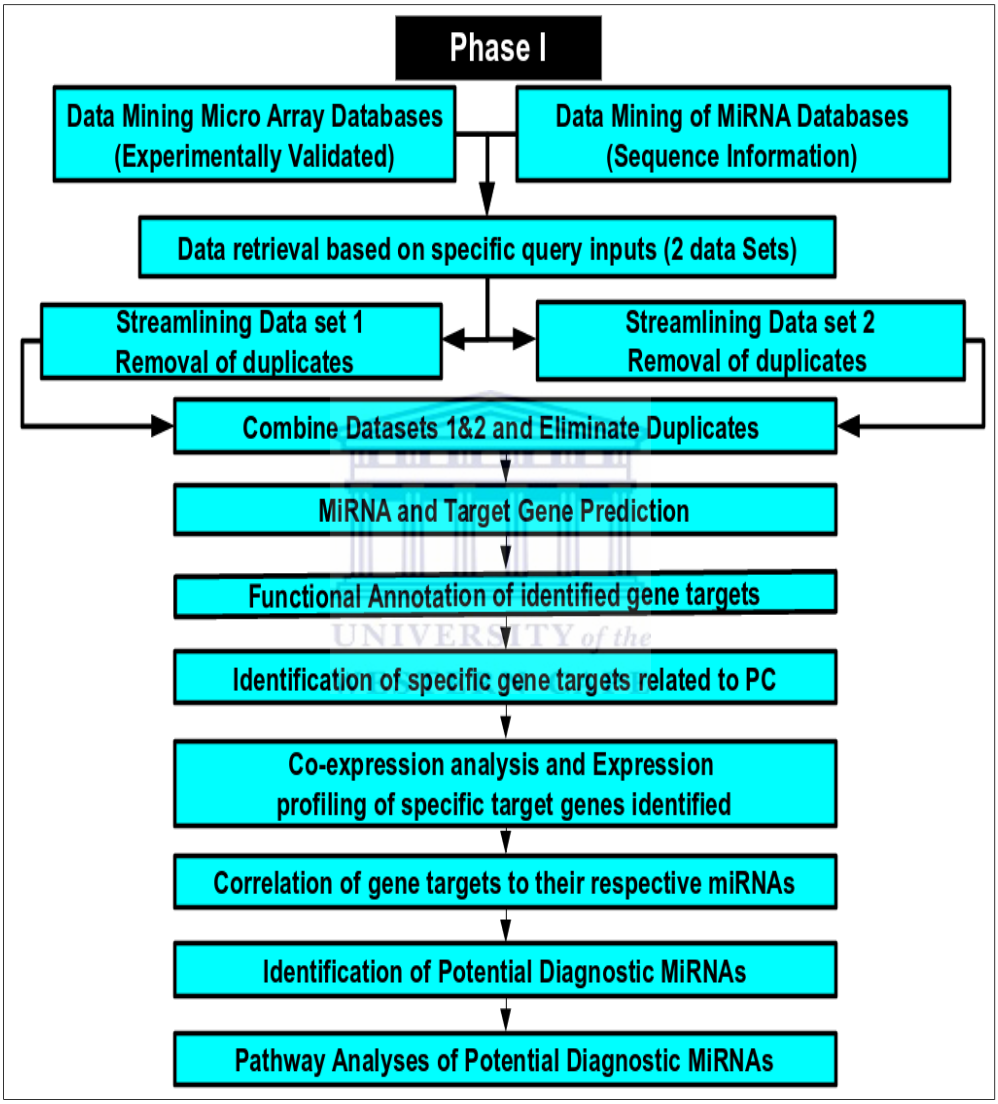


Figure 2.5: Flow Chart representing the *in silico* methods employed during Phase 1 of this study.

2.2.1 Database search and data retrieval

Publicly available databases such as MiRBase (http://www.mirbase.org/search.html) (Griffiths-Jones *et al.,* 2006) and the Gene Expression Omnibus (GEO) available at http://www.ncbi.nlm.nih.gov/geo (Barrett *et al.,* 2013) were scanned for miRNAs and their implication in PC. The miRBase database was searched for sequence specific info that implicated known miRNAs in PC. GEO was scanned for micro-array data available on known miRNAs that have previously been implicated in PC.

The miRBase database is a searchable database of published miRNA sequences and annotation. Each entry in the miRBase Sequence database represents a predicted hairpin portion of a miRNA transcript (termed mir in the database), with information on the location and sequence of the mature miRNA sequence (termed miR) (Griffiths-Jones *et al.,* 2008). Both hairpin and mature sequences are available for searching and browsing, and entries may be retrieved by name, keyword, references and annotation. All sequence and annotation data are also available for downloading from the

miRBase registry (Griffiths-Jones, 2004). The miRBase registry provides miRNA gene hunters with unique names for novel miRNA genes prior to publication of results (Griffiths-Jones, *et al.,* 2006). The MiRbase database was mined for all known miRNAs associated with PC. Parameters of the search included, but not limited to terms, such as human, prostate-specific, cancer, early stage etc.

The GEO is an international public repository that archives and freely distributes micro-array, next-generation sequencing, and other forms of high-throughput functional genomics data submitted by the research community (Barrett *et al.,* 2013). The repository supports MIAME-compliant data submissions. Tools are provided to help users query and download experiments and curated gene expression profiles. The three main goals of GEO are to provide a robust, versatile database in which to efficiently store high-throughput functional genomic data. In addition, to offer simple submission procedures and formats that support complete and well-annotated data deposits from the research community. Furthermore, to provide user-friendly mechanisms that allow users to query, locate, review and download studies and gene expression profiles of interests (Barrett *et al.,* 2013).

For the purpose of this study micro array datasets were used. These micro-array datasets contained data for human prostate cancer specific miRNAs. Datasets labelled GSE4574 and GSE45604 were retrieved from the GEO database.

Data from both databases were retrieved and a separate list for each database was compiled and streamlined. The miRNA list from GEO was prioritised based on the miRNA fold change expression and a p value of > 0.05. After prioritization 60 miRNAs were identified as being implication in PC from GEO and 50 from miRBase. Both the lists, one from miRBase and the other from GEO were combined and duplicates were eliminated using a Perl script (See Appendix A).

A combined list of 40 known miRNAs was generated that included their accession numbers and links to publications, describing the association of these miRNAs to PC. Data was exported or retrieved and saved in compatible formats for further analysis.

2.2.2 Target gene prediction using the 40 identified prostate miRNAs

The 40 miRNAs identified were used as individual queries to predict the target genes they are known to regulate. Target predictions were performed using the miRDB database available at http://www.mirdb.org/cgi-bin/search.cgi (Wong and Wang, 2015), the Gene Expression Omnibus database available at http://www.ncbi.nlm.nih.gov/geo/ (Barrett *et al.,* 2013) and the target scan database available at http://www.targetscan.org/ (Friedman *et al.,* 2009).

The miRDB is an online database for miRNA target prediction and functional annotations. All the targets were predicted by a bioinformatics tool, MirTarget, which was developed by analyzing thousands of miRNA-target interactions from high-throughput sequencing experiments. Common features associated with miRNA target binding have been identified and used to predict miRNA targets with machine learning methods. miRDB hosts predicted miRNA targets in five species: human, mouse, rat, dog and chicken. As a recent update, users may provide their own sequences for customized target prediction.

The GEO database as described in section 2.2.1, houses an array of information. Micro-array experiments present in the database incorporates information on gene targets that interact with the respective miRNAs of interest.

TargetScan predicts biological targets of miRNAs by searching for the presence of conserved 8mer and 7mer sites that match the seed region of each miRNA (Lewis *et al.,* 2005). As an further option, non-conserved sites can also be predicted. In addition, the tool can also identify sites with mismatches in the seed region that are compensated for by conserved 3' pairing (Friedman *et al.,* 2009). In mammals, predictions are ranked based on the predicted efficacy of targeting as calculated using the context+ scores of the sites (Grimson *et al.,* 2007 and Garcia *et al.,* 2011). As an option, predictions are also ranked by their probability of conserved targeting (P_{CT}) (Friedman *et al.,* 2009). TargetScanHuman considers matches to annotated human UTRs and their orthologs, as defined by UCSC whole-genome alignments. Conserved targeting has also been

detected within open reading frames (ORFs). TargetScan uses the miRNA ID information as a query to extract target gene information from all available experiments publicly available

Three separate databases were used in this study to improve robustness whilst adding value to the predictions that were made. Each miRNA was used as a query to identify specific target genes for each miRNA. Gene targets from all of the databases for each of the miRNAs were collated and saved for further analyses.

2.2.3 Prioritising the miRNA and target gene lists

For the purpose of this research the target gene lists containing 800 or less targets were included in the study. Only gene targets predicted with a confidence level of more than 80% was included for downstream analysis. Each of the separate files generated for the target genes of interest were catalogued using a command line interface. Duplications were eliminated from these files using a Perl script (See Appendix A), with this process being repeated twice to ensure that all duplicates in the final file were removed. A final gene list of 551 genes were identified as being targeted by the 40 identified prostate cancer miRNAs.

2.2.4 Functional Annotation of predicted gene targets using DAVID

Functional annotation of the 551 target genes identified was done using the The Database for Annotation, Visualization and Integrated Discovery (DAVID) version 6.7 (Huang *et al.,*2009a and Huang *et al.,* 2009b). DAVID consists of an integrated biological knowledge-base and analytic tools aimed at systematically extracting biological meaning from large gene/protein lists. DAVID, is a high-throughput and integrated data-mining environment, to analyse gene lists derived from high-throughput genomic experiments.

The initial procedure required uploading of a gene list containing any number of common gene identifiers followed by analysis using one or more text and pathway-mining tools such as gene functional classification, functional annotation chart or clustering and functional annotation table. By following this protocol, investigators are able to gain an in-depth understanding of the biological themes in lists of genes that are enriched in genome-scale studies. The database is available at the following URL (http://david.abcc.ncifcrf.gov/). The 551 gene targets IDs were used as a query in DAVID to identify which functional annotations were assigned to

the identified gene targets. Once the genes were uploaded multiple species were identified as being part of the gene list of interest. Only genes present in the species, *Homo Sapiens,* were subsequently analysed. A total of 497 genes corresponded to *Homo Sapiens*. The genes and their functional annotations were then subjected to K-means clustering a tool available in DAVID (Huang *et al.,* 2009b) where 77 clusters were generated.

Clusters generated were individually investigated to select clusters of genes that are involved in biological process, present in the membrane region and that were involved in pathways that are known to play a critical role in the onset and progression of cancer. From the 77 clusters, 3 clusters were found to be of interest namely cluster 7, 13 and 21. Gene lists representing clusters 7, 13 and 21 were combined and processed, using a Perl script (see Appendix A) to eliminate any duplicated genes within the gene lists, with a final list of 21 genes identified. The identified target genes were then cross-referenced back to the list of 40 identified miRNAs for their involvement in PC. A total of 13 miRNAs involved in PC were identified and used for the subsequent analyses in this study.

2.2.5 Co-expression analyses using the STRING databases

This section, describes the methodologies employed for the co-expression analyses of genes targeted by miRNAs implicated in prostate cancer using publicly available databases. Gene IDs for 21 genes, targeted by the 13 miRNAs implicated in prostate cancer were used as input for the generation of a gene network using the STRING DB version 9. (Snel *et al.,* 2000; Meiring, 2003; Franceschini *et al.,* 2013). STRING an online database (http://string-db.org/), for the identification of protein-protein interactions, focuses on functional protein association. Protein-protein interaction networks are an important ingredient for the systems-level understanding of cellular processes. Such networks can be used for filtering and assessing functional genomics data and for providing an intuitive platform for annotating structural, functional and evolutionary properties of proteins, all of which is made possible by interrogating protein-protein interactions between genes. Exploring the predicted interaction networks can suggest new directions for future experimental research and provide cross-species predictions for efficient interaction mapping.

This tool was used to generate gene networks of the protein-protein

interactions between co-expressed genes to determine if any known interactions exists between the 21 genes and also amongst the 21 genes and important proteins implicated in the progression of PC (Franceschini *et al.,* 2013). The 21 genes were used as driver genes to produce expression networks. To produce each of the expression networks, parameters were judiciously chosen as a follows: (i) a confidence level of 0.7, (ii) a network depth of 4 and (iii) restricting to show only the top 50 interactions between the 21 genes targeted by the identified miRNAs.

2.2.6 Expression profiling for genes targeted by miRNAs

Micro-array data files containing expression data for all 21 target genes were downloaded and extracted from Gene Expression Atlas (GEA) available at (http://www.ebi.ac.uk/gxa/). Files containing data on the expression level of the 21 genes in various cancers were extracted from GEA using a Perl script (see Appendix A). Fold change expression values for tumour vs normal tissue were extracted for each of the cancer types and documented in Excel spread sheets. These files were analysed using an R script (see Appendix A) to produce a graphical representation of the cancer type that indicate the activity of the target genes. Individual graphs for

each target gene were generated, plotting all the cancer types on one graph.

2.2.7 Pathway Analyses using mirPath in DIANA tools

Pathway analysis for this section of the project was achieved using the mirPath (Vlachos *et al.,* 2012) software tool in the DIANA database (DIANA-miRPath v2.0) accessible at http://www.microrna.gr/miRPathv2. New modules enable the DIANA-miRPath server to annotate all the predicted and experimentally validated miRNA targets in a selected molecular pathway (Reverse Search module). DIANA-miRPath is an efficient and yet easy to use tool that can be incorporated successfully into miRNA-related analysis pipelines.

DIANA-miRPath performs miRNA pathway analysis, providing accurate statistics and is able to accommodate advanced pipelines generated by the user. The miRPath tool can utilize predicted miRNA targets provided by the DIANA-microT-CDS (Maragkakis *et al.,* 2009) algorithm and/or experimentally validated miRNA interactions derived from DIANA-TarBase v6.0 (Vergoulis *et al.,* 2012). The tool is capable of combining results with merging and meta-analysis algorithms, performing hierarchical clustering of miRNAs and pathways based on their interaction levels, as

well as being able to generate sophisticated visualization of results, such as dendrograms or miRNA versus pathway heat maps.

From the 40 miRNAs identified further analyses revealed that 13 of the miRNAs were of interest of which 8 (MiR1-8) were tested during subsequent sections of this study and the remaining 5 were tested in a parallel study. This database (DIANA-miRPath) was used to determine the involvement and implication of the 8 selected miRNAs in cancer causing pathways and to further implicate them in the progression of PC.

2.3 Results

2.3.1 Data retrieval

Public databases such as miRBase and GEO were searched for miRNAs implicated in PC. Duplicates were removed using a Perl script (See Appendix A) and a total of 50 and 60 miRNAs were identified from miRBase and GEO respectively. Datasets were combined (110 genes) and duplicates were once again removed to produce a final list of 40 miRNAs implicated in PC. Results from Table 2.1 below indicates the process used to retrieve datasets from known miRNA databases and to identify miRNAs.

Table: 2.1: Tabular representation of miRNA dataset retrieval from 2 public databases

Database name	miRBase	GEO
Number of miRNAs identified	60	455
Number of datasets retrieved	1	2
Number of miRNAs after duplicates are removed	50	60
Datasets from both databases combined	110	
Total combined list of miRNAs after duplicates are removed	40	

2.3.2 Target predictions

The 40 miRNAs identified in section 2.3.1 were used as queries in three publicly available target search databases namely MiRDB, GEO and TargetScan. MiRNAs were individually used as a query to identify specific target genes associated with each miRNA. Table 2.2 indicates the combined total number of target genes identified from each database for the 40 identified miRNAs.

Table 2.2 Tabular representation of the number of target genes identified

Database Name	Initial No.of target genes identified
MiRDB	7063
GEO	14638
TargetScan	21986

2.3.3 Prioritization of target gene lists

Target gene lists for each of the miRNAs from the respective databases were prioritised using the following parameters : (a) only miRNA targeted gene lists generated that contained 800 or less genes were included in the study for further analyses. (b) only gene targets with a prediction score of

80% and more were taken into account. (c) eliminating duplicates. Table 2.3 shows the prioritised number of target genes for each database as well as the final number of target genes identified.

Table 2.3: Prioritization of target genes identified for the 40 miRNAs implicated in PC

Stage of prioritization	Database		
	MiRDB	GEO	TargetScan
Number of target genes after applying criteria (a)	3321	5865	7898
After applying criteria (c)	1805	3065	5923
Number of target genes after applying criteria (b)	803	925	1078
Number of target genes after combining and applying criteria (c)	551		

2.3.4 Functional Annotation using DAVID

The 551 genes identified in section 2.3.3 were subjected to functional annotation and cluster analyses. In total, 77 clusters were generated and the annotation terms associated with the target genes were characterised into 3 categories namely Biological Processes (BP), Molecular Function (MF) and

Cellular Components (CC). Three (7, 13, 21) of the 77 clusters, were found to be of interest and contained a total of 41 target genes. Gene duplicates were eliminated and a total number of 21 target genes (Figure 2.4), were associated with biological processes such as "regulation of Apoptosis", "Programmed cell death" and "regulation of cell death" as depicted in Figure 2.6. For the category Molecular Function (MF) it was shown that more than 40% of the identified target genes were associated with "metal ion binding", "metal binding", "ion binding" and "cation binding" as indicated in Figure 2.7. The target genes were also strongly associated to cellular components such as "membrane enclosed lumen" and intracellular organelle lumen" as depicted in Figure 2.8.

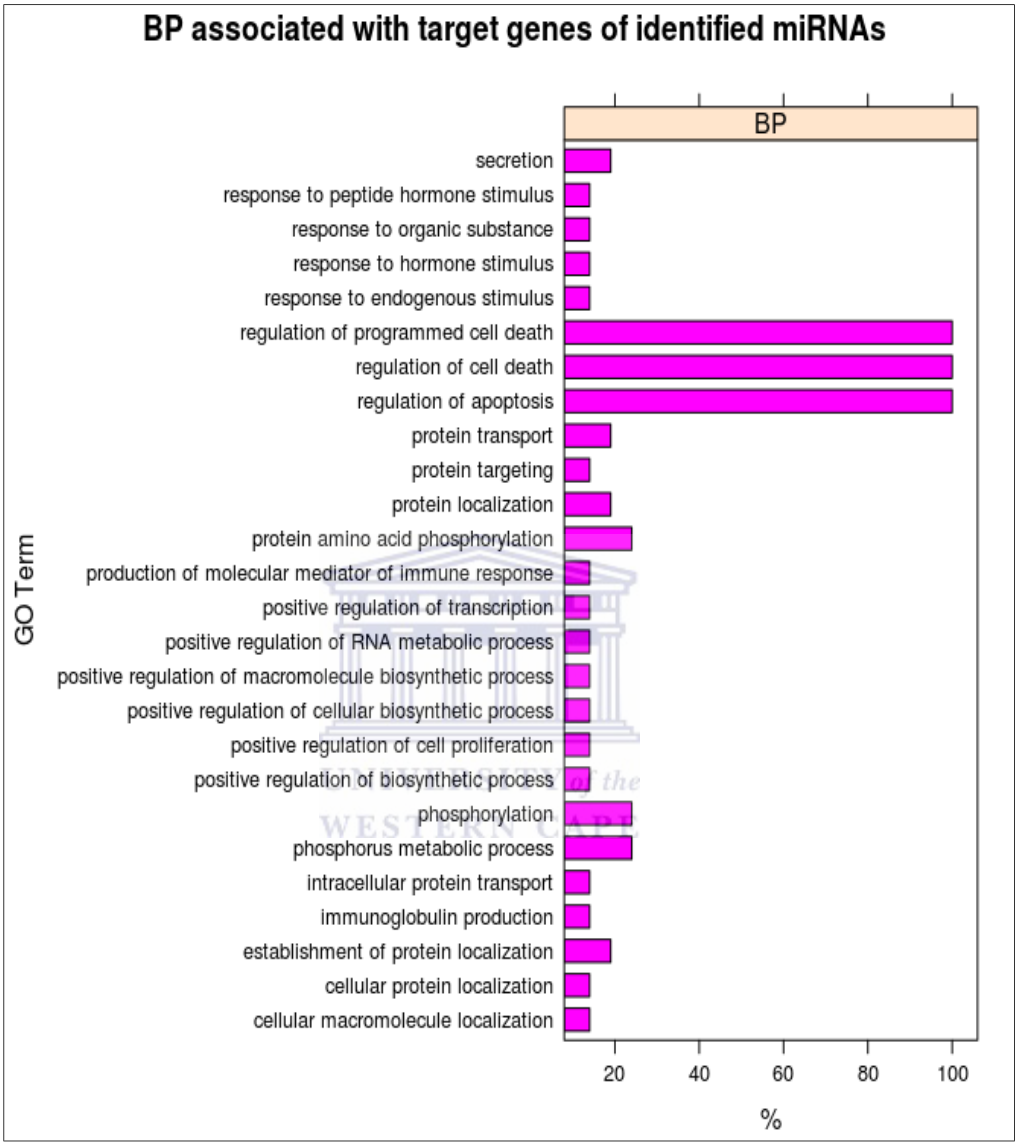


Figure 2.6: Graphical representation of the functional annotation terms under the category biological processes (BP), associated with the target genes for the identified miRNAs. The magenta bars represent the percentage of genes associated with a specific term.

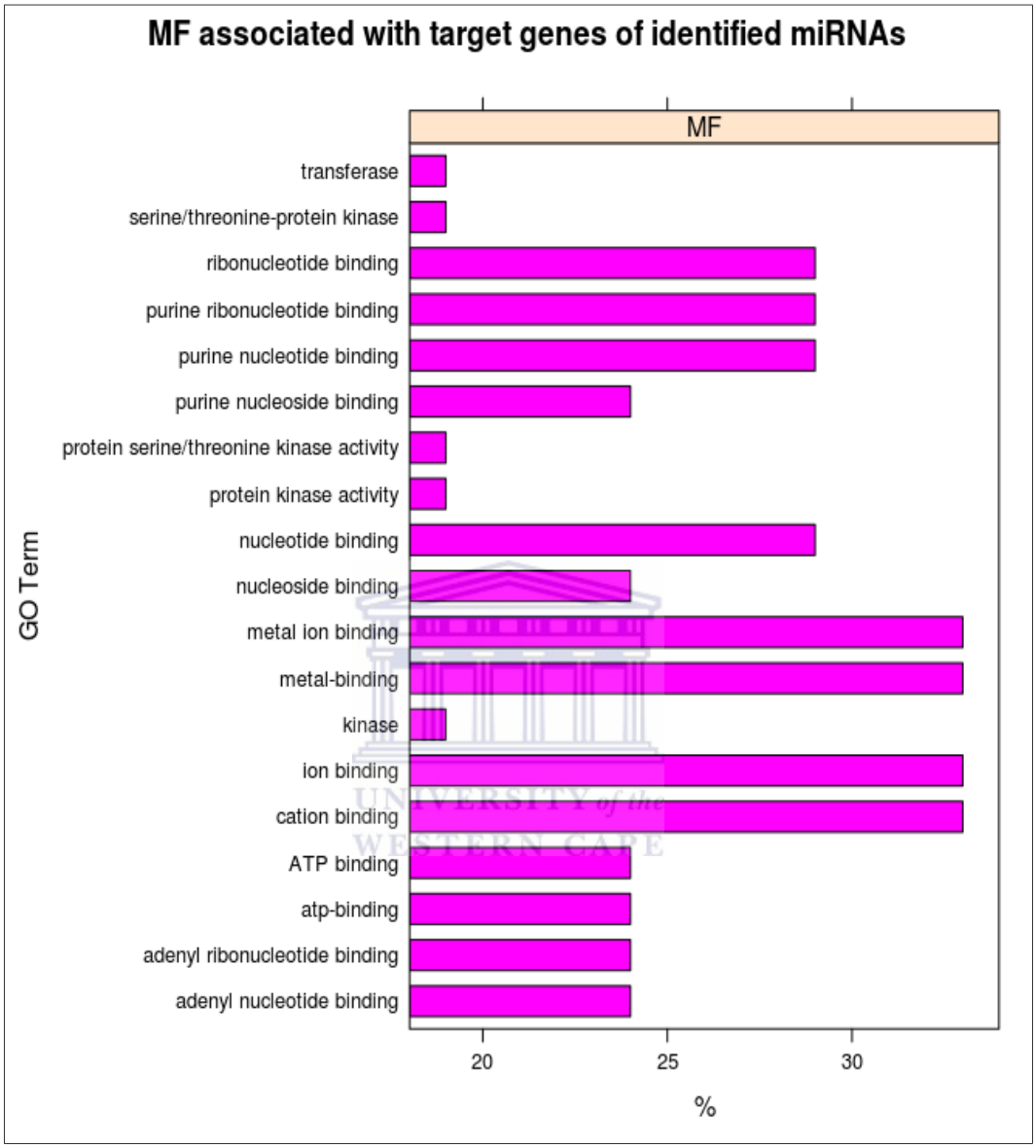


Figure 2.7: Graphical representation of the functional annotation terms under the category molecular function (MF), associated with the target genes for the identified miRNAs. The magenta bars represent the percentage of genes associated with a specific term.

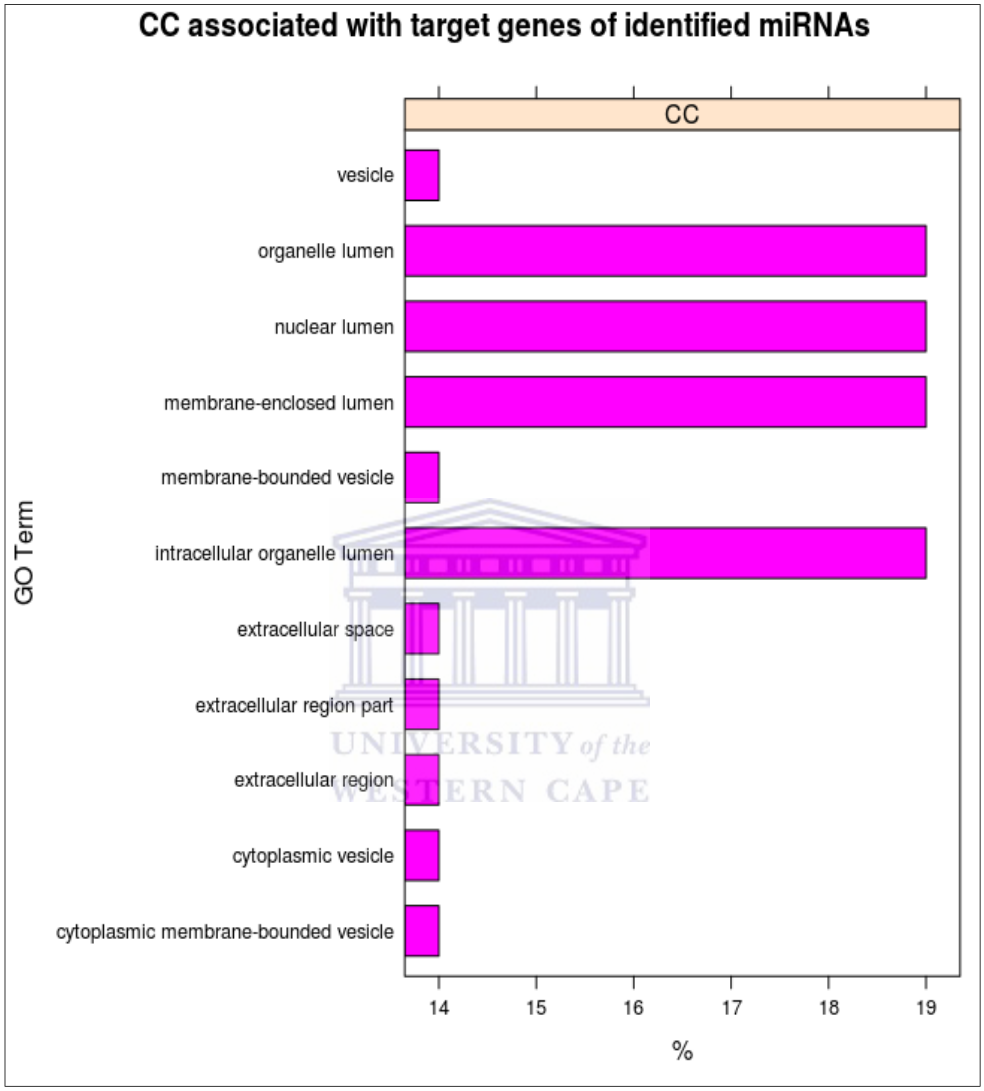


Figure 2.8: Graphical representation of the functional annotation terms under the category cellular components (CC), associated with the target genes for the identified miRNAs. The magenta bars represent the percentage of genes associated with a specific term.

Table 2.4: Tabular representation of the miRNAs and their identified gene specific targets: Column 1 indicates miRNAs, Column 2 represents the confidence level of the prediction Column 3 indicates the gene specific target for each of the potential diagnostic miRNAs and Column 4 represents the description of the targets identified.

MiRNA	Confidence %	Target	Gene Description
MIR9	96	YWHAZ	Tyrosine 3-monooxygenase/ tryptophan 5-monooxygenase activation protein, zeta polypeptide
MIR4	95	PRKCI	Protein kinase C, iota
MIR8	95	CFLAR	CASP8 and FADD-like apoptosis regulator
MIR5	94	ING4	Inhibitor of growth family, member 4
MIR4	93	PTPRC	Protein tyrosine phosphatase, receptor type, C
MIR7	88	CLN8	Ceroid-lipofuscinosis, neuronal 8
MIR3	87	SH3RF1	SH3 domain containing Ring Finger
MIR1	85	ADNP	Activity-dependent neuroprotector homeobox
MIR6	85	ATM	Ataxia telangiectasia mutated
MIR 12	84	TNFSF15	Tumor necrosis factor (ligand) superfamily, member 15
MIR2	84	BTG2	BTG family member 2
MIR9	84	TNFSF13B	Tumour necrosis factor
MIR 10	83	LIG4	Ligase IV, DNA, ATP-dependent
MIR1	82	PAK7	p21 protein (Cdc42/Rac)-activated kinase 7
MIR 11	81	FOXC1	Forkhead box C1
MIR 11	81	MNT	MAX binding protein
MIR1	81	RAB27A	RAB27A, member RAS oncogene family
MIR3	81	BFAR	Bifunctional Apoptosis regulator
MIR8	81	CSRNP3	Cysteine-serine-rich nuclear protein 3
MIR 10	80	TMX1	Theoredoxin-related transmembrane protein
MIR1, MIR6	80	ACVR1C	Activin A receptor, type IC

2.3.5 Co-expression Analysis Using STRING

The 21 genes were used as driver genes to produce expression networks using the parameters as described in section 2.2.5. The results in Figure 2.9 shows a clear link between 12 of the 21 target genes as shown in the blue network present in the overall gene network. Gene targets BFAR and CFLAR are seen to be connected to one another. The remaining 7 genes did not have any connection to one another or the rest of the target genes as seen by the nodes in the upper right quadrant of Figure 2.9.

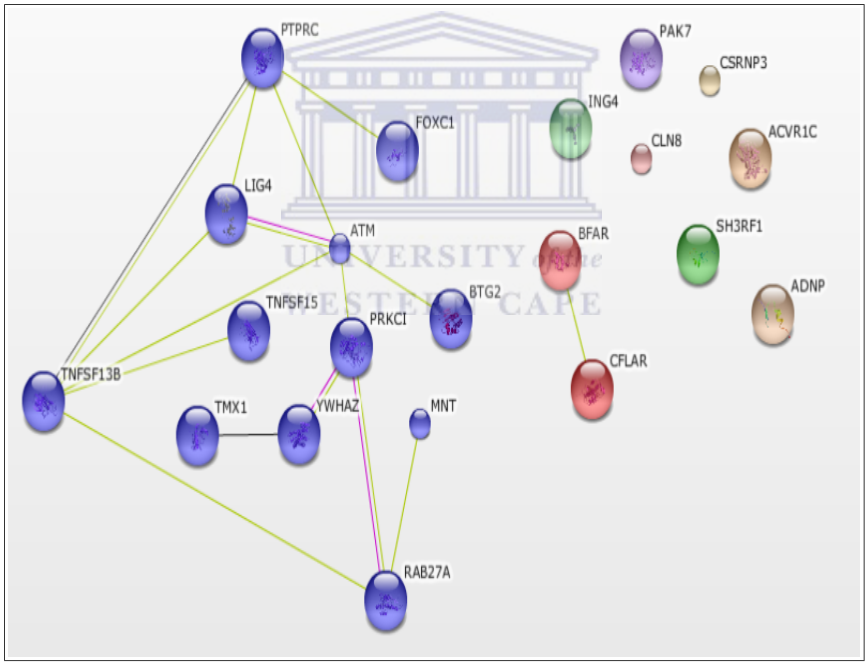


Figure 2.9: Protein Network Visualization generated by STRING. STRING analysis shows the interactions of the 21 miRNA targeted genes clustered together using MCL clustering. The nodes represent genes and the lines joining them represent the evidence available for the connection between genes.

2.3.6 Expression profiling of 21 target genes using GEA

Figure 2.10 indicates the fold change expression of 20 of the target genes in various types of cancer tissue. The fold change was calculated relative to the genes expression value obtained in cancer vs normal tissue. From the graph it is clear that 3 types of cancer showed a higher degree of expression for the target genes. These cancers included breast, liver and prostate cancer. A number of genes showed a significant increase in expression in cancerous tissue compared to normal tissue. In breast cancer tissue some of the target genes (SH3RF1, MNT, PRKCI, RAB27A and LIG 4) showed a 2.5-4 fold change in expression compared to normal breast tissue. Whereas ADNP, ATM, BFAR and BTG2 showed a fold change of between 6 and 8 in breast cancer tissue compared to normal breast tissue (Figure 2.10). In liver cancer two genes (CFLAR and BFAR) showed a fold change expression of 14 and 11 respectively when compared to normal liver tissues. In PC tissue, the expression of 6 target genes (SH3RF1, TMX1, MNT, PRKCI, CFLAR and ING4) were increased as opposed to normal prostate tissue (Figure 2.10). Amongst these six genes whose expression levels were significantly increased in prostate cancer tissues, SH3RF1 and ING4 showed the lowest fold change in expression. MNT had a fold change value of 6, whereas genes TMX1, PRKCI and CFLAR recorded the

highest fold change (7.5) in PC

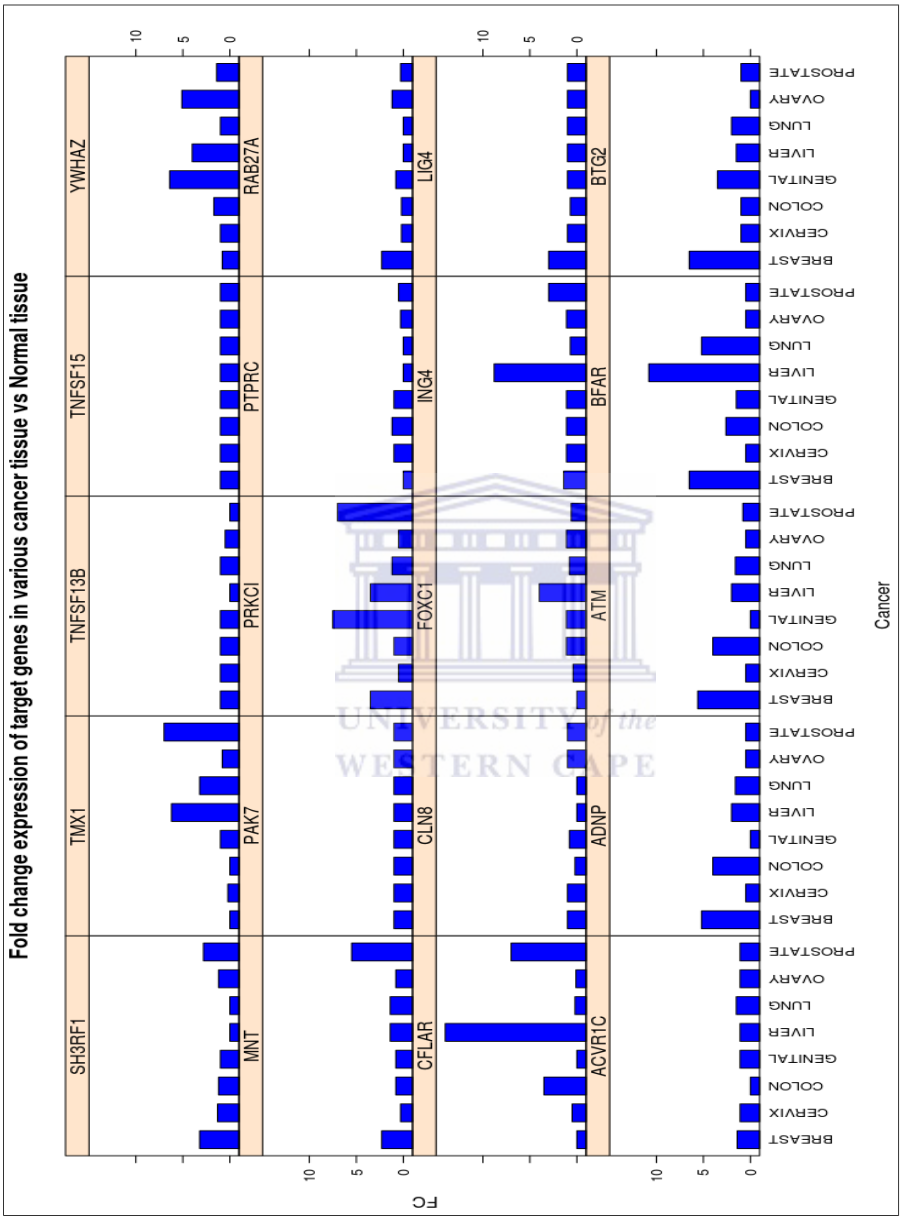


Figure 2.10: Graphical representation of the fold change expression in tumour Vs normal tissue for 20 target genes of the identified miRNAs.

2.3.7 Pathway Analysis using DIANA-mirPath

In silico pathway analysis tools (DIANA-miRPath v2.0, and target database microT-CDS) were utilized to obtain predicted gene targets and associated KEGG (Kyoto Encyclopedia of Genes and Genomes) (Kanehisa and Goto 2000) pathways of MiR1-8. In total, 33 KEGG pathways were identified that are associated with MiR1-8 as shown in Figure 2.11. Pathways were selected based on the following criteria: smallest *P*-values for the union of common pathways targeted by all miRNAs and the largest number of genes represented in those pathways. Some of the pathways associated with miRNA1-8 included pathways in cancer (prostate cancer, colorectal cancer, basal cell carcinoma, small lung cancer), PI3K (phosphoinositide 3-kinase)/Akt signalling pathway, TFG-beta signalling pathway, cell structure-related pathways, (focal adhesion), cell function pathways (ubiquitin proteolysis). A complete list of pathways and their associated p-values is included in Appendix A. Collectively these results lend support to the identified miRNAs potential link to cancer and more specifically PC.

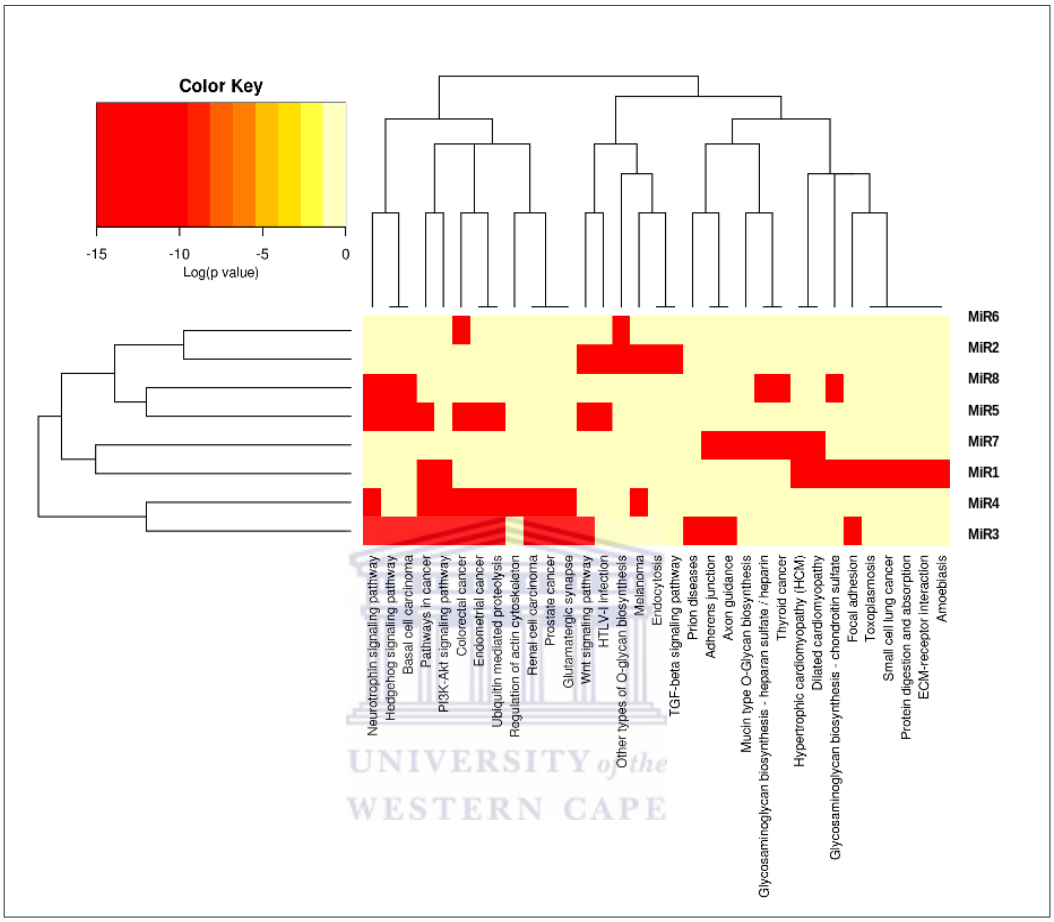


Figure 2.11: Graphical representation of the association of miRNA1-8 and various cancer causing related pathways. The red blocks indicated the strongest association (associations with log *p-values* smaller the -9.5). Orange blocks show intermediate association and yellow blocks (with log *p-values* ranging between 0 and -1.5) show the poorest association between miRNA1-8 and the pathways identified.

2.4 Discussion

MiRNAs have been found to be down-regulated in a number of tumours and in some cases the re-introduction of these miRNAs have been shown to impair the viability of cancer cells (Davis & Hata, 2009; Sotiropoulou *et al.,* 2009; Pritchard *et al.,* 2012). The value of miRNA profiles in tumour diagnostics is well established for example, 16 miRNAs have been shown to be strongly down regulated in breast cancer (Aqeilan *et al.,* 2010). These markers may be used in the development of drug resistant and treatment selection tests. MiRNAs have also been shown to be involved in other processes besides translational gene silencing. Currently, evidence to support this are through studies that show that mature mammalian miRNAs can be imported into the nucleus and secreted from the cell (Garzon *et al.,* 2006). These results suggest that miRNAs may regulate transcription or paracrine signalling.

Unlike siRNAs, miRNAs are endogenous and therefore have the potential to enhance the understanding of the regulation of particular genes. In addition, miRNAs are now seen as an additional layer of gene regulation

which can be dysregulated in disease (Ozen *et al.,* 2008; Clape *et al.,* 2009; Fu *et al.,* 2010; Schaefer *et al.,* 2010).

From the *in-silico* methods employed as described in section 2.2, 40 miRNAs were identified as being previously implicated in prostate cancer as indicated in Table 2.1. A new pipeline was generated using selective terms and parameters to identify miRNAs and their target genes associated with PC. Information from various databases was pooled together to generate a list of potential miRNAs as biomarkers for PC. The retrieved list was further refined using various *in silico* tools which incorporates statistical analysis to prioritise the miRNA list.

This proved that our "in-house" technique was more specific for the identification and the prioritisation of PC-associated miRNAs. Using the 40 miRNAs as queries, 551 target genes were identified for the list of input miRNAs. Bioinformatics analysis which included functional annotation and cluster analyses revealed a prioritised target gene list of 21 genes of interest. These target genes were found to be targeted by 13 of the initial 40 miRNAs identified as depicted in Table 2.4.

The gene targets for these 13 miRNA as indicated in Table 2.4 showed a strong association to known processes involved in apoptosis and cell cycle regulation processes. MiRNA-8 was predicted to have a strong association of 95% with CFLAR (Table 2.4), a gene that is known to be involved in the regulation of Caspase-8 and FADD (Goa *et al.,* 2006), two crucial genes in the regulation of apoptosis.

Studies by Gao and colleagues in 2005 and 2006 demonstrated that the Androgen receptor pathway directly targets the CFLAR gene, this finding resulted in the identification of three functional antioxidant responsive elements (AREs) such as par-4 in the CFLAR promoter (Gao *et al.,* 2005). They have further illustrated that par-4 functions as an AR cofactor for androgen-driven CFLAR gene expression. In order to validate their findings Goa (2006) conducted experiments to observe the activity of a par-4 enhanced androgen-driven CFLAR promoter in a transient transnfection assay (Goa *et al.,* 2006). Chromatin immunoprecipitation analysis indicated that par-4 was capable of directly binding to the promoter region of CFLAR during androgen treatment when using the

LNCaP cell line. These results suggested that par-4 directly targeted CFLAR.

The androgen pathway exerts a protective effect in the prostate gland (Hakimi *et al.,* 1996) and in androgen-sensitive human prostate cancer LNCaP cells (Coffey *et al.,* 2002). Although the mechanisms underlying these effects have not been clearly defined, androgen's effects on both pro-apoptotic and anti-apoptotic gene expression have been demonstrated (Coffey *et al.* 2002). Gao and colleagues (2006) observed that AR and par-4 levels gradually decreased and increased respectively in all epithelial cells following castration, they also found that a few epithelial cells lose CFLAR expression following castration (Gao *et al.*, 2006). Most interestingly and significantly, from this study was the fact that apoptosis was only detected in the epithelial cells that did not express CFLAR. These finding may suggest that predictions made earlier could possibly indicate, that the target gene CFLAR, is indeed involved in the relevant cancer causing pathways as well as PC.

Since a clear link has been established between dysregulation of apoptosis

and cancer, this miRNA would be expected to potentially be involved in the regulation of apoptosis and therefore play a role in cancer development. This miRNA also regulates the gene, CSRNP3, a crucial gene in the regulation of apoptosis and thus a definitive link of the miRNAs to cancer is expected.

In addition to results in Table 2.4, miRNAs 3 and 5 showed association to genes involved in apoptosis and cell growth regulation. The data as summarized in Table 2.4 would suggest that the targets of the miRNAs selected in this study have strong association with genes that are associated with crucial processes in cell growth, cell death regulation and therefore potentially cancer development.

Functional annotation analyses showed that the targets of these selected miRNA could be implicated in processes linked to cancer as shown in the Gene Ontology (GO) analysis performed in this study. The target genes were found to be highly associated to biological processes such as “regulation of transcription”, “regulation of apoptosis” and programmed cell-death as depicted in Figure 2.6. With respect to biological processes

identified for these selected targets, the process of apoptosis is crucial in cancer, as dysregulation of this process is known to be involved in the development of cancer (Ozen *et al.,* 2008; Clape *et al.,* 2009; Fu *et al.,* 2010; Schaefer *et al.,* 2010). All 21 gene targets identified for the 13 miRNAs were associated with these processes as depicted in Figure 2.6.

Furthermore, Figure 2.7 indicates that more than 40% of the target genes were found to be associated with Molecular functions such as “metal ion binding”, “ion binding”, “metal binding” and “cation binding” this finding was significant as many diseases including cancer are associated with metal ion imbalance (Williams, 2014). A metal ion imbalance may be regarded as a deficiency of the essential metals, overload of either essential or non-essential metals or perturbation of the overall natural balance. Results as depicted in Figure 2.8 shows that the targets of the miRNAs were also strongly associated with membrane bound proteins. This finding was significant as molecules that are membrane bound are easily excreted into bodily fluids such as blood, saliva and urine which makes them good targets for diagnostic purposes and drug targeting.

STRING analysis as described in section 2.3 and illustrated in Figure 2.9 showed links based on experimental evidence between some of the target genes associated with the miRNAs. However, a number of the 13 miRNA targets (7 miRNA targets) showed no association with any other known protein networks. The reason for this could be that the database was in the process of being updated or that these genes have not previously been linked to cancer causing pathways and specifically PC. The targets that showed strong links with one another in a biological network were all associated with apoptosis regulation and cell growth. Two of the genes, LIG4 and ATM, showed a strong association with more than one of the targets of the miRNAs (Figure 2.9).

Five of the target genes, PTPRC, LIG4, ATM, TNFSF15 and RAB27A showed a direct association and two target genes YWHAZ and PRKCI showed an indirect association with the tumour necrosis factor (TNFSF13B) as depicted in Figure 2.9 in the blue network cluster. This suggests that these targets could be potential regulators of tumour growth. A recent study by Ishiguro and colleagues (2009) using the prostate cancer cell line DU145, revealed that PRKCI is involved in prostate cancer growth

both *in vivo* and *in vitro* (Ishiguro *et al.,* 2009). Furthermore, depletion of PRKCI in DU145 cells suppressed NFκB and AP-1 activities, transcription and secretion of IL-6, as well as suppressing cell growth, but not IL-6 signalling. The group concluded that enhanced PRKCI expression in prostate cancer cells results in overproduction and secretion of a prostate growth factor, IL-6, at the transcription level (Ishiguro *et al.,* 2009). The study further reported that this forms an autocrine loop contributing to the growth of prostate cancer (Ishiguro *et al.,* 2009, Nguyen *et al.,* 2014). The PRKCI-dependent expression of IL-6 mRNA is also observed for another androgen-independent prostate cell line, PC-3, suggesting the generality of this regulation (Ishiguro *et al.,* 2009, Nguyen *et al.,* 2014). The specific over-expression of PRKCI in epithelial cells but not in stromal cells of the prostate further supports such an autocrine mechanism. It is known that most of androgen-independent prostate cancer tissues over-express and/or express mutated AR that is still activated by lower concentrations of androgen, estrogen and anti-androgen drugs (Ishiguro *et al.,* 2009, Nguyen *et al.,* 2014).

Taken together, the pathway might cooperate with the deregulated AR

system to regulate proliferation of hormone-independent prostate cancer cells. These results allow us to confidently establish a link between the target genes and the miRNAs identified in this chapter and by implications to pathways involved in PC. This evidence supports our hypothesis that these miRNAs might have a role in the regulation of the development of tumour cells, and thus possible implications in the progression of PC such that they are potential biomarkers for PC diagnosis.

Furthermore, experimental evidence looking at the ubiquitously expressed YWHAZ protein (a target gene of MiR9) with a 96% confidence as depicted in Table 2.4. Recent research suggests that YWHAZ may play a central role in regulating multiple pathways responsible for cancer initiation and progression.

A study by Murata *et al.,* 2012 showed that the expression of YWHAZ was up-regulated in LNCaP cells treated with androgen. Murata and colleagues (2012) also found that this gene provided favourable conditions for the survival of prostate cancer as it contributed readily to cell proliferation, motility and acquired resistance to etoposide-induced apoptosis as a result

of its ectopic expression in LNCaP cells (Murata *et al.,* 2012; Rüenauver *et al.,* 2014). Murata (2012) and colleagues further investigated the clincopathological influence YWHAZ had on PC (Murata *et al.,* 2012; Rüenauver *et al.,* 2014). This study revealed that YWHAZ had an increased level of expression during the later stages of PC including lymph-node metastasis (Murata *et al.,* 2012; Rüenauver *et al.,* 2014). YWHAZ was described as an androgen-regulated tumour-promoting factor in PC cells by this study (Murata *et al.,* 2012; Rüenauver *et al.,* 2014). These finding strongly suggest a link between the target gene YWHAZ and miRNAs identified lending more credibility to the miRNAs as biomarkers for PC.

No direct links to known regulators of cancer were identified for ING4, PAK7, CLN8, CSRNP3, ACVR1C, ADNP using STRING. However, in a recent study by Berger and colleagues (2014) they identified the chromatin remodelling protein, ING4, as a crucial switch downstream of Myc and Pten that is required for human prostate epithelial differentiation. Myc-induced transient expression of ING4, is required for the differentiation of basal epithelial cells into luminal cells, while sustained ING4 expression induces apoptosis (Berger *et al.,* 2014). The study showed that ING4 expression is lost in >60% of human primary prostate tumours.

They further established that by blocking ING4 expression, Pten loss, prevents differentiation and this phenomena is then rescued by ING4 re-expression (Berger *et al.,* 2014). Pten or ING4 loss generates tumour cells that co-express basal and luminal markers, indicating prostate oncogenesis occurs through disruption of an intermediate step in the prostate epithelial differentiation program (Berger *et al.,* 2014). This finding led to the identification of a new epithelial cell differentiation switch involving Myc, Pten, and ING4, which when disrupted leads to prostate tumorigenesis (Berger *et al.,* 2014). There has been no previous studies indicating that Myc over-expression and Pten loss are common genetic abnormalities in prostate cancer. The study by Berger (2014) is the first demonstration that transient ING4 expression is absolutely required for epithelial differentiation, its expression is dependent on Myc and Pten, and it is lost in the majority of human prostate cancers.

The study by Berger and colleagues in 2014 is the first demonstration that the loss of function of ING4, either directly or indirectly through loss of Pten function, promotes Myc-driven oncogenesis by deregulating

differentiation (Berger *et al.,* 2014). The results from Berger (2014) is thus the first evidence that links the function of the identified target genes in relation to PC and by implication, supports the notion that the miRNAs that regulate the expression of these target genes are vital in the proliferation of PC. Based on the ontology data retrieved from the target search for these genes, they showed strong association to apoptosis regulation and cell growth. Therefore, they were selected as potential targets of the identified miRNAs as shown from the analysed data in Table 2.4 and Figure 2.6, 2.7 and 2.8.

From the expression profiling results as described in section 2.2.6 and as depicted in Figure 2.10, of the 21 genes targeted by the selected 13 miRNAs, it is clear that there is a link between the expression of the target genes and the progression of cancer. Furthermore, there is clear evidence that these target genes are further implicated in PC as seen in the level of fold change expression in PC tissue as depicted in Figure 2.10. As a result of these genes being targeted by the 13 miRNAs of interest we could deduce that these miRNAs could play a potential role in the progression of disease and specifically in the progression of PC.

Pathway Analyses performed using DIANA as described in section 2.2.7 and presented in Figure 2.11 represents a heat map of the 8 unique miRNAs that were not associated with PC in the public databases. Five miRNAs (MiR1, 3, 4, 5 and 8) showed a very strong association with PC based on log scores retrieved from the database, whereas MiR2, 6 and 7 showed a slightly weaker association with PC. Based on the data shown in Figure 2.11 the five miRNAs that showed strong association with PC, also have predicted associations with other cancers. These findings in combination with the expression profiling, functional annotation and cluster analysis results further implicates the identified miRNAs in cancer causing and PC pathways.

CHAPTER 3

Molecular validation of identified miRNAs as biomarkers for the early detection of Prostate Cancer (PC) using qRT-PCR

3.1 Introduction

3.2 Methods

3.2 Molecular Techniques

3.2.1 Cell Culture

3.2.2 Preparation of sample material

3.2.3 miRNA isolation technique: two column protocol

3.2.4 Primer Design

3.2.5 Reverse transcription of miRNA to cDNA using stemloop sequence specific primers

3.2.6 Validation of miRNA expression levels in cancer and control cell lines using PCR and qRT-PCR.

3.3 Results

3.3.1 Determining the PCR efficiency for miRNAs 1-8 using a standard curve:

3.3.2 Calculating the PCR Efficiency

3.3.3 Melting curve analysis

3.3.4 Expression Profiling

3.4 Discussion

CHAPTER 3

Molecular validation of identified miRNAs as biomarkers for the early detection of Prostate Cancer (PC) using qRT-PCR

3.1 Introduction

MicroRNAs (miRNAs) are short strands of RNA that act as post-transcriptional regulators. Current research on miRNA expression in various disease processes from cancer to cardiovascular disease indicates that miRNAs play a vital role in disease pathogenesis and have potential as biomarkers and therapeutic agents (Buckingham, 2003; Nilsen, 2007; Zhao and Srivastava, 2007; Tanzer *et al.,* 2008; Sotiropoulou *et al.,* 2009; Sato *et al.,* 2011; Martinez-Sanchez and Murphy, 2013).

Routine diagnostic methods for the early detection of PC include digital rectal examination (DREs) and prostate-specific antigen (PSA) testing (American Cancer Society, 2012; CANSA, 2013). The PSA test is non-specific, due to the fact that elevated PSA levels measured in benign prostatic hyperplasia (BPH), infection, and/or chronic inflammation, may lead to confounding outcomes. Other blood based biomarkers such as human glandular kallikrein 2 (hK2), urokinase plasminogen activator (uPA)

and its receptor (uPAR), transforming growth factor-beta 1 (TGF- β1); interleukin-6 (IL-6) and its receptor (IL-6R) have been studied alone or in combination with PSA and suggested for diagnosis, staging, prognostics, (Ishiguro *et al.,* 2009, Nguyen *et al.,* 2014). and monitoring of prostate cancer. However, since we are studying a disease that has a highly heterogeneous nature, there is a an urgent need to identify additional biomarkers for better prediction of disease progression and prognosis to aid in clinical decision making with respect to treatment options.

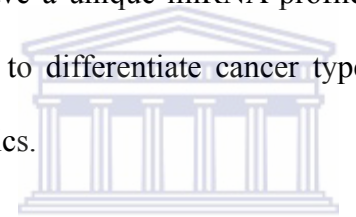
MicroRNAs (miRNAs) are small (18 to 25 nucleotides), highly conserved, non-coding RNA molecules that regulate gene expression post-transcriptionally (Buckingham, 2003; Nilsen, 2007; 2007; Tanzer *et al.,* 2008; Sotiropoulou *et al.,* 2009; Sato *et al.,* 2011). Computational analysis performed on miRNA expression profiles suggests that more than 60% of gene transcripts in mammals are potentially regulated by miRNA either directly or indirectly (Zhao and Srivastava, 2007; Martinez-Sanchez and Murphy, 2013). miRNAs play key regulatory roles in cell cycle control, cell proliferation and differentiation, and the regulation of apoptosis (Buckingham, 2003; Nilsen, 2007; Zhao and Srivastava, 2007; Tanzer *et al.,* 2008; Sotiropoulou *et al.,* 2009; Sato *et al.,* 2011; Martinez-Sanchez

and Murphy, 2013). Recent studies have implicated various miRNAs in the development and progression of multiple human cancers and also as potential biomarker in cancer diagnosis and prognosis (Li *et al.,* 2009).

In the present study, we utilized qRT-PCR for the expression analysis of miRNAs identified in chapter 2, which offers many advantages for miRNAs with known sequences. Quantitative real-time PCR is considered the gold standard for miRNA expression profiling due to its unparalleled sensitivity and specificity (van Rooij, 2011; Dedeoğlu, 2014; Stokowy *et al.,* 2014; Usó *et al.,* 2014). Some of the main advantages of using this methods includes (a) high sensitivity to the single cell level, (b) high specificity to the disease of interest, (c) it does not require a large volume of RNA (Usó *et al.,* 2014) and (d) that it is a high throughput technique. Additionally, this method is less time-consuming compared to microarrays and NGS and the results do not require to be analysed/processed by bio-statisticians (Usó *et al.,* 2014).

Expression profiling of the identified miRNAs was performed in 3 prostate cell lines a Benign prostatic hyperplasia (BPH1), metastatic cancer (LNCaP) and 1 corresponding normal control (PNT2C2) cell line to study

the expression of miRNAs associated with the detection of PC. The identified miRNAs were further validated in a larger sample set consisting of an additional 6 non PC cancer cell lines and two adjacent normal control cell lines. Our main aim therefore was to establish which of the *in silico* identified miRNAs were specific for PC, and secondly whether a definitive miRNA profile for PC could be established to distinguish it from the rest of the cancer cell lines using qRT-PCR. The expectation was that each cancer cell line will have a unique miRNA profile, and that these unique profiles could be used to differentiate cancer types from each other and potentially for diagnostics.



The main aim of this section of the study is thus to molecularly validate the identified miRNAs (in Chapter 2) as potential biomarkers for the early detection of PC using qRT-PCR.

3.2 Molecular Methods:

For the molecular validation part of this study 11 human cell lines were cultured and tested in Phase II of the study.

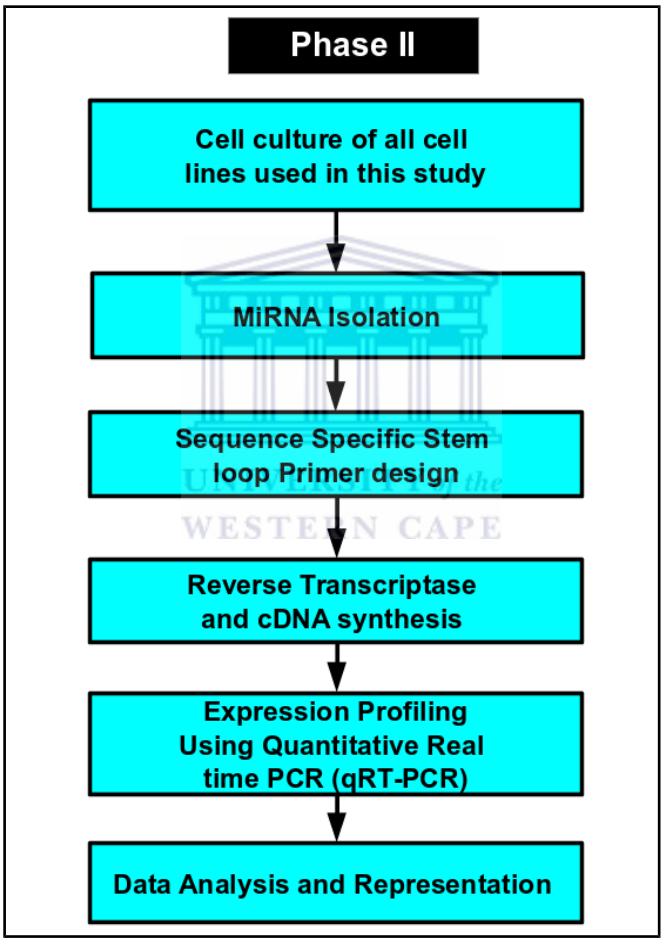


Figure: 3.1 Flow chart representing the methodologies employed in Phase II of the study.

Table 3.1: Tabular representation of the cell lines used to investigate the specificity and sensitivity of miRNA 1-8 in cancer. Column 1 represents the ATCC cell line name. Column 2, 3 and 4 gives the description and status of each cell line at the point of immortalization, respectively. Column 5 and 6 indicates the media and supplement requirement for optimal growth.

Name	Description	Status	Cell type	Media used	Supplement
BPH1	Benign Prostatic Hyperplasia	Benign	Prostate adherent	RPMI 1640	FBS, Penstrep
LNCaP	AR-sensitive human prostate adenocarcinoma	Malignant	Prostate suspension	RPMI 1640	FBS, Penstrep
PNT2C2	Human prostate epithelial	Normal	Prostate adherent	RPMI 164 0	FBS, Penstrep
MCF7	Michigan cancer foundation 7	Cancer	Breast adherent	DMEM	FBS, Penstrep
MCF12 A	Michigan cancer foundation 12 epithelial	Normal	Breast mammary gland adherent	DMEM-F12	FBS, Penstrep, Insulin, hydro-cortisone
HePG2	Liver Carcinoma	Cancer	Liver, adherent	DMEM	FBS, Penstrep
A549	Adenocarcinoma human alveolar basal epithelial cells	Cancer	Lung, adherent	DMEM	FBS, Penstrep
Caski	Cervical squamous carcinoma	Cancer	Cervical adherent	RPMI 1640	FBS, Penstrep
CaOV3	Adenocarcinoma primary ovarian	Cancer	Ovarian adherent	DMEM	FBS, Penstrep
HT29	Colorectal adenocarcinoma	Cancer	Colon, adherent	DMEM	FBS, Penstrep
KMST	Embryonic Fibroblast	Normal	Skin, adherent	DMEM	FBS, Penstrep

Table 3.1 above represent the cell lines used in this study. The panel was selected to contain 3 prostate cell lines two of which were classified as malignant PC and benign Prostate (LNCaP and BPH1 respectively). The third cell line PNT2C2 served as the control cell line for prostate cancer. These 3 cell lines were selected to represent different types of prostate tissue including PC as well as to determine whether miRNA 1-8 were differentially expressed in PC.

The other cell lines indicated in the table represented different cancers. These cell lines were selected to evaluate the differential expression of the eight miRNAs. The main purpose of this was to establish whether the selected miRNA was specific to prostate cancer or if it was expressed in other cancers.

3.2.1 Cell Culture

- Thawing frozen cells for tissue culture

The following protocol describes a general procedure for thawing cryo preserved cells. For detailed protocols, refer to the cell-specific product insert. Cryovials containing the liquid nitrogen frozen cells were removed

from storage and placed in a 37 °C water bath. Cells were thawed rapidly for less than 1 minute until only a small piece of ice remained in the tube. The tubes were transferred to a laminar flow hood and wiped down with ethanol at a concentration of 70% before opening the vials. A volume of 1 ml of pre-warmed complete media was added drop wise to the cryovial containing the thawed cells. The cell suspension was centrifuged at approximately 200 × g for 5–10 minutes using a Sorval H400 TC6 centrifuge (American Instrument Exchange, Inc.). The actual centrifugation speed and duration varies depending on the cell type. Following centrifugation the clarity of the suspension and the visibility of the pellet was checked. Thereafter the supernatant was aseptically decanted without disturbing the pellet that was formed following centrifugation. The pellet was then re-suspended in complete growth media and transferred to a T25 cell culture flasks and incubated at 37 °C in 5% CO_2 for 24 -72 hours depending on the cell line

- Counting cells using a haemocytometer

The procedure below provides a general description on the use of the hemacytometer. The chamber and cover slip was cleaned with alcohol, thereafter it was dried and fixed in position. Cells were then harvested and

10 µL of cells were added to the cover slip. The chamber was then placed under an inverted light microscope (Nikon TMS-F) with a Leica EC3 digital Camera (Leica Microsystems Ltd) attached and visualised using the 10X magnification setting and phase contrast to distinguish the cells. Cells were then counted in the large, central gridded square (1 mm^2). In order to estimate the amount of cells per mL the counted cells were then multiplied by 10^7. Duplicates were prepared in order to average the cell count

- Cryopreservation of mammalian cells

Mammalian cells are cryopreserved to avoid loss by contamination, to minimize genetic change in continuous growth of cell lines, and to avoid ageing and transformation in finite cell lines. Before cryopreservation, cells were characterized and checked for contamination through visual inspection.

Cryopreservation media generally consists of a base medium, cryopreservative, and a protein source. The cryopreservative and protein protect the cells from the stress of the freeze-thaw process. A serum-free medium has generally low or no protein; but can still be used as a base for a cryopreservative medium. Cryopreservation was carried out on suspension as well as adherent cultures.

- Suspension cultures

For suspension cultures the number of viable cells to be cryopreserved was counted. Cells should be in the log phase of the growth cycle. The cells were centrifuged at ~200 to 400 × g for 5 min to pellet the cells. Thereafter a pipette was used to remove the supernatant down to the smallest volume without disturbing the pelleted cells. Cells were then re-suspended in preservation media to a cell count of 1 x 10^7 to 5 x 10^7 cells/ml for serum containing medium, or 0.5 x 10^7 to 1 x 10^7 cells/ml for serum-free medium. Cells were then aliquoted into cryogenic storage vials and placed on ice in order to start the freezing process. Cells were then frozen slowly at -1 °C /min, this was achieved by placing the vials into an insulated box and placing it in a -70 °C to -90 °C freezer.

- Adherent cultures

Cells were detached from the substrate with a dissociation agent namely trypsin, this procedure was done as gently as possible to avoid irreparable damage to cells. Trypsinisation is the process of *cell* dissociation using trypsin, a proteolytic enzyme which breaks down proteins, to dissociate

adherent cells from the vessel in which they are being cultured. When added to a cell culture, trypsin breaks down the proteins which enable the cells to adhere to the vessel. Cells were then re-suspended in complete growth medium and a viable cell count was established. Cells were then centrifuged at ~200 g for 5 min to pellet cells. Using a pipette, the supernatant was removed down to the smallest volume without disturbing the pelleted cells. Cells were then re-suspended in preservation media to a concentration of 5 x 10^6 to 1 x 10^7 cells/ml and then aliquoted into cryogenic storage vials. Vials were then placed on ice in order to start the freezing process. Cells were then frozen slowly at -1°C /min, this was achieved by placing the vials into an insulated box and placing it in a -70°C to -90°C freezer.

3.2.2 Preparation of sample material

Preparation of cells for RNA extraction was done as follows. Cultured animal cells were harvested at a cell count of 10^6 cells. Cells were transferred to a centrifuge tube and transported on ice to the extraction laboratory. Cells were centrifuged at a low centrifugal force of 200 × g for 10 min at 15-25 °C. Thereafter, the supernatant was removed carefully to not disturb the pellet that was formed. The cells were then lysed by adding

150 µL of a 20 % binding buffer solution and homogenised by pipetting the suspension up and down until the solution become clear. The lysate was then passed through a 20-gauge needle of 0.9 mm thickness which was attached to a sterile syringe. The homogenate was then centrifuged for 2 min at maximum speed using a bench top centrifuge to ensure that all cells were homogenised and the supernatant was transferred to a sterile tube for step one of the RNA isolation process. Similar procedures were used for sample preparation of suspension and adherent cultures.

All the cell lines were grown under specific conditions as described above in section 3.2.1. Briefly, cell lines were cultured under general tissue culture conditions at 37°C. in 5% carbon dioxide for 24-96 hours depending on cell type. Cells were cultured until confluency was reached. Once cells reached confluency cells were divided, sub-cultured or cryopreserved as described above in section 3.2.1. The remainder of the cells were used for further experimentation. Cells in the remaining flasks were prepared for miRNA extraction as described in section 3.2.3. A cell count was done using a haemocytometer as described in section 3.2.1 Flasks containing no more than 1000000 cells were used for miRNA

extraction

3.2.3 miRNA isolation technique: two column protocol

For the two column protocol 150 µL of the cell lysate was transferred to a sterile micro-centrifuge tube. A volume of 312 µL of the binding buffer was added to the lysate and the tube was vortexed briefly thrice at 5 s intervals. A high pure filter and collection tube was combined and the homogenate and binding buffer was placed into the upper reservoir of the filter. The combined tubes were then centrifuged for 30 s at 13000 × g and the flow-through collected. Thereafter 200µL of binding enhancer was added to the flow-through. A new high pure filter and collection tube was assembled and the mixture was transferred to the filter tube in the upper reservoir. The tubes were then centrifuged for 30s at 13000 × *g* and the flow-through was discarded. Thereafter 500 µL of wash buffer was added to the upper reservoir and the tubes were centrifuged for 30s at 13000 × *g* and again the flow-through was discarded. Following this step, 300 µL of wash buffer was added and once again the tubes were centrifuged for 30 s at 13000 × *g* and the flow-through discarded. Once the filter is optimally washed the filter was dried by centrifuging for 1 min at *13000* × *g*. The

filter was then transferred to a clean 1.5 ml microcentrifuge tube and 100 µL of Elution buffer was passed through the filter. The filter and elution buffer was subsequently incubated for 1 min at 15-25 °C. Thereafter the tube and filter was centrifuged for 1 min at 13000 × *g*. The microcentrifuge tube containing the eluted miRNA was used in upstream molecular techniques and was stored at -80 °C. After completion of the miRNA extraction the quality and concentration of the miRNAs extracted was measured using the Qbit assay (according to the manufacturer's instructions), as the miRNAs could not be visualized by normal gel electrophoresis due to its small size.

3.2.4 Primer Design

Parameters for Primer design of miRNAs for RT-PCR were obtained using the Cotton estate Database which is available at the following URL: http://www.leonxie.com/miRNAprimerDesigner.php#. MiRNA sequences were used as the input for this primer design database. The database identified primers specific for the sequences uploaded by using a built in algorithm that scans available public data that have been experimentally validated. The miRNA primer designer has been developed for favouring designing of miRNA primers and qRT-PCR primers for miRNA expression

profile assay. The following parameters were used to design primers for the identified miRNAs (Table 3.2). All primers were designed to have an average Tm of 60 ºC.

Table 3.2 : Tabular representation of the parameters used for *in silico* primer design

Reverse primer:	**ATCCAGTGCAGGGTCCGAGG**
RT adaptor	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGA
Forward adaptor:	GCGGCGG
Length of RT primer	6
Optimum length of forward primer:	13
Minimum length of forward; primer:	15
Maximum length of forward primer:	17

3.2.5 Reverse transcription of miRNA to cDNA using stemloop sequence specific primers

For cDNA synthesis the protocol used was according to the manufacturer's instruction, with the only alteration being the use of specific stem-loop primers as opposed to the oligo-dT and random hexamers provided by the manufactures (Transcriptor First Strand cDNA synthesis kit, Roche Life

Science). In a total volume of 20 µL, 1 µg miRNA was mixed with 2.5 µM stemloop primer, 1 x Transcriptor Reverse Transcriptase Reaction Buffer (containing 8 mM $MgCl_2$), 20 U Protector RNase Inhibitor, 4 mM Deoxynucleotide Mix and 10 U Transcriptor Reverse Transcriptase. The reaction was incubated at 55 °C for 30 min followed by a final inactivation step of 5 min incubated at 85 °C. The concentration of the synthesised cDNA was determined with a NanoDrop Spectrophotometer ND1000.

3.2.6 Validation of miRNA expression levels in cancer and control cell lines using PCR and qRT-PCR.

- End-point PCR protocol

For the end point protocol a PCR master mix was prepared by scaling down the volumes to 20 µL. In order to overcome pipetting errors, a 10% coverage was included. Internal controls were prepared by adding nuclease free water in place of the cDNA products. Nuclease free water, PCR buffer with a final concentration of 1X, dNTP mix with a final concentration of 10 mM, forward and reverse primers with a final concentration of 1µM each and Taq Polymerase 2 (1X), was added to a microcentrifuge tube to make up the master mix. aliquotes of 19 µL of the master mix and 1 µL of cDNA with a final concentration of 250 ng/µL was added to individual PCR tubes.

Reactions were then placed in a preheated (94°C) thermal cycler heating block and incubated at 94°C for 2 min, followed by 20–40 cycles of 94°C for 15s and 60°C for 1 min and 72ºC for 10 min. The reaction products were then analysed by electrophoresis on a 2% agarose gel in 1× TAE buffer.

- miRNA SYBR Green I assay protocol for qRT-PCR

A 5× LightCycler FastStart SYBR Green I master mix was prepared (Kappa Sybr fast Master Mix) according to the manufacturers instructions. In all qRT-PCR reactions, a standard reaction mix was prepared containing KAPA SYBR FAST® Optimized for LightCycler®480 (KAPA Biosystems). miRNA-specific stem-loop primers were designed as described in section 3.2.4 and optimised with a standard end-point PCR using a T_A of 60°C for all primers. PCR master mixes were prepared by scaling the reaction volumes to 20 µL. In order to overcome pippetting errors a 10% coverage was included. Internal controls were prepared by adding nuclease free water in place of the cDNA products. Nuclease free water, Kappa Master Mix and 1µM of both the forward and reverse primers were added to a microcentrifuge tube to make up the master mix. Master mixes were prepared for each primer set that was to be tested. A 96 well Real Time

PCR plate was placed on ice and used for the RT-PCR reactions. Aliquotes of 16 µL master mix was pipetted into each well of the 96 well plate, 4 µL of the respective cDNAs was added to each of the wells. The plate was then covered with a clear foil seal and centrifuged to bring all the solutions to the bottom of the wells of the plates . The plate was then transferred to the Light Cycler 480 instrument and incubated at 95°C for 5 min, followed by 45–50 cycles of 95°C for 15 s, 60°C for 15 s, and 72°C for 10 s. For melting curve analysis, the samples were denatured at 95°C, then cooled to 65°C at 20°C per second. Fluorescence signals were collected at a 530 nm wavelength continuously from 65°C to 95°C at 0.2°C per second. The results were analysed using the REST software package (Pfaffl *et al.,* 2002).

- Standard Curves

A standard curves for each primer set was prepared by using a dilution series of the experimental sample expected to have the highest expression. Alternatively, standard curves could be prepared from a dilution series of an appropriate RNA oligonucleotide, e.g. when detecting artificial miRNAs or siRNAs.

- Relative Quantification

The relative standard curve quantification method was employed to quantify the expression levels of miRNA in all the cell lines. This method relies on the use of both a target and reference gene(s) to be measured in a serial dilution format, in order to construct a standard curve which will be used to calculate the relative expression of each miRNA. Standard curves were generated using a dilution series in the concentration range 500 ng-0.05 ng. The data generated was analysed using the publicly available Microsoft® Excel®-based software REST© (Pfaffl *et al.,* 2002).

Pfaffl developed a mathematical formula widely used for the relative quantification of gene expression in Real-time PCR (Pfaffl, 2001a). This model combines gene quantification and normalization with an amplification efficiency of the target and reference genes. This calculation can be based on one sample (Souazé et al., 1996;; LightCycler® Relative Quantification Software, 2001) or multiple samples (Pfaffl, 2001a,, 2001b) and their formulas are given in Eqs. 1-6 respectively. Reactions for the determination of efficiencies of the genes should be run in a 5 or 10-fold serially diluted sample. In new approaches, multiple reference genes is used to obtain more stable and reliable results (Pfaffl, 2002). An efficiency

corrected calculation models, based on multiple samples and reference genes (so-called REF index), should consist of at least three reference genes (Pfaffl, 2001b).

Analysis of the raw data in precise mathematical and statistical manner should be performed rationally in gene expression analysis. To calculate the expression of a target gene in relation to an adequate reference gene various mathematical models are established. Calculations are based on the comparison of the distinct cycle determined by various methods, e.g., crossing points (CP) and threshold values (Ct) at a constant level of fluorescence; or CP acquisition according to established mathematics algorithm Two types of relative quantification models are available and published: (1) without efficiency correction (see Eqs.1 and 2) (Livak and Schmittgen, 2001)

$$R = 2^{-[\Delta CP\ sample - \Delta CP\ control]} \quad (1)$$

$$R = 2^{-\Delta\Delta CP} \quad (2)$$

and (2) with kinetic PCR efficiency correction (Eqs.3-6). For a group-wise comparison for more samples (up to 100), e.g., REST and REST-XL. The

relative expression ratio of a target gene is computed, based on its real-time PCR efficiencies (E) or a static efficiency of 2, and the crossing point (CP) difference (Δ) of one unknown sample (treatment) versus one control (ΔCP control - treatment). Using REST and REST-XL, the relative calculation procedure is based on the MEAN CP of the experimental groups (Eq. 4)

(3)
$$ratio = \frac{(E_{target})^{\Delta CP_{target}\,(control - sample)}}{(E_{ref})^{\Delta CP_{ref}\,(control - sample)}}$$

(4)
$$ratio = \frac{(E_{target})^{\Delta CP_{target}\,(MEAN\ control - MEAN\ sample)}}{(E_{ref})^{\Delta CP_{ref}\,(MEAN\ control - MEAN\ sample)}}$$

In these models the target-gene expression is normalized by a non-regulated reference gene expression, e.g., derived from classical and frequently described housekeeping genes. The crucial problem in this relative approach is that the most common reference-gene transcripts from so-called housekeeping genes, whose mRNA expression can be regulated and whose levels vary significantly with treatment or between individuals (Pfaffl et al., , 2002, Pfaffl et al., 2001). However, relative quantification can generate useful and biologically relevant information when used appropriately.

(5) $$\text{ratio} = \frac{(E_{ref})^{CP_{sample}}}{(E_{target})^{CP_{sample}}} \div \frac{(E_{ref})^{CP_{calibrator}}}{(E_{target})^{CP_{calibrator}}}$$

(6) $$\text{ratio} = \frac{\text{conc}_{(\text{target sample})} / \text{conc}_{(\text{ref sample})} * \text{MF}}{\text{conc}_{(\text{target cal.})} / \text{conc}_{(\text{ref cal.})} * \text{CF}}$$

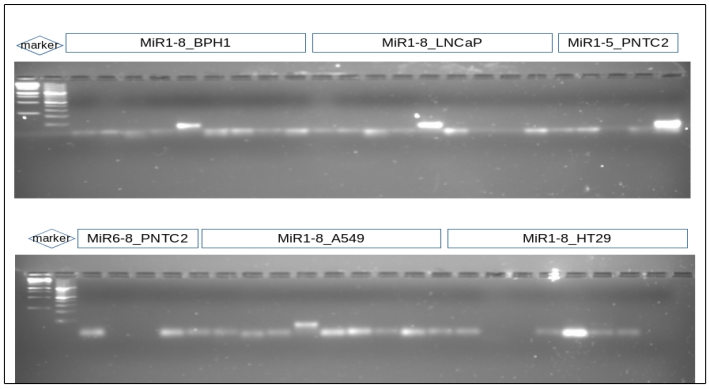
The selection of an appropriate constitutively expressed reference gene for qRT-PCR analysis is crucial when determining the expression levels of genes and miRNA. It is necessary to compensate for differences, such as inter-sample variances and amplification differences which can easily affect the reliability of qRT-PCR results. An appropriate reference gene would exhibit constant expression in all tissues and under all conditions and would thus allow for the comparison of samples (Kubista *et al.,* 2006; VanGuilder *et al.,* 2008). Quite a large selection of reference genes is available however, none are as effective in miRNA studies when compared to measuring normal gene expression levels. In this study two housekeeping miRNAs, Mir191a and miR17a, were selected and the housekeeping gene (HKG), GAPDH, was used as an additional control. PTEN expression levels were also measured as an additional control for expression of a gene under cancerous conditions. The results of the analysed data were plotted as

X versus Y and included the mean/ standard deviation (SD) for each analysis.

3.3 Results

Conventional PCR was performed to determine whether the stem-loop primers designed were capable of exponentially amplifying the miRNA fragments in the cDNA synthesised from the RT method described in section 3.2.5. The PCR products were separated on a 2 % agarose gel and visualised using the UVP system. In Figure below it is clear that miRNA 1-8 fragments were amplified successfully and that different intensities in the bands produced is visible as indicated in Figure 3.2.

A



B

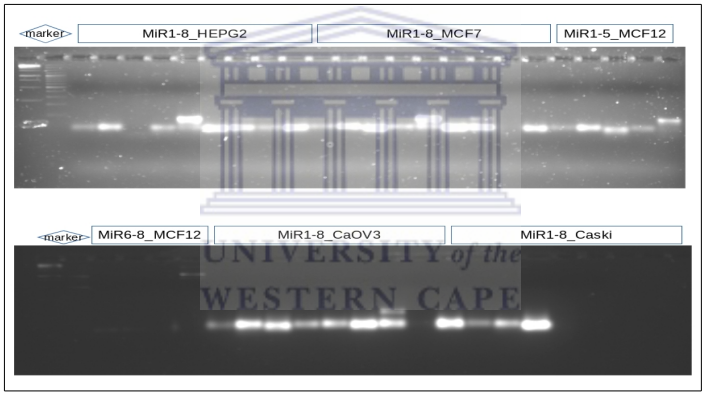


Figure 3.2: Representation PCR amplification of miRNAs with sequence specific primers for each cell line used in this study. Bands represent miRNA 1-8 in various cell lines. In (A) BPH1, LNCaP and PNTC2 represent prostate cell lines, HePG2 and A549 represent a colon and a lung cell lines respectively. In (B) MCF7 and MCF12 represent breast cell lines and HepG2, CaOV3 and Caski represents lung, ovarian and cervical cancer cell lines respectively.

3.3.1 Determining the PCR efficiency for miRNAs 1-8 using a standard curve:

In order to normalize the relative quantification of the miRNAs to internal controls (miRNA191A, miRNA17A, GAPDH and PTEN), the sets of crossing points for the reference genes were imported into REST (Relative Expression Software Tool). The chosen crossing points were considered during the calculation process of the software, which allowed for normalization of the target genes with the reference genes. The amplification plot for the reference gene GAPDH is shown in Figure 3.3 (a). Figure 3.3 (b) shows the amplification plot for the miRNA5, plots for the rest of the miRNAs and controls are present in Appendix B. In both Figure 3.3 (a) and (b), evenly spread out slopes were observed.

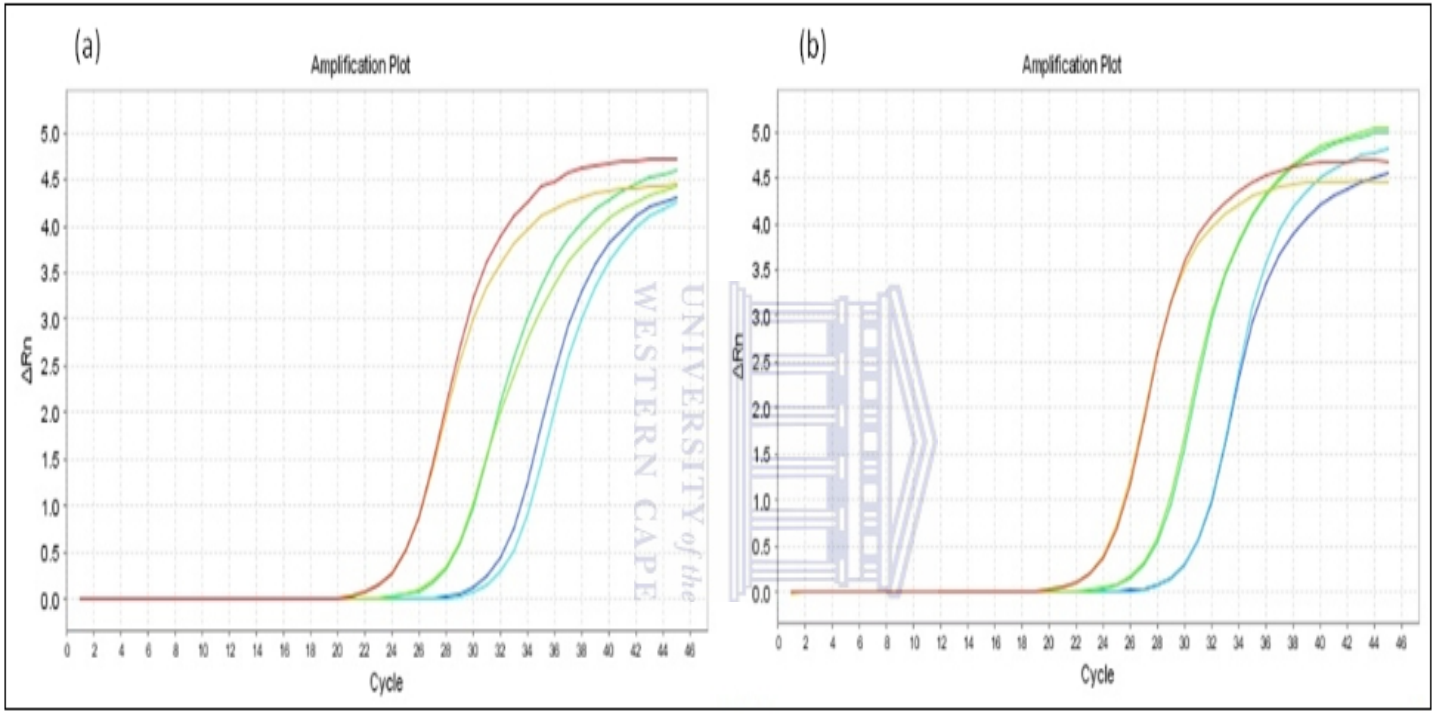


Figure 3.3 (a): Amplification plot for the housekeeping gene, GAPDH. The red, green, and blue slopes indicate the amplification of GAPDH at the 1:10 1:100, and 1:1000 dilutions respectively.

Figure 3.3 (b): Amplification plot for the miRNA5. The red, green, and blue slopes indicate the amplification of miRNA5 at the 1:10, 1:100, and 1:1000 dilutions respectively.

3.3.2 Calculating the PCR Efficiency

The amplification efficiencies of the qRT-PCR reactions, performed with serial dilutions of the templates, are based on the slopes of the standard curves obtained. Table 3.3 shows the efficiencies of all the miRNAs and housekeeping genes in the normal-, prostate cancer- as well as the other cancer cell lines. These values range from 91-98 % and are highly similar, indicating that a similar rate of amplification occurred for each miRNA. The amplification efficiency for each miRNA was used in the REST software to determine whether they were differentially expressed in cancer cell lines when compared to their control non-cancerous cell lines.

Table 3.3: Amplification efficiencies of the MiRNAs and housekeeping genes in the normal and cancer cell lines.

Cell line / MiRNA / gene	PNT2C2	BPH1	LNCaP	MCF-12A	MCF-7	KMST1	A549	HT29	HepG2	Caski	CaOV3
GAPDH	97	95	94	94	96	93	93	94	94	96	95
MiRNA1	93	95	95	96	94	95	97	93	95	94	93
MiRNA2	93	95	92	93	94	96	96	96	94	97	94
MiRNA3	96	94	93	92	94	94	97	92	96	97	93
MiRNA4	95	93	94	94	97	96	97	94	93	95	97
MiRNA5	95	96	98	94	94	94	98	97	98	96	92
MiRNA6	95	94	94	93	93	94	95	96	93	96	93
MiRNA7	94	92	91	93	93	97	95	91	91	94	92
MiRNA8	95	93	92	93	97	95	98	93	92	94	97
MiRNA191A	97	96	94	95	94	93	94	96	92	94	95
MiRNA17A	94	92	91	94	92	94	97	95	96	95	94
PTEN	94	96	98	94	93	93	94	93	94	92	96

3.3.3 Melting curve analysis

A melting curve ranging from 60°C to 95°C was constructed for every reference and target gene. The obtained melting curves were used to determine whether any contamination, mis-priming (referring to the annealing of primers to complementary sequences on non-target DNA), primer-dimers (primers annealing to themselves), or other inconsistencies occurred relating to the amplification process. The melting curve for the reference gene GAPDH and miRNA5 are shown in Figure 3.4 (a) and 3.4 (b). From the results it is evident that only one peak is observed (one T_m per gene/miRNA) for both GAPDH and miRNA5. Melting curves for the rest of the miRNAs as well as the controls is present in Appendix B.

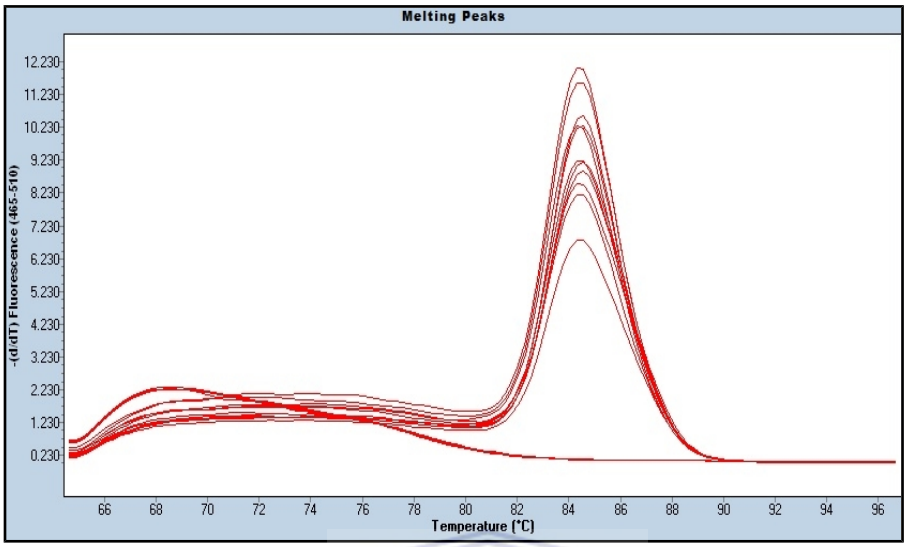


Figure 3.4 (a): Melting curve for the housekeeping gene, GAPDH, ranging from 65 °C to 95 °C.

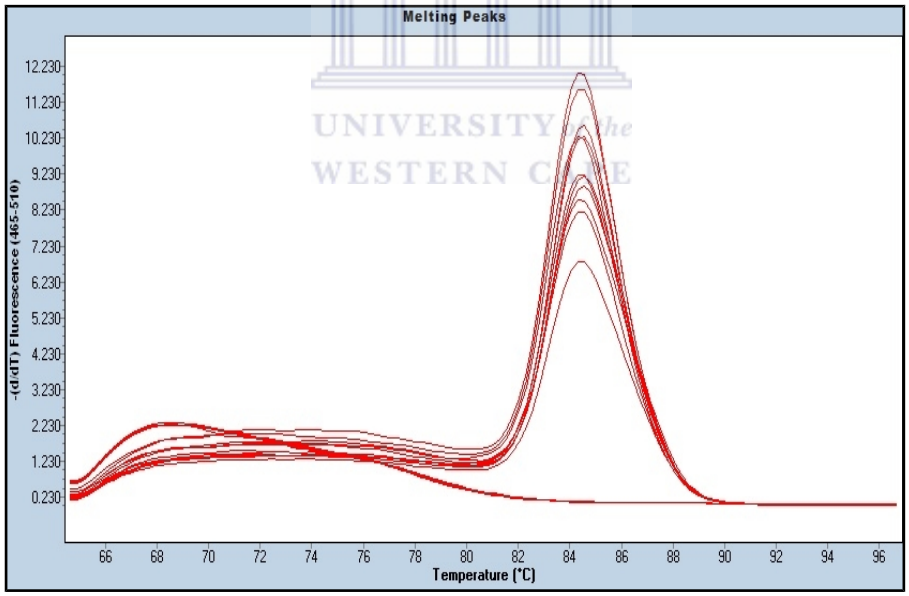


Figure 3.4 (b): Melting curve for miRNA5, ranging from 65 °C to 95 °C.

3.3.4 Expression Profiling

3.3.4.1 Determining the expression levels of internal controls across cancer cell lines and control cell lines.

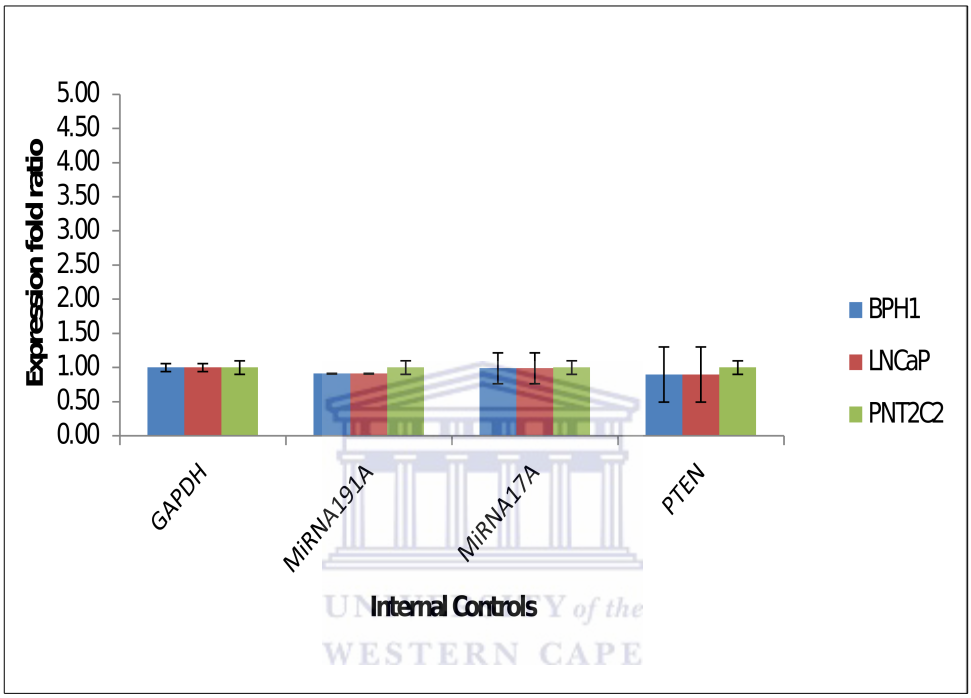


Figure 3.5: Relative expression of internal controls in the BPH1, LNCaP and PNT2C2 prostate cell lines. Where BPH1 is shown in blue bars, LNCaP is shown in red bars and PNT2C2 is shown in green bars.

From Figure 3.5 the housekeeping miRNA's miRNA191A and miRNA17A remain stable across the 3 tested prostate cell lines. GAPDH a known house keeping gene shows no change in its expression in any of the cell lines tested. Similarly PTEN a known general cancer biomarker (Chu and

Tarnawski, 2004), also showed no change in expression levels in the 3 tested prostate cell lines.

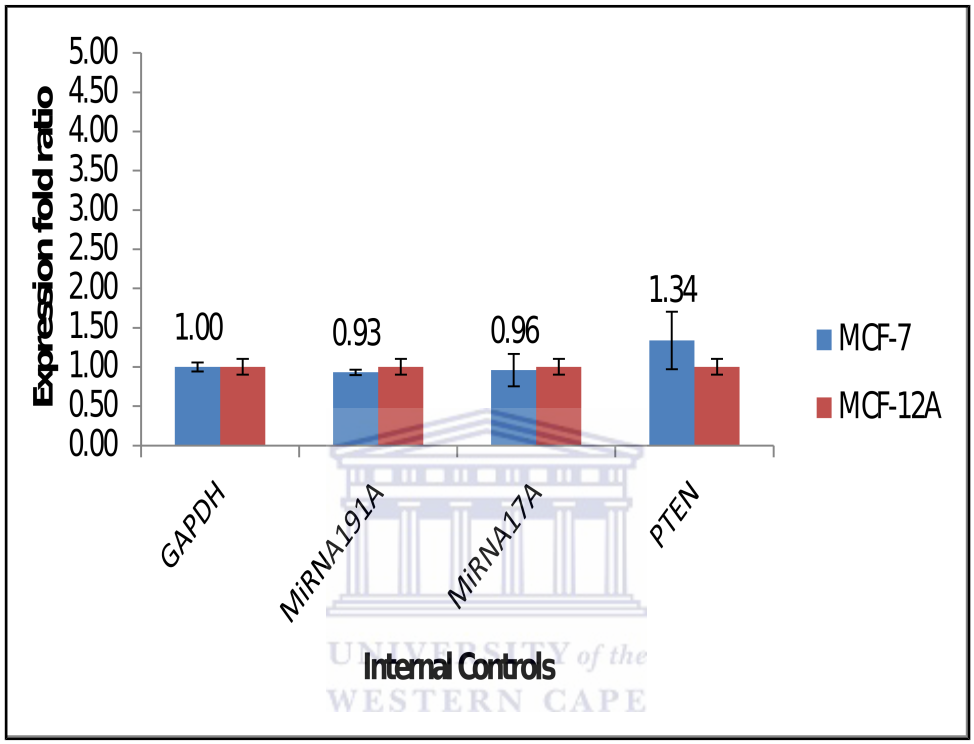


Figure 3.6: Relative expression of internal controls in MCF7 and MCF12A breast cell lines. Where MCF7 is shown in blue bars and MCF12A is shown in red bars.

Figure 3.6 shows the relative expression of the 4 internal controls which included 2 miRNA controls namely miRNA191A and miRNA17A,

GAPDH as well as PTEN in MCF7 and MCF12 A. From the graphs in Figure 3.6 it is clear that GAPDH, miRNA191A and miRNA17A remained stable across cell lines. PTEN showed a slight increase in expression in the MCF7 (breast cancer) cell line however it was not considered significant.

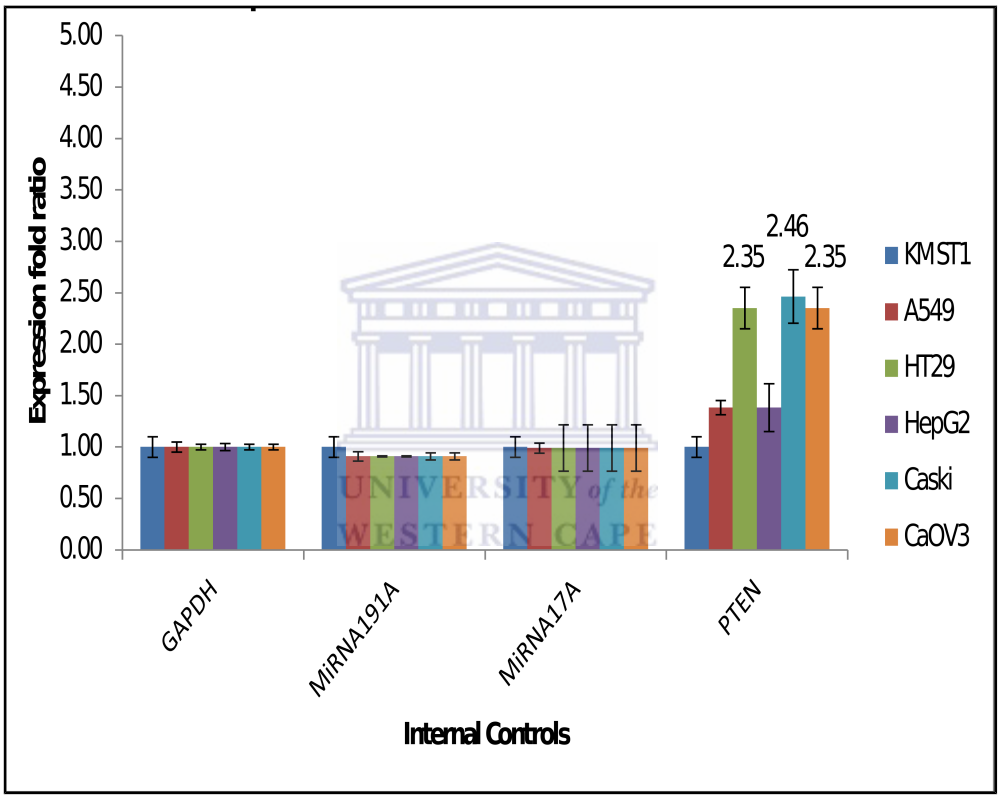


Figure 3.7: Relative expression of internal controls across 5 cancer cell lines namely A549 (lung), HT29 (colon), HepG2 (liver), Caski (cervical) and CaOV3 (ovarian) and KMST1 (normal) fibroblast cell line.

Figure 3.7 shows the relative expression of the 4 internal controls namely, miRNA191A, miRNA17A, GAPDH as well as PTEN, across several cancer cell lines compared to a KMST a normal cell line. From the data in the graph it is shown that GAPDH, miRNA191A and miRNA17A expression remained stable across cell lines. The internal control PTEN showed increased level of expression across all of the cell lines. PTEN showed a 2.46 fold expression ratio in the Caski cell line. In the HT29 and CaOV3 cell lines an expression ratio of 2 .35 fold for PTEN was observed . A549 and HepG2 showed a 1 fold increase in expression for this marker.

3.4.1.2 Differential Expression analysis of MiRNA 1-8 across cell lines when compared to normal cell lines.

Table 3.4 Differential expression of miR1-8 in BPH1 in comparison to PNT2C2 cell lines.

MiRNA/ Gene	Cell line	Mean Ct	Std error	Fold regulation	p-value
GAPDH	PNT2C2	18.19	0.04	0	
	BPH1	18.05	0.11	0	0
miRNA1	PNT2C2	34.67	0.33		
	BPH1	34.67	0.33	-1.098	0.178
miRNA2	PNT2C2	34.67	0.32		
	BPH1	34.67	0.27	1.046	0.214
miRNA3	PNT2C2	27.74	0.06		
	BPH1	25.24	0.21	5.129	0.002
miRNA4	PNT2C2	31.97	0.07		
	BPH1	31.93	0.03	-1.073	0.156
miRNA5	PNT2C2	28.17	0.03		
	BPH1	26.61	0.18	2.686	0.001
miRNA6	PNT2C2	35	0		
	BPH1	35	0	-1.098	0.321
miRNA7	PNT2C2	29.2	0.06		
	BPH1	29.03	0.03	1.022	0.551
miRNA8	PNT2C2	31.21	0.12		
	BPH1	31.84	0.19	-1.707	0.416

Table 3.4 is a representation of the descriptive statistical analysis performed following the determination of the expression levels for each miRNA. In this table the expression levels of miRNA1-8 was compared between the control prostate cell line, PNT2C2, and the benign prostate cell line, BPH1. The data is represented by bar graphs in Figure 3.8 with the specific fold

change ratio indicated for each miRNA. The data showed that two miRNAs, miRNA3 and miRNA5 were significantly up-regulated in the benign prostate cell line by factors of 5 and 2.6 respectively as seen by the highlighted values in Table 3.4 and blue bars in Figure 3.8. These identified miRNA using a bioinformatics pipeline were not shown to be up-regulated in the prostate previously by *in silico* expression analysis using GEA as described in section 2.2.7 and shown in Figure 2.10.

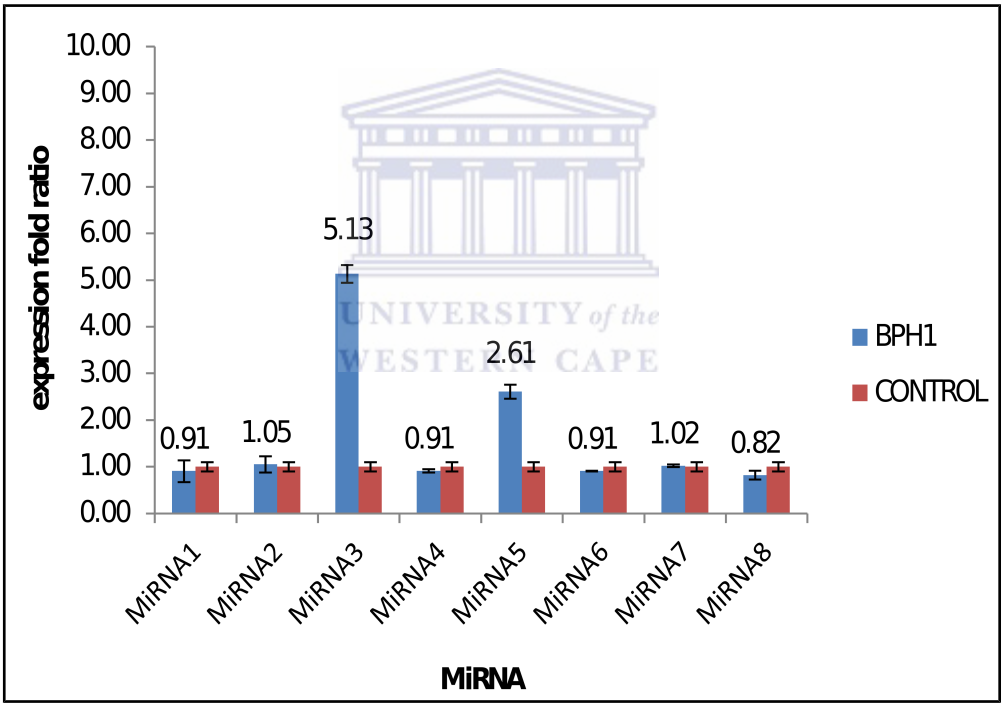


Figure 3.8: Relative expression ratio plot comparing the expression of miRNA 1-8 in the BPH1 (Benign Prostate Hyperplasia) cell line compared to PNT2C2 (normal).

Table 3.5 Differential expression analysis of miR1-8 in LNCaP in comparison to PNT2C2 cell lines.

miRNA	Cell line	Mean Ct	Std error	Fold regulation	p-value
GAPDH	PNT2C2	18.19	0.04	0	
	LNCaP	18.05	0.11	0	0
miRNA1	PNT2C2	34.67	0.33		
	LNCaP	35	0.17	-1.38	0.170
miRNA2	PNT2C2	34.67	0.32		
	LNCaP	34.57	0.09	-1.02	0.201
miRNA3	PNT2C2	27.74	0.06		
	LNCaP	26.30	0.21	2.47	0.003
miRNA4	PNT2C2	31.97	0.07		
.	LNCaP	32.03	0.09	-1.15	0.150
miRNA5	PNT2C2	28.17	0.03		
	LNCaP	25.96	0.04	4.263	0.001
miRNA6	PNT2C2	35	0		
	LNCaP	35	0	1.1	0.301
miRNA7	PNT2C2	29.2	0.06		
	LNCaP	29.07	0.07	-1.0	0.540
miRNA8	PNT2C2	31.21	0.12		
	LNCaP	28.05	0.12	8.1	0.001

Table 3.5 is a representation of the descriptive statistical analysis performed for the expression levels of miRNA1-8 for LNCaP and PNT2C2 cell lines and the data represented as a bar graph as shown in Figure 3.9. The data shows that three miRNAs, miRNA3 miRNA5 and miRNA8 were significantly up-regulated in the cancer cell line by factors of 2.5, 4 and 8

respectively as seen by the highlighted values in Table 3.5 and the blue bars in Figure 3.9 below. Tables containing the descriptive statistics for miRNA 1-8 across the remaining cell lines tested in the study are available in Appendix B. From the *in silico* expression analysis using GEA as described in chapter 2 section 2.2.7 in Figure 2.11 miRNA3 was shown to be up-regulated in PC however, miRNA5 and miRNA8 were not found to be previously up-regulated in PC using *in silico* analysis.

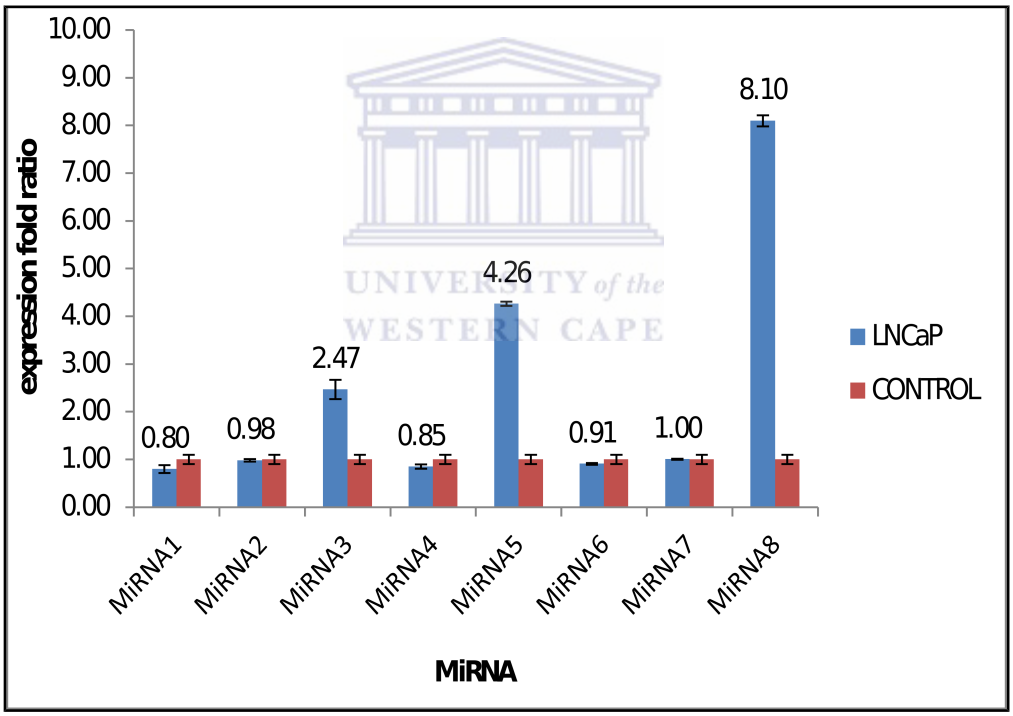


Figure 3.9: Relative expression ratio plot comparing the expression of miRNA 1-8 in the LNCaP (Metastatic prostate carcinoma) cell line compared to PNT2C2 (normal prostate)

Figure 3.9 shows the comparative expression analysis of miRNA1-8 between the control cell line, PNT2C2, and the prostate cancer cell line, LNCaP. The cell line LNCaP is different from the BPH1 prostate cell line, in that it is an adenocarcinoma, whereas, BPH1 is a benign prostatic hyperplasia. The expression of miRNA8 differs in these two cell line, as it is highly up-regulated in LNCaP, whereas it is unregulated in BPH1. The two miRNAs, miRNA3 and miRNA5 are also up-regulated in LNCaP. This finding is very important as it allows one to differentiate between Benign prostatic hyperplasia and metastatic prostate cancer when compared to the normal PNT2C2 prostate line using miRNA8.

PTEN as shown in Figure 3.5 is also stably expressed between these cell lines and the control cell line, indicating that it is not a useful biomarker to distinguish between the benign (BPH1), malignant (LNCaP) and normal prostate (PNT2C2) cell lines. MiRNA8 is also highly specific to only LNCaP as it does not show any differential expression in any of the other cancer cell lines evaluated in this study. However, this miRNA must still be evaluated in prostate tissue to ensure its specificity and sensitivity. Also,

additional cancer cell lines need to be tested to ensure that miRNA8 is not up-regulated or down-regulated in other cancer cell types.

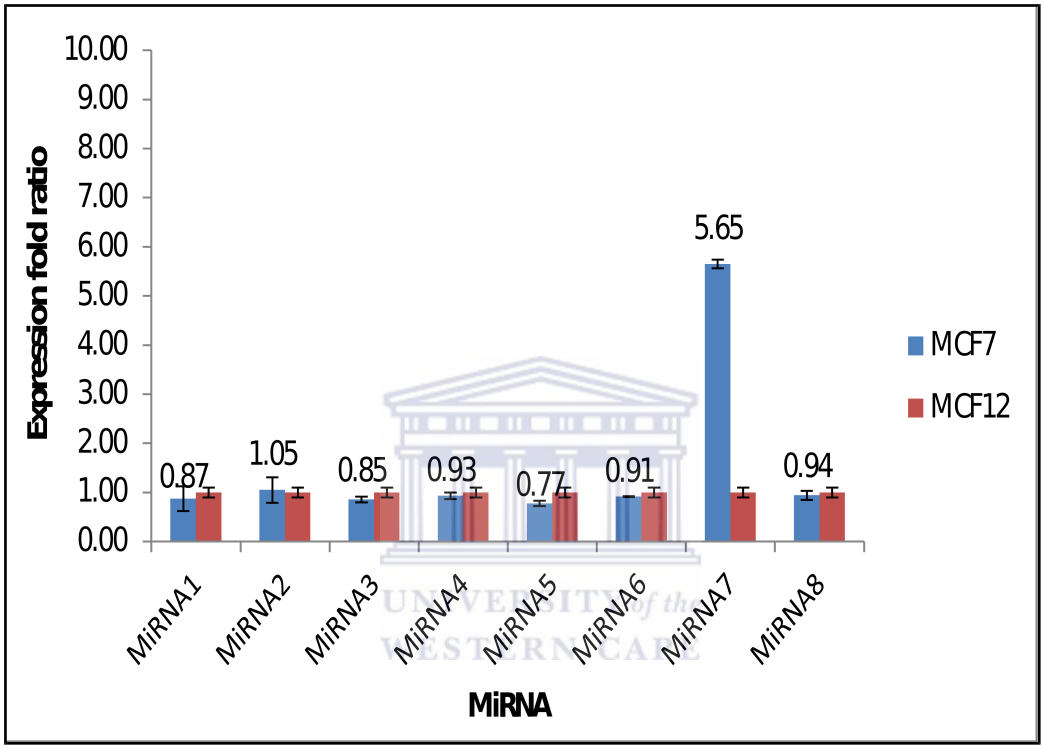


Figure 3.10: Relative expression ratio plot comparing the expression of miRNA 1-8 in the MCF7 (Breast) cell line compared to MCF12A.

Figure 3.10 shows the comparative expression analysis of miRNA1-8 between the normal breast control cell line, MCF12A, and the breast cancer cell line, MCF7. Only one of the miRNAs, miRNA7, showed differential

expression between the two breast cell lines. The rest of the miRNAs are all expressed at the same levels between the MCF7 and MCF12A cell line.

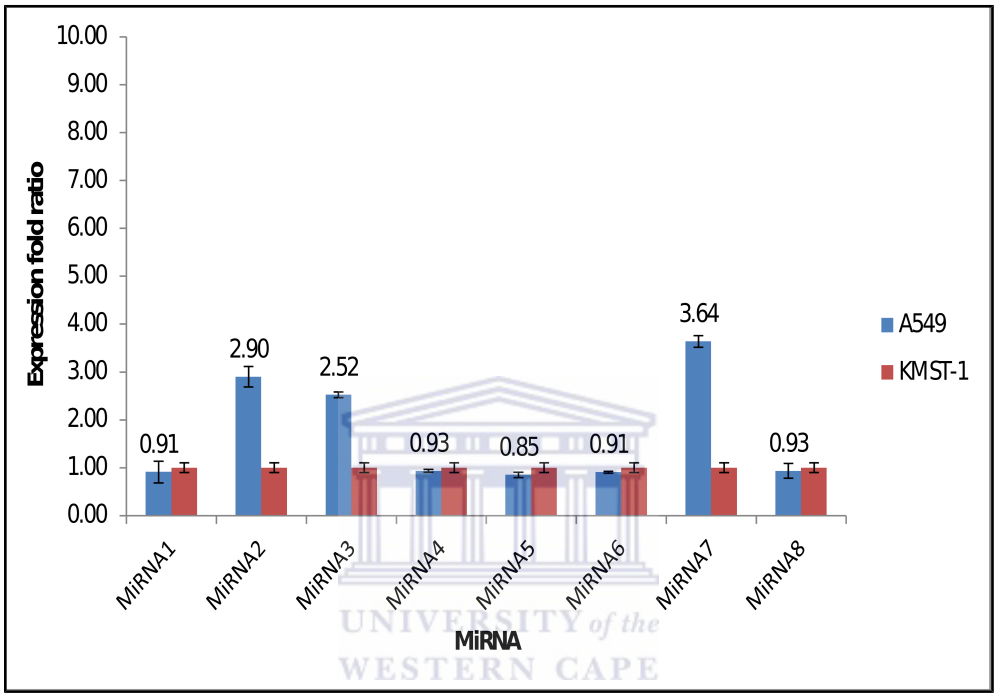


Figure 3.11: Relative expression ratio plot comparing the expression of miRNA1-8 in the A549 (Lung) cell line compared to KMST1 (normal)

Figure 3.11 shows the comparative expression analysis of miRNA1-8 between the KMST1 control cell line, and the lung cancer cell line, A549. Three of the miRNAs: miRNA2, miRNA3 and miRNA7, showed up-regulation in the lung cancer cell line. The rest of the miRNAs are all

expressed at the same level.

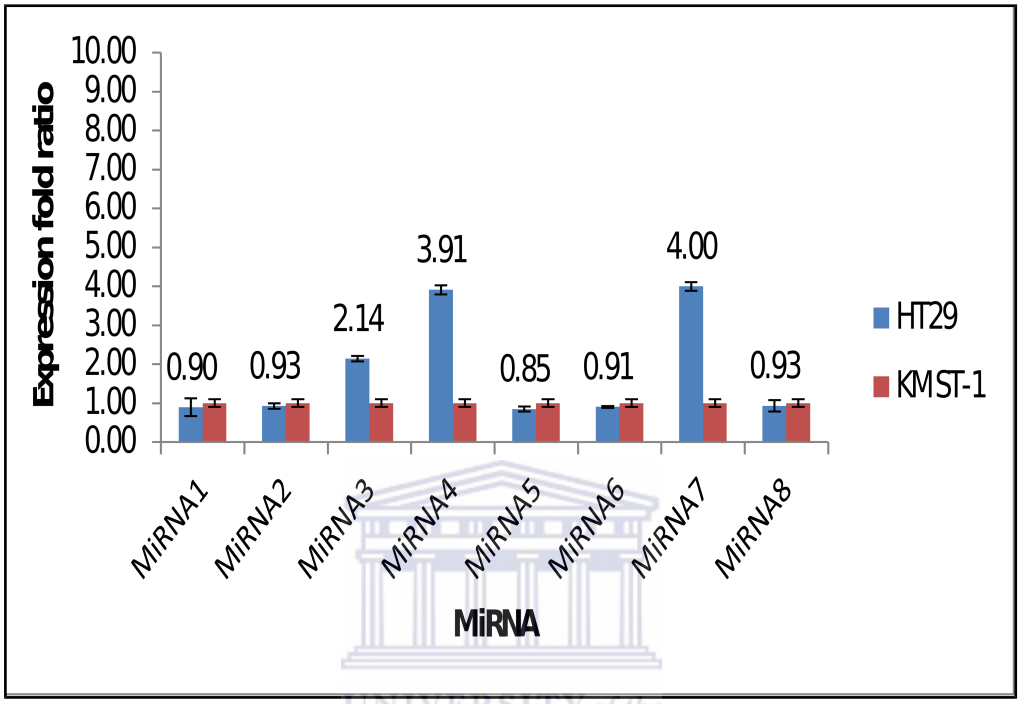


Figure 3.12: Relative expression ratio plot comparing the expression of miRNA 1-8 in the HT29 (Colon) cell line compared to KMST1 (normal) fibroblast cell line.

Figure 3.12 shows the comparative expression analysis between the KMST1 control cell line, and the colon adenocarcinoma, HT29. Three of the miRNAs: miRNA-3, miRNA-4 and miRNA-7, showed up-regulation in

the colon cancer cell line. The rest of the miRNAs are all expressed at the same level when comparing their expression in HT29 to KMST.

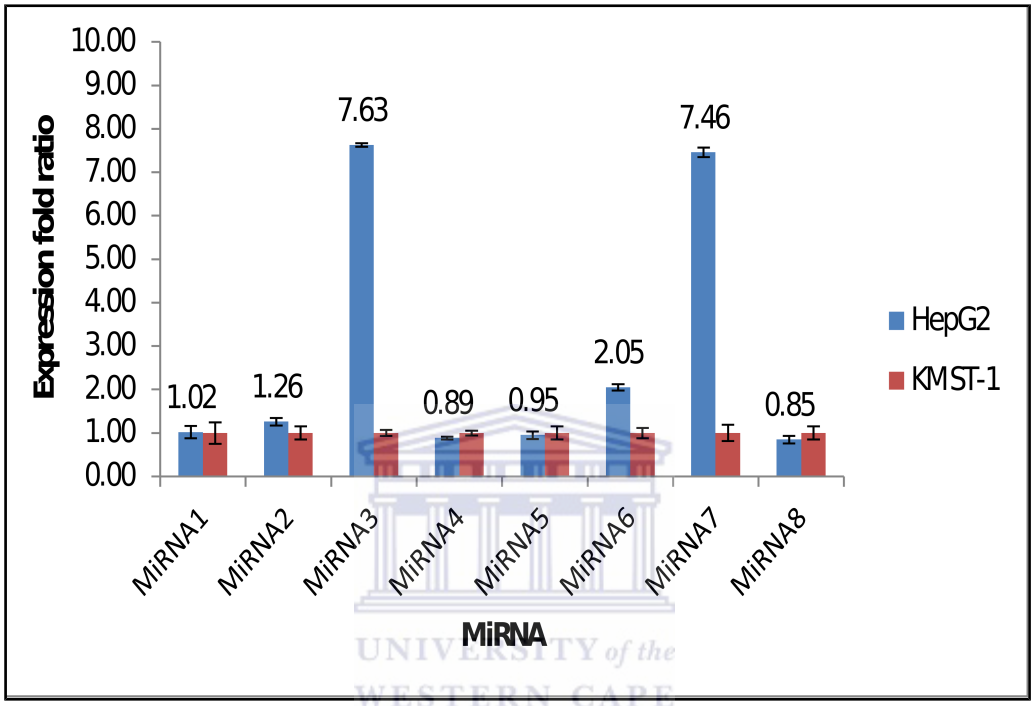


Figure 3.13: Relative expression ratio plot comparing the expression of miRNA 1-8 in the HepG2 (Liver) cell line compared to KMST1 (normal)

From the data shown in Figure 3.13 two of the miRNAs: miRNA3 and miRNA7, showed up-regulation in the liver cancer cell line. The rest of the miRNAs are all expressed at the same level when comparing their

expression in HepG2 to KMST. The known cancer marker PTEN was significantly up-regulated in HepG2 as seen in Figure 3.7.

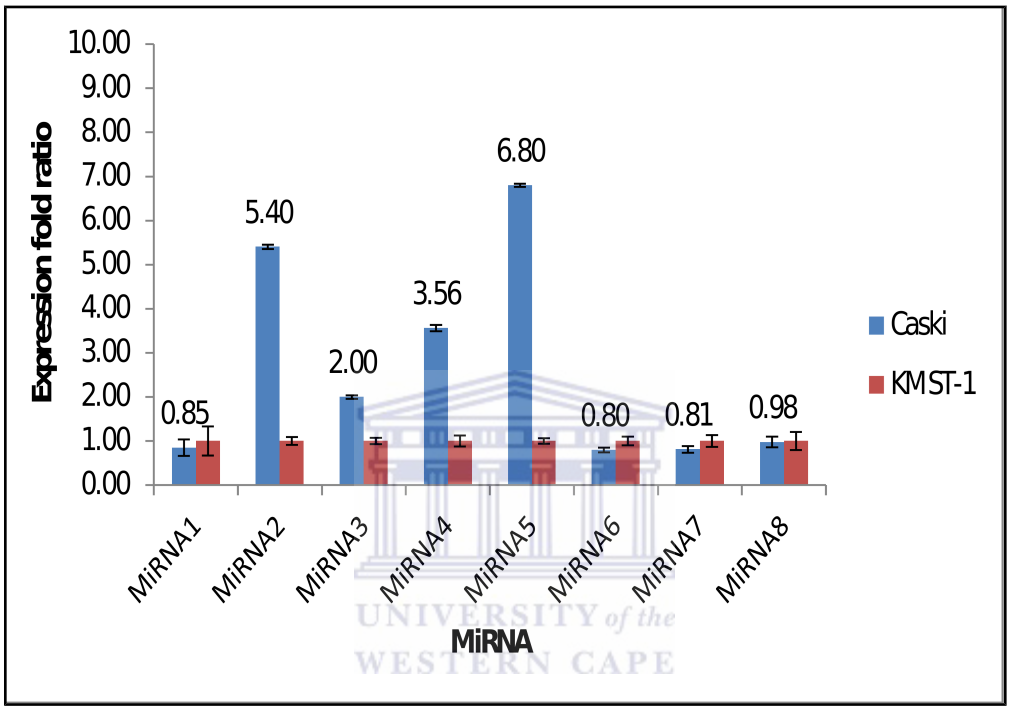


Figure 3.14: Relative expression ratio plot comparing the expression of miRNA 1-8 in the Caski (Cervical) cell line compared to KMST1

Figure 3.14 shows the comparative expression analysis of miRNA1-8 between the KMST1 control cell line, and the cervical carcinoma, Caski. Three of the miRNAs: miRNA2, miRNA4 and miRNA5, showed an expression fold change of 5.4, 3.6 and 6.8 respectively in the cervical

cancer cell line. miRNA3 was up-regulated by a factor of 2 and needs further investigation as to whether it can be considered a biomarker for cervical cancer. The rest of the miRNAs are all expressed at the same levels when comparing their expression in Caski to KMST. The known cancer marker PTEN (Chu and Tarnawski, 2004) was significantly up-regulated in Caski as shown in Figure 3.7.

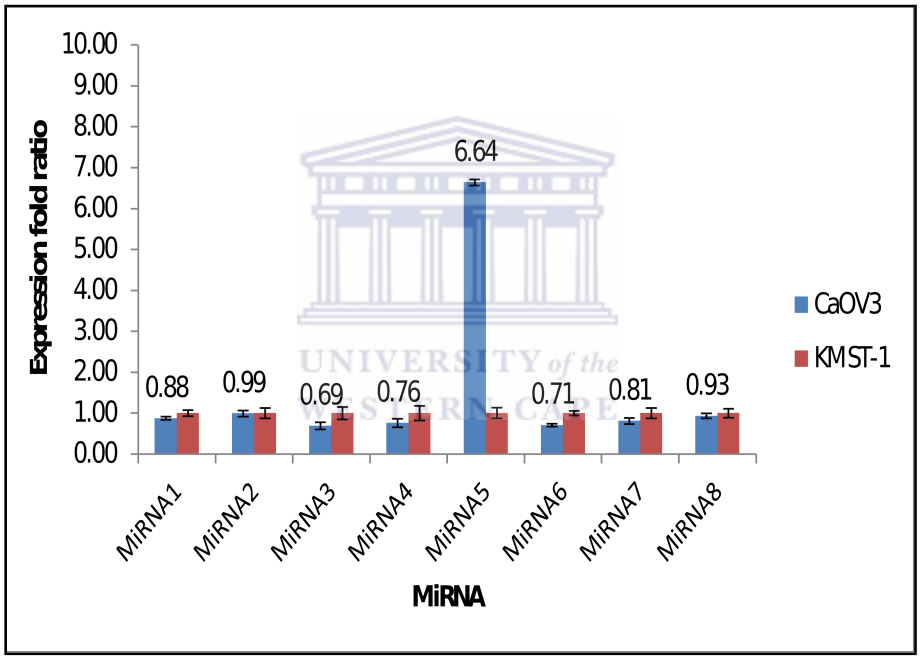


Figure 3.15: Relative expression ratio plot comparing the expression of miRNA 1-8 in the CaOV3 (Ovarian) cell line compared to KMST1 (normal)

Figure 3.15 shows the comparative expression analysis of miRNA1-8 between the KMST1 control cell line, and the ovarian carcinoma, CaOV3. Only one of the miRNAs, miRNA5, showed up-regulation in the ovarian cancer cell line. The known cancer marker PTEN was significantly up-regulated in CaOV3 as shown in Figure 3.7.

Table 3.6 Summarized table of expression profiles for each miRNA in the different cancer cell lines investigated.

	PNT2C2	LNCaP	BPH1	MCF7	MCF12	KMST1	CaOV3	Caski	HepG2	A549	HT29
miRNA1	0				0	0					
miRNA2	0				0	0		+		+	
miRNA3	0	+	++		0	0		+	+	+	+
miRNA4	0				0	0		+			+
miRNA5	0	+	+		0	0	+	+			
miRNA6	0				0	0			+		
miRNA7	0			+	0	0			+	+	+
miRNA8	0	++			0	0					
miRNA17A	0				0	0					
miRNA191A	0				0	0					
GAPDH	0				0	0					
PTEN	0				0	0	+	+	+		+

KEY: 0 - no expression

+ - over-expression

++ - highly over-expressed

3.4 Discussion

The aim of this chapter was to molecularly validate the putative miRNA biomarkers identified in Chapter 2 using qRT-PCR. The 8 miRNAs studied in vitro were selected from an initial 40 based on their association with PC. Based on the identified miRNAs strong association with PC, these miRNAs were selected as potential biomarkers. The relative expression levels of the miRNAs were determined with qRT-PCR as described in section 3.2.6. Based on the number of variations in expression patterns that could occur as a result of (a) method of sample preparation or (b) cDNA synthesis, it is pivotal to select good control miRNAs and/or genes to compensate for such variations and to improve the reliability of qRT-PCR experiments.

An appropriate reference gene would exhibit constant expression in all tissues and under all conditions and would thus allow for the comparison of samples (Kubista et al., 2006; VanGuilder et al., 2008). Quite a large selection of reference genes is available however, none are as effective in miRNA studies when compared to measuring normal gene expression levels. During the course of this study 4 controls were selected namely miRNA191A, miRNA17A, GAPDH and PTEN. GAPDH is a suitable

housekeeping gene as it is present in all nucleated cell types and mRNA expression of this gene was found to be stable in different cell lines including cancers (Hao *et al.,* 2011; Mattie *et al.,* 2006).

The PTEN gene is known as a 'tumour suppressor gene' because deletion of the PTEN gene increases the likelihood of developing cancer. PTEN makes a protein that switches off the PI3K/AKT/mTOR signalling pathway, which is responsible for cell growth, survival, and movement (Chu and Tarnawski, 2004). Cells require two copies of the PTEN gene to ensure that they function properly. In prostate cancer cells, one or both copies of the PTEN gene is frequently lost, thus allowing the PI3K/AKT/mTOR signalling pathway to trigger uncontrolled cell growth, survival, and movement. Recently, doctors have found that a loss of PTEN that occurs in combination with an abnormal fusion of the genes TMPRSS2 and ERG is a prognostic indicator of poorer outcomes (Chu and Tarnawski, 2004). PTEN deletions have also been reported in glioma, melanoma, endometrial cancer, kidney cancer, breast cancer, lung cancer, prostate cancer, and upper respiratory tract cancers. As a result of the above evidence PTEN is a currently known cancer biomarker (Chu and Tarnawski, 2004).

Conventional PCR as shown in Figure 3.2 shows clear bands in each lane

which is an indication of optimal amplification of cDNA using stem-loop sequence specific primers. PCR results revealed that the RT method works successfully for samples from cell culture as well as from tissue samples (tissue results not shown here). Preliminary results showed that miRNAs extracted from both cell lines and tissue sample are reverse transcribed successfully as shown in Figure 3.2. It is evident from the figure that the miRNA 1-8 differentially expressed across cell lines based on the intensity of the bands seen. The next step was to evaluate their expression levels in three prostate lines BPH1, LNCaP and PNT2C2 a benign, cancer and normal prostate cell line respectively. In addition, 6 other cell lines (five cancer and one control) were also evaluated to determine the expression levels and specificity of the miRNAs investigated.

The identified biomarkers were molecularly validated using qRT-PCR as described in section 3.2.6. PCR efficiencies were calculated and tabulated as indicated in Table 3.3. These values for each of the miRNAs as well as the controls, ranged from 91-98 % and are highly similar across cell lines. These results indicate that a similar rate of amplification occurred for each miRNA as well as for the controls.

During the qRT-PCR assays performed, amplification efficiencies between 3.3 and 3.5 were obtained for the respective miRNAs as seen in Figure 3.3 (a) and 3.3 (b).These results indicate that an optimised qRT-PCR reaction was performed for all primer sets tested in this study. All results for this part of the analysis is displayed in Appendix B.

The melting peak and curve analysis for all the miRNAs showed one distinctive peak, Melting peaks for GAPDH and miRNA5 are shown in Figure 3.4 (a) and 3.4 (b), melting curves for the rest of the miRNAs can be viewed in Appendix B. As a result of the similarity between the distinctive peaks produced for both GAPDH and miRNA5 it can be deduced that only the expected amplicon was amplified and that no contamination, mis-priming or primer-dimers occurred or are present in the PCR reaction. Therefore, no non-specific amplification occurred and it is evident that accurate quantification of the targets of interest has been achieved through optimized qRT-PCR.

The randomization test were performed using the REST software (Pfaffl, 2002), to determine whether normalization via the reference miRNAs and/

or genes were useful. The results of the randomization test displayed the factor of regulation of the standards and level of significance at which the genes are being expressed, which indicated that the reference genes used were suitable in this experimental trial (see Appendix B).

The relative expression of the 4 controls were measured across all cell lines tested in this study. Figure 3.5 shows the expression of the 4 controls in the 3 prostate cell lines tested. From the data shown it is evident that the expression of all the controls remained stable in benign (BPH1), cancerous (LNCaP) as well as in the normal prostate cell line (PNT2C2). Similarly in the breast cancer cell line MCF7 the expression patterns of the 4 controls remain stable as seen in Figure 3.6. However when investigating the relative expression of the 4 controls in the remaining cell lines tested during this study, a different expression pattern was observed for PTEN, whilst the rest of the controls remain stable across cell lines. PTEN was found to be up regulated in all 5 of the additionally tested cell lines as seen in Figure 3.7. This finding was expected as PTEN is a known cancer biomarker (Chu and Tarnawski, 2004) and its expression levels are expected to be higher in most cancer cell lines.

In this part of the study we determined the expression levels of all eight miRNA in two prostate cell lines, BPH1 (benign) and LNCaP (metastatic), and compared it to the expression levels in the prostate control cell line, PNT2C2 as seen in Figure 3.8 and 3.9 . This was to establish whether the miRNAs were differentially expressed in benign and prostate cancer. The data showed that miRNA3 and 5 were up-regulated in the BPH1 and LNCaP when compared to PNT2C2 (Figure 3.8). In addition miRNA8 was also shown to be up-regulated in LNCaP (Figure 3.9). Based on these results we could establish a miRNA profile to distinguish between the BPH1 and LNCaP prostate cell lines. miRNA8 was specific to LNCaP (Table 3.6), and this result is supported by a study conducted by Coffey *et al.,* 2002, where CLFAR a target of miRNA8 as indicated in table 2.4 was strongly associated with androgen-sensitive human prostate cancer (LNCaP). This study indicated that the androgen receptor pathway directly targets CFLAR although the molecular mechanisms are not clearly defined. MiRNAs 3 and 5, are implicated in cell growth and apoptosis regulation respectively, which are common molecular mechanisms associated with cancers in general suggesting that BPH1 could be a very early precursor of

PC. This notion is further established by a recent study by Hung and associates where the association of BPH and PC was demonstrated in men with an ethnic low-risk of PC. Asian men who were diagnosed with prostate cancer over a 10-year period and all of whom were over the age of 50 years had a strong association with prostatitis and/or BPH . There was a higher odds ratio (OR) for prostate cancer associated with BPH (26.2, 95% confidence interval (CI) 20.8-33.0) than with prostatitis (10.5, 95% CI = 3.36-32.7) (Hung et al., 2013).

Recent studies have further re-iterated the link between BPH and PC with genetic evidence. With the utilization of technology, such as next-generation sequencing, it has provided us the ability to uncover disease-related variants by processing large genomic intervals in a rapid and thorough manner. Mutations in the HBOX13 gene have been established to be associated with the development of prostate cancer. Saaristo and colleagues assessed 100 single-nucleotide polymorphisms (SNP) in over 500 men with BPH who either remained with this disease or later developed prostate cancer (Saaristo *et al*. 2013). As expected, a mutation in HBOX13 was strongly associated with the prostate cancer specimens.

However, BPH patients carrying this mutation had a 4.6 times greater risk of developing prostate cancer compared to non-carriers (OR 4.56, CI 95% 1.29-16.11, P = 0.0098). Interestingly, the average PSA levels at baseline biopsy in both groups were very similar (BPH = 7.3 μg/l (range 0.5-44 μg/l) and 8.0 μg/l (range 2.1-75 μg/l) respectively). This was the first study to report genetic predisposition to developing PC in BPH patients and gives a future potential of predicting those men with BPH who will later go on develop PC (Saaristo *et al*. 2013).

In addition, we also determined the expression of the identified miRNAs in breast cancer by comparing their expression profiles in MCF7, a cancer cell line, and MCF12A a normal breast cell line (Figure 3.10). In MCF7 miRNA7 was shown to be up-regulated when compared to the control cell line, MCF12A. In the breast cancer cell line the effect of miRNA7 on the development of cancer is not clear as the molecular mechanisms underlying its target CLN8 is not defined. However, studies have suggested its role in the possible regulation of apoptosis as one of CLN8's main functions (Szafranska-Schwarzbach *et al.,* 2011). Numerous studies additionally indicates the genes involvement in altered lipid metabolism, oxidative and

ER stress, mitochondrial dysfunction and inflammation which have previously been described as hallmarks of cancer in chapter 1. This result indicated that we could differentiate between breast cancer and PC by using a combination of miRNA 3, 5 and 7.

Furthermore, we also compared the expression levels of the miRNAs in 5 other cancer cell lines which included A549 (lung), HT29 (colon), HEPG2 (liver), Caski (cervical) and CaOV3 (ovarian) and compared it to the control cell line, KMST1 derived from human fibroblast as seen in Figures 3.11-3.15. The data as depicted in Figures 3.11-3.15 showed that miRNA2 was up regulated in the A549 and Caski cell lines. It was further evident from the data in the above-mentioned figures that miRNA3 was up-regulated in all the additional cell lines tested except CaOV3 when compared to the normal KMST1 cell line. MiRNA 4 was found to be up-regulated in the HepG2, HT29 and Caski cell lines. MiRNA 5 was shown to be up-regulated in the Caski and CaOV3 cell lines as seen in figure 3.14 and 3.15 similarly this miRNA was seen to up-regulated in BPH1 and LNCaP (prostate cell lines), these results may suggest that miRNA5 may be a regulator of reproductive mechanisms that may lead to tumour formation

and disease progression. MiRNA7 which targets the CLN8 gene as seen in Table 2.4 was found to be up-regulated in the HepG2, HT29 and A549 as seen in Figures 3.11-3.13 as well as in the MCF7 cell line as depicted in Figure 3.10. Based on the key function of its target gene CLN8 being the regulation of apoptosis as previously mentioned (Szafranska-Schwarzbach et al, 2011), these results may suggest that miRNA7 could also play an integral role in the regulation of the the apoptosis process.

Finally, the present study has shown as indicated in Table 3.6 that we can use an individual as well as a combination of all eight miRNA to create unique profiles for all the cancer cell lines tested in this study. For example, we could distinguish between Caski, a cervical cancer cell line, and CaOV3 an ovarian cancer cell line using a combination of different miRNAs e.g. miRNAs2, 3 and 4 for Caski (Table 3.6). Similarly a combination of miRNA2 and 3 as opposed to miRNA3 and 6, could be used to distinguish between A549 and HepG2 respectively. This is an important finding which could have potential diagnostic applications in clinical practice.

The results obtained suggest that one miRNA on its own could be

sufficient to differentiate between benign and cancer cell lines, however it should not be disregarded that by creating a unique profile for each cancer cell line by using a combination of miRNAs is a possible approach as well. The results also suggest that some of these cancers could share common molecular mechanism as the same miRNAs were differentially expressed across cancerous cell lines (Table 3.6). The identified miRNAs were shown to regulate the expression of the same target genes as depicted in Table 2.4 in chapter 2 .

CHAPTER 4

General Discussion

CONCLUSIONS AND FUTURE WORK

CHAPTER 4

General Discussion

Prostate cancer (PC) is the most frequent tumour in men and a major cause of cancer-related morbidity and mortality. The current marker of choice is PSA, but this marker lacks sensitivity and specificity. The recent controversial recommendation against the use of PSA in PC screening has highlighted the importance of developing new clinical biomarkers for PC (Gulati *et al.,* 2012). The diagnostic application of PSA has led to widespread over-diagnosis and subsequent over-treatment of clinically insignificant tumours (Simmons *et al.,* 2011; Stavridis *et al.,* 2010; Wolf *et al.,* 2010. This problem will very likely increase with rising life expectancy as more men become more elderly (Simmons *et al.,* 2011).

The use of biomarkers in PC detection and screening is crucial to the diagnosis and management of the disease. Currently, no biomarker on its own is able to overcome the issue with specificity with regards to PC detection, and in particular early detection of PC (Dayeti *et al.,* 2012). This

study attempted to address some of the concerns surrounding the need of novel biomarkers that can aid in the early detection of PC. In this study the aim was to identify miRNAs that are implicated in PC and to utilize these miRNAs as a tool to identify the presence of PC by determining their expression profiles in PC cells and comparing it to non-PC cells. Their expression levels in other cancer cell lines were also evaluated to determine specificity of these markers in PC.

There is considerable excitement surrounding the use of miRNAs in the biomarker field, which is well founded given their stability and relative ease of detection (Li *et al.,* 2007; Park *et al.,* 2009). Unlike mRNA, miRNAs show high stability in different types of biological samples, i.e.,formalin-fixed, paraffin-embedded clinical tissue, fresh snap-frozen material, plasma or serum, saliva and urine (Li *et al.,* 2007; Chen *et al.,* 2008; Mitchell *et al.,* 2008; Park *et al.,* 2009) . This high stability is due to their resistance to endogenous and exogenous RNAse activity, extreme temperatures, extremes of pH (pH 1 to 13), extended storage in frozen conditions, and reheated freeze-thaw cycles (Mitchell *et al.,* 2008; Gilad *et al.,* 2008). Resistance to such hazardous conditions has been attributed to

the encapsulation and association with protein complexes (Vickers *et al.,* 2011; Arroyo *et al.,* 2011; Valadi *et al.,* 2007; Collino *et al.,* 2010 Zernecke *et al.,* 2009). However, currently their use is limited by conflicting data between studies, due to the lack of standardisation in methodology and the lack of suitable reference genes for normalisation (Dayeti *et al.,* 2012). It is unlikely that any single miRNA will achieve the desired level of diagnostic or prognostic accuracy, as evidence is indicating, that increased levels of particular miRNAs can be associated with several different types of tumours (Vickers *et al.,* 2011; Arroyo *et al.,* 2011; Valadi *et al.,* 2007; Collino *et al.,* 2010 Zernecke *et al.,* 2009). However a panel of these biomarkers can serve as a signature of specific cancers.

There is evidence from *in vitro* and *in vivo* studies that alteration in miRNA function plays a role in prostate carcinogenesis (Berger *et al.,* 2014 and Murata *et al.,* 2012). MiRNA dysregulation influences a number of critical cellular processes involved in carcinogenesis, including but not limited to: stimulation of the cell cycle, avoidance of apoptosis, epithelial-mesenchymal transition and modulation of AR-mediated signalling (Ozen *et al.,* 2008; Clape *et al.,* 2009; Fu *et al.,* 2010; Schaefer *et al.,* 2010).

Further understanding of the functional importance of miRNA dysregulation may allow the development of novel diagnostic and therapeutic strategies involving miRNA augmentation or inhibition in the future. Circulating miRNA profiling in prostate cancer patients has been carried out by various investigators, but studies thus far have involved small numbers of patients, and a variety of methodologies, and have yielded heterogeneous results (Dayeti *et al.,* 2012). The potential to detect circulating miRNAs in serum and potentially, in urine, clearly exists (Chen *et al.,* 2008; Mitchell *et al.,* 2008; Li *et al.,* 2007; Park *et al.,* 2009). Furthermore, despite the variability of results in previous studies, the present study indicates that there are a number of miRNA targets that have both a demonstrable functional role in prostate pathogenesis and have demonstrated over- or under-expression when comparing normal prostate cells to prostate cancer cells (Coffey *et al.* 2002 and Ishiguro *et al.,* 2009). Further investigation of these miRNAs, either singularly or as part of a panel, in larger, prospective patient cohorts will help to define their potential role as diagnostic and prognostic biomarkers in the future. It is our belief that it will not be long before a miRNA-based diagnostic test is launched for prostate cancer, as has recently been done for pancreatic

adenocarcinoma (Szafranska-Schwarzbach *et al.,* 2011).

Initially, a rigorous bioinformatics approach was used to identify novel miRNA biomarkers associated with PC. The study identified 40 miRNA using an "in-house" generated bioinformatics pipeline compared to the 20-30 miRNAs already housed in public databases using alternative methods. The pipeline established "in-house" allowed for the identification of additional miRNAs that were not present in the public databases or showed a connection to PC previously. These 40 miRNA were narrowed down to 13 potential PC miRNA biomarkers based on the gene targets of these miRNAs that are linked to PC, or other processes that are important in the development of cancers. This prioritization was done by Gene Ontology analyses and Functional Annotation studies performed using DAVID. These included processes such as cell growth, proliferation, regulation of cell death and apoptosis. STRING analysis was also performed to identify possible protein-protein interactions based on experimentally validated data, amongst the 21 gene targets of the identified miRNAs.

Of these 13 miRNA, 8 miRNA were tested in this study. The remaining

miRNAs formed part of a parallel study within our group.

QRT-PCR analysis was performed to determine the expression of each individual miRNA in PC cell lines, as well as other selected cancer cell lines. Our results suggest that we could identify one specific miRNA biomarker that is unique to a specific cancer for example miRNA8 which was found to be specific for the LNCaP prostate cancer cell line. Additionally, we could also establish unique miRNA profiles or fingerprint for each of the cancers investigated in this study as indicated in Table 3.6. This finding could potentially lay the groundwork for future and additional studies within the group to refine a definitive fingerprint/miRNA profile for each cancer, and specifically for the early detection of PC. Although the study has managed to create unique profile to distinguish between the cancer cell lines investigated in this study, the study needs to be expanded to a larger cancer panel. Furthermore, these identified miRNAs should also be tested in both tissue and biological fluids such as urine. This will determine if these markers can be of clinical significance in the future for PC risk management, diagnosis, prognosis as well as response to treatment of PC.

CONCLUSIONS AND FUTURE WORK

A comparative bioinformatics analyses performed using an "in-house" pipeline to retrieve information from databases using slightly different parameters from those used in public databases, identified miRNAs that have not previously been described nor showed association to PC. In addition miRNAs were identified that were not found in miRNA databases.

In the present study the expression profiles of eight potential miRNA biomarkers for the possible detection of PC was determined, to distinguish PC from other cancers. These selected miRNAs were identified using different bioinformatics tools, and were confirmed to be novel miRNAs associated with PC. The pipeline has been optimized to identify miRNAs in a very short space of time. This allows for the identification of miRNAs as these databases are updated regularly, permitting a cost effective and less labor intensive miRNA identification pipeline possible.

In the panel of cancer cell lines that were investigated in this study it is

clearly shown that miRNA8 over expression is unique to LNCaP, giving some indication that in isolated incidences one miRNA may be sufficient for the detection of a condition. However, additional cell lines as well as patient samples must still be evaluated to establish the specificity of the over-expression of miRNA8 in PC. Our hypothesis is that a profile of miRNA expression in each cancer cell line should be unique.

The study suggests a unique profile of miRNAs for the cell lines investigated, and recognize the need for additional analysis on a larger set of cell lines to confirm their specificity. In addition, future experiments on the expression of these miRNAs in urine samples will be performed to determine their efficiency as potential markers for PC.

This study positively contributes to the field of molecular biomarkers that can aid in the detection of PC. While none of the recently identified biomarkers in literature can confidently replace PSA as the marker of choice for PC, we propose to evaluate the efficiency and the specificity of the identified miRNAs in urine samples by comparing it to PSA. The advantage of using miRNA biomarkers are that they are non-invasive as

they are present in urine, whereas the PSA test requires blood. From literature it is also suggested that miRNAs are stable in urine (Chen *et al.,* 2008; Mitchell *et al.,* 2008; Li *et al.,* 2007; Park *et al.,* 2009) and this negates the use of tissues and plasma to determine/verify the presence of the miRNA markers. Given the heterogeneity of cancers and in particular PC, the use of singular or a combination of miRNA as performed in the miRNA profiling in this study will potentially improve the predictive accuracy or prognostics as well as treatment outcomes in PC. The use of these markers will play an important role in screening for PC, however the fundamental goal remains to reduce the number of unnecessary biopsies performed and to limit the invasive procedures performed to differentiate between normal cells, benign tumours and aggressive PC.

APPENDIX A

Supplementary Information for Chapter 2

1. Scripts

2. Genelist

3. Pathway information

APPENDIX B

Supplementary Information for Chapter 3

1. Amplification Curves

2. Melting Curves

3. Descriptive Statistics

APPENDIX A

Supplementary Information for Chapter 2

1. Scripts

Extract Expression Data.pl

```
#############################################################
#!/usr/bin/perl                                             #
#Firdous Khan                                               #
#BRG LAB                                                    #
#Biotechnology Department                                   #
#University van wes-kaapland                                #
#Generates a tab delimited file of mean expression, gene id #
#and     probe name                                         #
#############################################################
use strict;
use warnings;


#Absolute path of infile2
my $infile2 = "path";
open (IF2, $infil2) || die "cannot open $infile2:$!";
open(OF2, ">mean.csv");
my @line2 = <IF2>;
for each my $line2(@line2)
{
        #Removes the new line character at the end of each line
        chomp $line2;
        #Slits each line into an array using tab meta-character
```

```
        my @col2 = split(/\t/, $line2);
        my $geneid = shift@col2;
        my $probe name = shift@col2;
        my $len =scalar@col2;
        my $total = 0;
        while (my $num = shift@col2)
        {
                $total += $num ;
        }
        my $mean = $total/$len;
        print OF2 $geneid."\t".$probe name."\t".$mean."\n";
}

close OF2
```


Data_Formatting.pl

```
#################################################################
#Firdous Khan                                                   #
#BRG LAB                                                        #
#Biotechnology Department                                       #
#University van wes-kaapland                                     #
#################################################################
#Converting csv-txt and removing unnecessary delimiters         #
#################################################################

sed 's/"//g' filename.csv > filename2.txt
sed 's/,/\t/g' filename2.txt > filename3
```

Repeat this for each dataset to ensure all the files
####in the same format

Annotation_graphs.R

###
#Firdous Khan #
#BRG LAB #
#Biotechnology Department #
#University van wes-kaapland #
###
#Initializing R and attaching input files #
#Plotting annotation graphs #
###

##Initialize R
R
library(lattice)
getwd()

#Ensure you are in the directory where the files are located#
input=read.table(file=file.choose(), header=T, sep="\t")
input path
eg: /home/user/annotation_data/catergory
attach(input)

###Plot graph

barchart(Heading1 ~ Heading2 | Heading3, data=input, col = "colour1", main="title1", ylab="label", xlab="%"))

barchart(Heading1 ~ Heading2 | Heading3, data=input, col = "colour2", main="title2", ylab="label", xlab="%")

barchart(Heading1 ~ Heading2 | Heading3, data=input, col = "colour3", main="title3", ylab="label", xlab="%")

png(filename="filename.png")

plot(filename)

dev.off()

rm(input)

Gene_expression_graph.R

###

#Firdous Khan #

#BRG LAB #

#Biotechnology Department #

#University van wes-kaapland #

###

##converting csv-txt and removing unnecessary delimiters #

#Initializing R and attaching input files #

#Plotting Gene expression graph #

###

##converting csv-txt and removing unnecessary delimiters

sed 's/"//g' filename1.csv > filename2.txt

```
sed 's/,/\t/g' filename2 > filename3

##Initialize R
R

library(lattice)
getwd()

#Ensure you are in the directory where the files are located#
input=read.table(file=file.choose(), header=T, sep="\t")
input path
eg:  /home/user/annotation_data/catergory
attach(input)

###Plot graph in R
barchart(Heading2 ~ Heading1 | Heading3, data=input, col = "colour",
horizontal=FALSE, scales=list(x=list(rot=90)),
index.cond=list(c(1,2,3,4,5,6)), main="title  ", xlab="label1", ylab="label")

barchart(Heading1 ~ Heading2 | Heading3, data=input, col="purple",
layout =c(1,2), main="Title", ylab="name1", xlab="name2")
Remove_duplicates.pl
```

##
 # Firdous Khan #
#BRG LAB #
#Biotechnology Department #
#University van wes-kaapland #
##
##converting csv-txt and removing unnecessary delimiters #
#Catalogue gene ids #
#Eliminate duplicates from files #
##

##converting csv-txt and removing unnecessary delimiters
sed 's/"//g' filename1.csv > filename2.txt
sed 's/,/\t/g' filename2 > filename3

Initialize the terminal and catalogue gene lists
cat [options] [filenames][-] > [filename]

##Eliminating duplicates

cat [option][file1][file2] | [sort-u] > [file3]

2. Genelist

Table 3: Tabular representation of the 551 target genes identified using *in silico* methods.

Target genes identified					
USP32	HBP1	DTHD1	SH3TC2	MBOAT1	FMR1
USP6	IFI30	FAM105A	SLAMF1	NAA50	GTF3C3
ZNF24	IL1RL1	FMR1	SLC30A7	ZNF26	HDHD1
RBFOX1	FOXE1	DUSP4	PRR16	ZNF562	FERMT2
SLC2A5	FOXP4	F2RL2	PTPN12	AFF4	SLC5A12
RARRES3	FA2H	SOCS6	ZNF148	CHD4	PDE1A
RNF170	LRRTM3	SRPK2	LPP	CSNK1G3	PRKCB
SH3TC2	OSBPL2	TULP4	AHCYL2	EIF4H	RBMS2
SLC31A2	PAFAH1B2	PGR	ZNF652	FAM46A	RFFL
SNRNP48	RAPGEF2	RSPO4	ONECUT2	INO80D	RIMKLA
UNC80	SLC12A2	SDHC	PKP2	LIG4	RUNDC3B
VAPA	SPTLC2	STAG2	PAPOLA	LMO7	SH3TC2
VAPB	SRSF11	ZNF673	PICALM	MIB1	ZNF589
ZBTB8A	TNPO1	PRKCI	PPFIA1	PAK7	BRWD1
ZNF780B	TXLNG	FAM86B1	PTPRC	PLSCR1	CHAC1
MAPK1IP1L	ITFG3	FAM86C1	PLXNA4	PPARG	CHORDC1
RAB27A	FOXN2	SAR1A	DAZ4	GNAL	SGIP1
RASAL2	FZD5	SBNO1	DCAF7	GPR180	SUB1
RBPJ	GADL1	SFMBT2	E2F7	GSDMC	TAPT1
RFK	UBQLN2	DCX	ENTPD1	HDAC9	TNFSF13

					B
SH3TC2	VEZF1	ING4	ZFP3	KCMF1	TOX
ACOX1	UQCRB	MPEG1	RAPGEF2	KCNT2	TSPAN2
CLMN	USP37	LPP	GLCCI1	KIAA2022	UQCRB
ENPP1	ZNF721	BRWD1	TNPO3	LMAN1	USP10
FAM20B	ANKRD52	C12orf73	URGCP	LOC100507161	UTRN
G3BP2	C1orf190	DKK2	ZFP106	ACVR1C	RRAD
HNF4G	CNTNAP1	CACNG8	ARL17A	ADNP	SH3TC2
MAGEA11	CREB1	CLN8	BACE1	AQP4	SLC5A12
PIAS2	CXADR	DDAH1	BCAT1	BCLAF1	SMAP1
PKNOX2	ELK1	TCP11L1	PPP1R9B	BMPR2	TLK2
SLC24A2	FOXC1	UBTF	RSU1	BTBD11	TMTC4
SLC39A14	SLC28A1	WIZ	CEPT1	CEP112	USP45
ENTPD4	ZNF704	IL5RA	FNDC3B	CLNK	USP46
ERBB4	COX11	IREB2	FYTTD1	CPEB2	VEZF1
EIF4H	ZNF559	LRRTM2	WDR33	SNRK	RIMS3
EP300	ZNF681	MED13	WWTR1	SYNPO2	SH3TC2
ESRRG	FAM46D	JAZF1	GUCY1A3	TAL1	SPECC1L
FMNL3	NKTR	LYSMD1	KCTD21	TSPAN2	SUPT3H
GABARAP	SLC17A8	NID2	KIAA0232	VANGL1	ZDHHC8
GCA	SLC25A21	PLEKHG7	KIF5C	ZBTB38	BCKDHB
MAVS	TOX	RABEP1	LYPLAL1	ZNF608	RNF41
MCART6	ACPL2	RLF	MGA	ANKIB1	RPRD2
MICAL3	ANKRD27	RRAS2	MLLT4	ARL17A	ACVR1C
MTPN	ATAD2B	SESTD1	MMS22L	C2orf63	AMD1

OAS2	BTG2	SMTNL2	NECAP1	CBX8	ARAP2
PARK2	C1orf96	SPARC	NEMF	LOC100294341	ASB8
PCNX	CCNT2	ADSS	NIPSNAP3B	LOC100506255	ATM
PDC	CEP97	AFAP1L2	OTX2	NHSL1	C5orf24
PURA	COL1A1	AFF1	PCDHAC2	PDK3	C5orf47
RAB5C	COL1A2	BFAR	RAB3B	SALL1	CALCR
RC3H1	COL3A1	BTBD7	RAB9B	TMX1	CCBL2
SIM2	COL4A1	C16orf52	RANBP2	UTP23	CCDC88A
SLC22A23	COL4A5	CNBP	RAP1A	ASB1	CGGBP1
SLC31A2	COL5A2	CNTNAP3B	RAPH1	C20orf177	CHML
SMG6	DYNLT1	CPLX3	RBBP6	INPP5A	COL6A6
SMU1	EIF3J	DDX17	RSBN1	KCND1	DEXI
SUFU	FOXJ2	DIS3	RTF1	KSR2	ENOPH1
SYK	GOLGA7	DOK6	SESTD1	LRRC4	FAM120A
EPC2	ADAM22	MGP	XPO1	PAM	AAK1
FASTKD2	AMMECR1L	MR1	YWHAZ	PAN3	AFF3
GBP3	ASF1A	MTX3	ZFX	PCDHB16	ATP2B4
HDX	ASPH	NOVA1	ZNF236	PRDM10	BSN
IGSF10	BET1	NSL1	ZNF492	PRKACB	CDK6
LAMP2	C22orf39	PAK7	DDAH1	RAB11A	COL4A4
LSM14B	C9orf123	PI15	DDX18	RAB8B	DAZ1
MIER3	CCDC34	PSMA8	DENND1B	RNF144A	DAZ2
MILR1	CREBZF	NRCAM	CYTIP		

3. Pathway information

Table 2: Tabular representation of significant pathways associated with miRNA1-8 *silico* methods

KEGG pathway	p-value	#genes	#miRNAs
ECM-receptor interaction	<1e-16	14	1
Prion diseases	<1e-16	3	1
Amoebiasis	<1e-16	14	1
Protein digestion and absorption	1.187939E-014	15	1
Focal adhesion	2.055885E-010	29	2
Colorectal cancer	5.361692E-010	19	4
PI3K-Akt signaling pathway	1.408314E-008	51	3
Endometrial cancer	9.908088E-008	13	3
Glycosaminoglycan biosynthesis - chondroitin sulfate	1.737486E-007	2	2
Wnt signaling pathway	0.000010579	29	3
Pathways in cancer	0.0001553293	61	4
Glycosaminoglycan biosynthesis - heparan sulfate / hepari	0.0001909948	6	2
Hypertrophic cardiomyopathy (HCM)	0.0002814309	12	2
Dilated cardiomyopathy	0.0004519795	12	2
Adherens junction	0.000476915	13	2
Hedgehog signaling pathway	0.0005539027	14	3
Neurotrophin signaling pathway	0.000855797	26	4
Prostate cancer	0.0009389909	13	2
Basal cell carcinoma	0.002328241	12	3
Axon guidance	0.004194514	16	2
Mucin type O-Glycan biosynthesis	0.004534804	3	1
Endocytosis	0.00574708	8	1
Thyroid cancer	0.006504963	3	2
Ubiquitin mediated proteolysis	0.007459014	26	3
Small cell lung cancer	0.009583545	6	1
TGF-beta signaling pathway	0.009684528	5	1
Regulation of actin cytoskeleton	0.01302628	12	1
Melanoma	0.01380284	9	2
Glutamatergic synapse	0.01984405	14	2
Renal cell carcinoma	0.02505617	13	2
Toxoplasmosis	0.03384567	7	1
HTLV-I infection	0.03759023	23	2
Other types of O-glycan biosynthesis	0.04270389	3	2

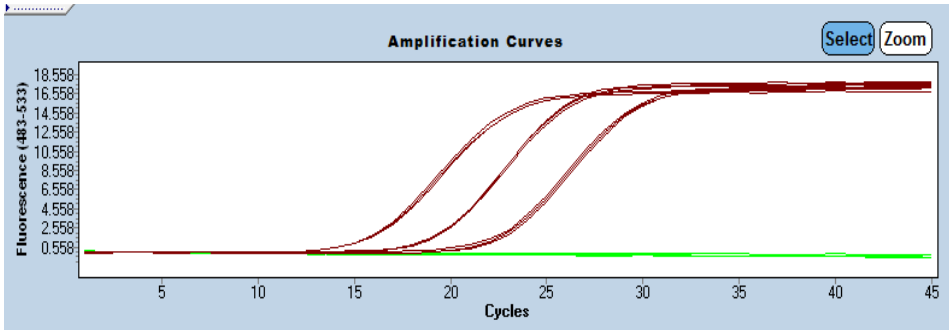
Table 3: Tabular representation of additional identified pathways associated to miRNA1-8

KEGG pathway	p-value	#genes	#miRNAs
ErbB signaling pathway	0.05102382	7	1
p53 signaling pathway	0.05323238	8	2
mRNA surveillance pathway	0.05493907	14	2
Transcriptional misregulation in cancer	0.06032142	10	2
Measles	0.06556488	7	1
Influenza A	0.0683851	8	1
Valine, leucine and isoleucine biosynthesis	0.07245746	1	2
Arrhythmogenic right ventricular cardiomyopathy (ARVC)	0.08613729	6	1
Adipocytokine signaling pathway	0.09080646	11	2
Cytokine-cytokine receptor interaction	0.1103693	4	1
Hepatitis B	0.1216237	11	1
N-Glycan biosynthesis	0.1407615	2	1
Shigellosis	0.1483155	10	2
Hepatitis C	0.1505176	6	1
Long-term potentiation	0.1707405	4	1
Fc gamma R-mediated phagocytosis	0.174731	3	1
T cell receptor signaling pathway	0.1762456	10	1
Cholinergic synapse	0.1771551	12	1
Glycosaminoglycan biosynthesis - keratan sulfate	0.1906681	1	1
Melanogenesis	0.2044662	8	1
Acute myeloid leukemia	0.2193944	4	1
RNA degradation	0.2241099	8	1
Non-small cell lung cancer	0.2306316	5	1
Drug metabolism - cytochrome P450	0.2575857	1	1
Dopaminergic synapse	0.259873	10	1
Spliceosome	0.265297	6	1
Epstein-Barr virus infection	0.2709317	7	1
Dilated cardiomyopathy	0.3149708	13	1
Pantothenate and CoA biosynthesis	0.319202	2	1
Tuberculosis	0.3267261	14	1
RIG-I-like receptor signaling pathway	0.3275958	7	1
HIF-1 signaling pathway	0.3532609	9	1
B cell receptor signaling pathway	0.3727899	8	1
Insulin signaling pathway	0.4067456	10	1
Protein processing in endoplasmic reticulum	0.4514956	13	1
Dorso-ventral axis formation	0.464945	3	1
Valine, leucine and isoleucine degradation	0.4992077	4	1
Bacterial invasion of epithelial cells	0.5384656	10	2
Lysine biosynthesis	0.5471215	1	1
Herpes simplex infection	0.5661742	8	1
Fatty acid elongation	0.5756327	1	1
Oocyte meiosis	0.5758602	10	1
Retrograde endocannabinoid signaling	0.5802484	8	1
Calcium signaling pathway	0.6007089	13	1
Viral carcinogenesis	0.6146964	10	1
Nicotine addiction	0.6633905	3	1
Mineral absorption	0.6807361	4	1
Toll-like receptor signaling pathway	0.6830615	4	1
Homologous recombination	0.7044528	2	1
Folate biosynthesis	0.750011	2	1
Malaria	0.796024	3	1
Metabolism of xenobiotics by cytochrome P450	0.8069897	3	1
Viral myocarditis	0.8287709	4	1
Glycerolipid metabolism	0.8361433	3	1
One carbon pool by folate	0.8499083	2	1
Bile secretion	0.8583768	3	1
Gap junction	0.865365	7	1
Tyrosine metabolism	0.8672197	2	1
Gastric acid secretion	0.8715899	7	1
Purine metabolism	0.9279241	5	1
Selenocompound metabolism	0.9685885	2	1
Thiamine metabolism	0.9784377	1	1

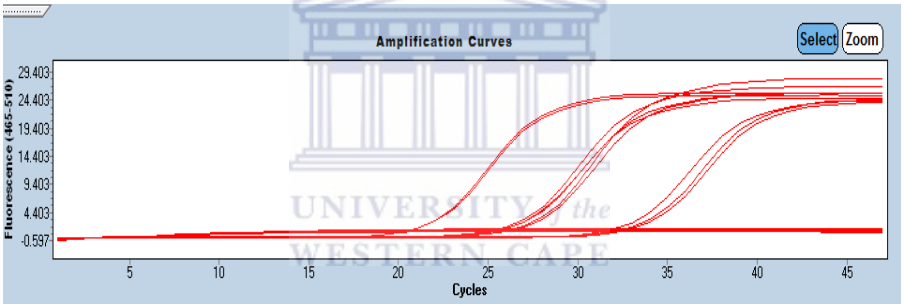
APPENDIX B

Supplementary Information for Chapter 3

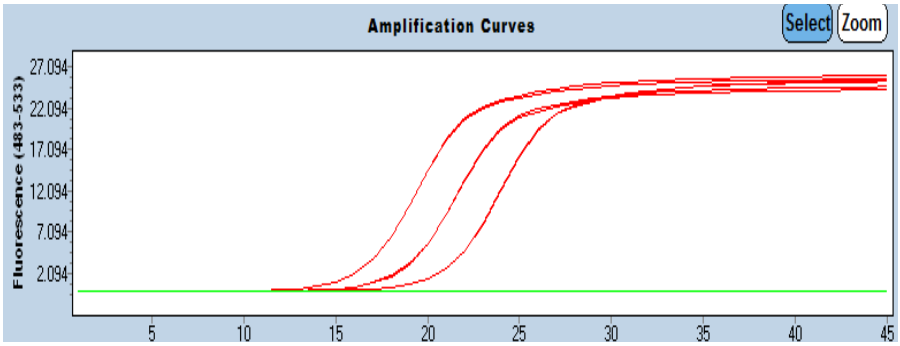
1. Amplification Curves



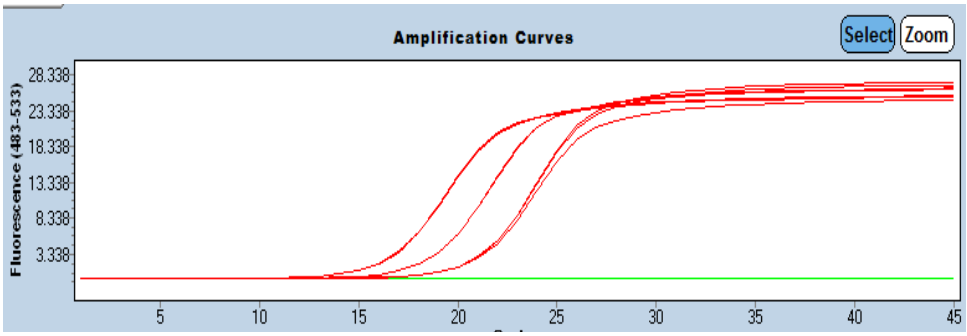
MiRNA 17A – Amplification curve



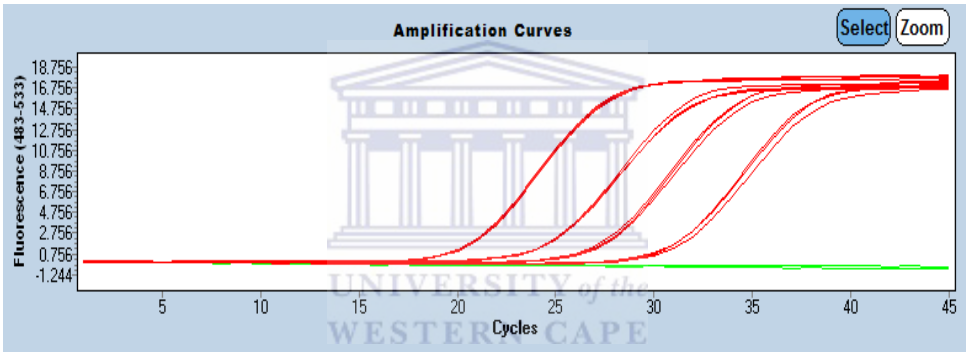
MiRNA 191A – Amplification curve



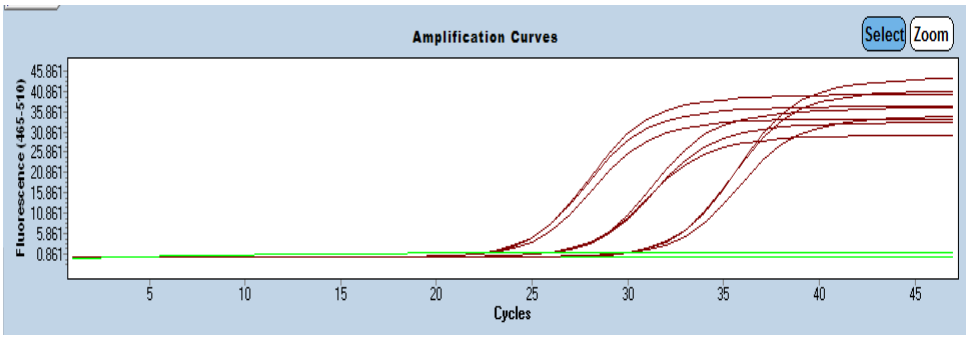
MiRNA 1 – Amplification curve



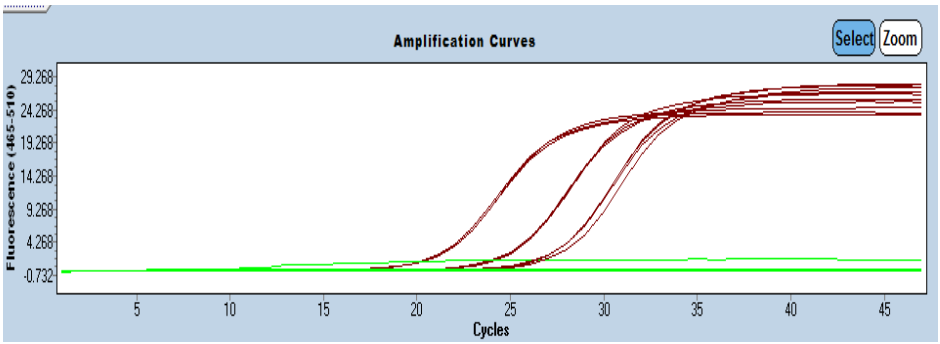
MiRNA 2 – Amplification curve



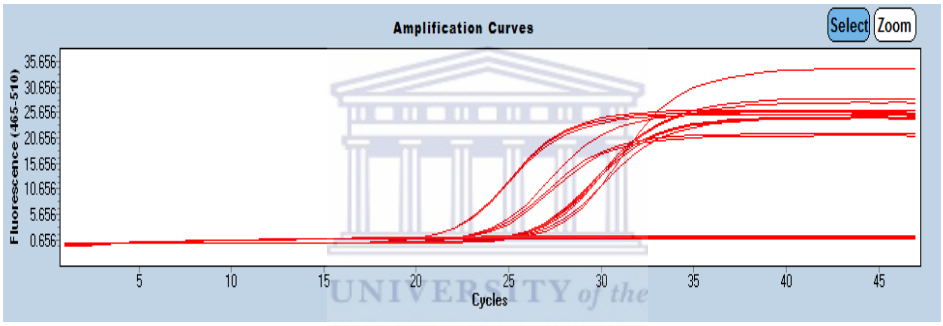
MiRNA 3 – Amplification curve



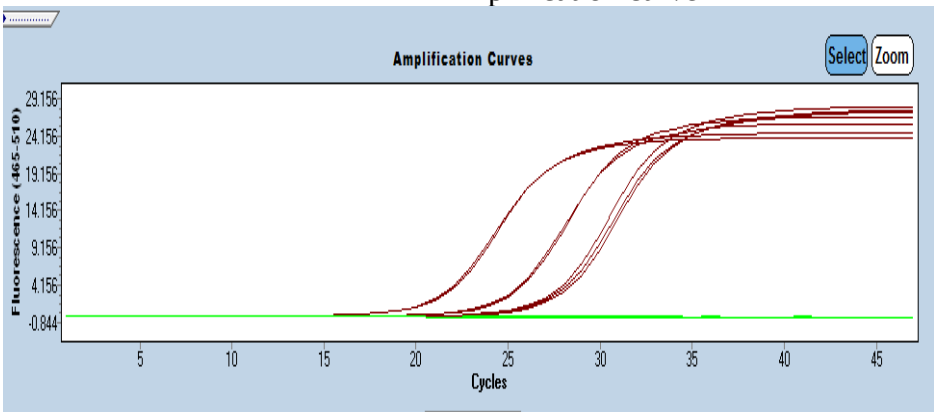
MiRNA 4 – Amplification curve



MiRNA 6 – Amplification curve

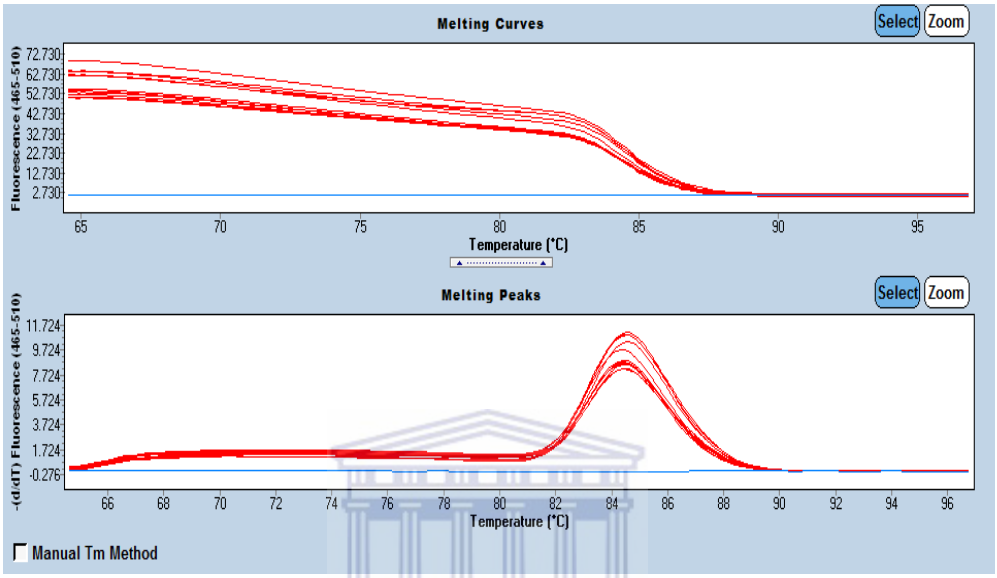


MiRNA 7 – Amplification curve

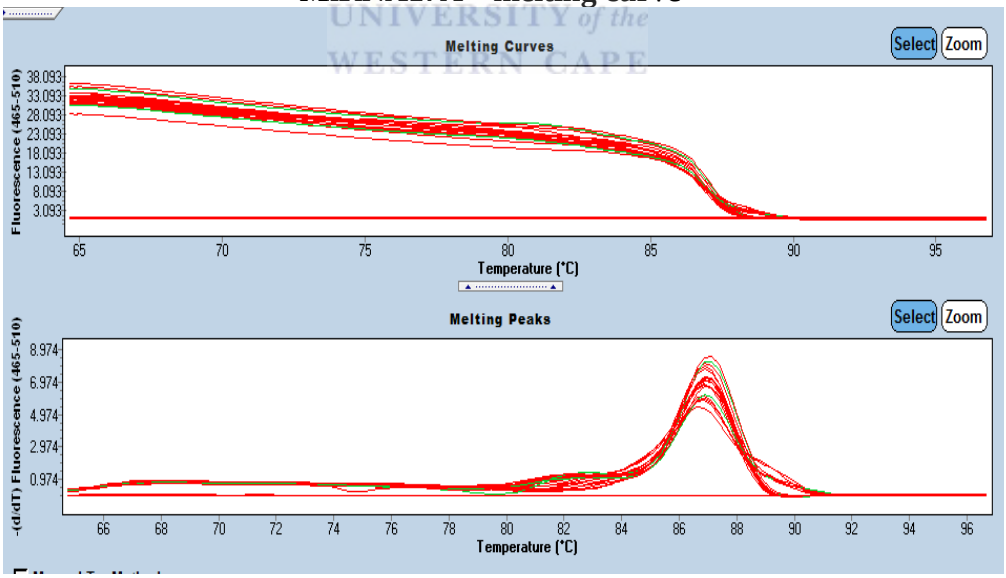


MiRNA 8 – Amplification curve

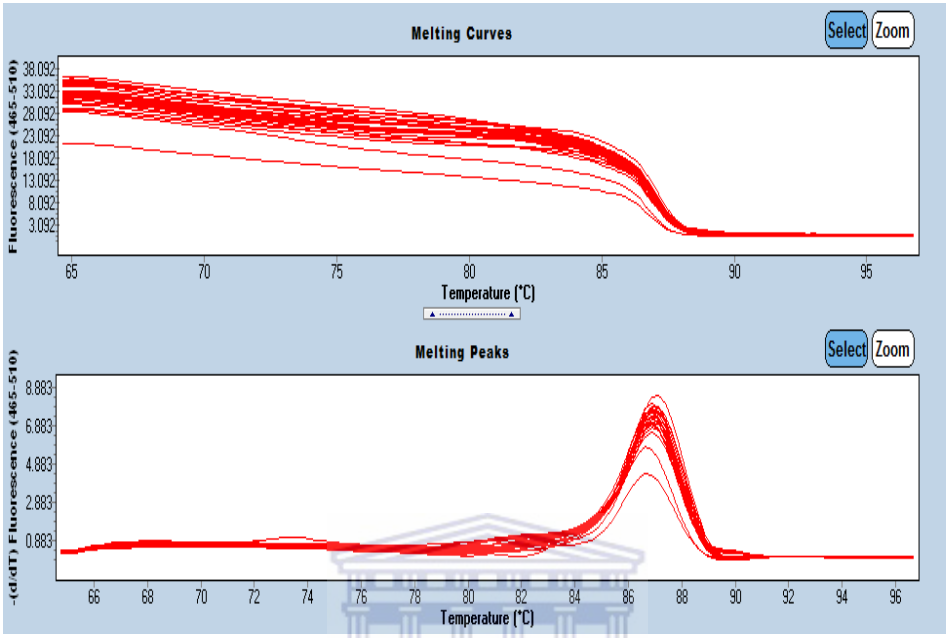
2. Melting Curves



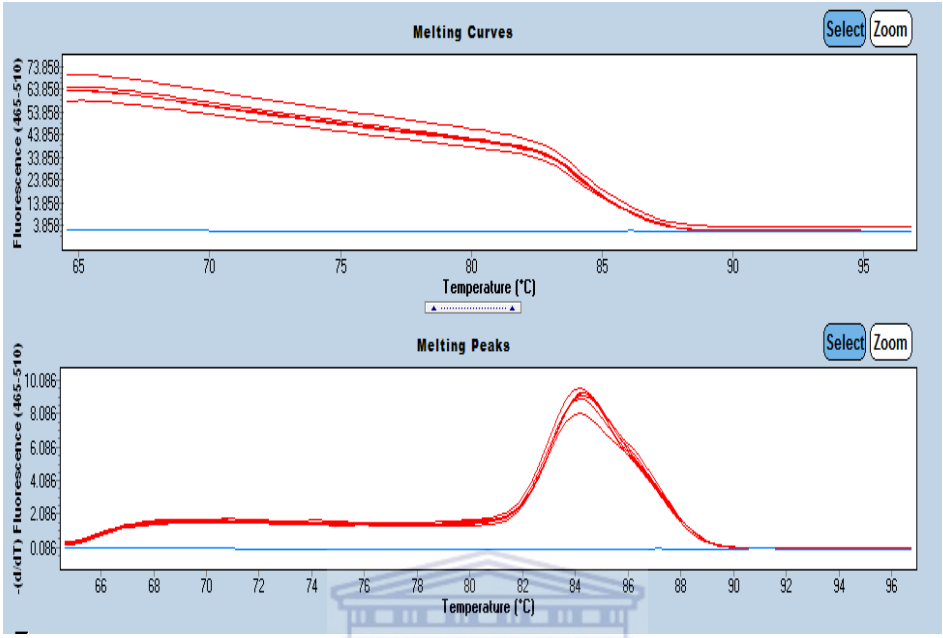
MiRNA17A – melting curve



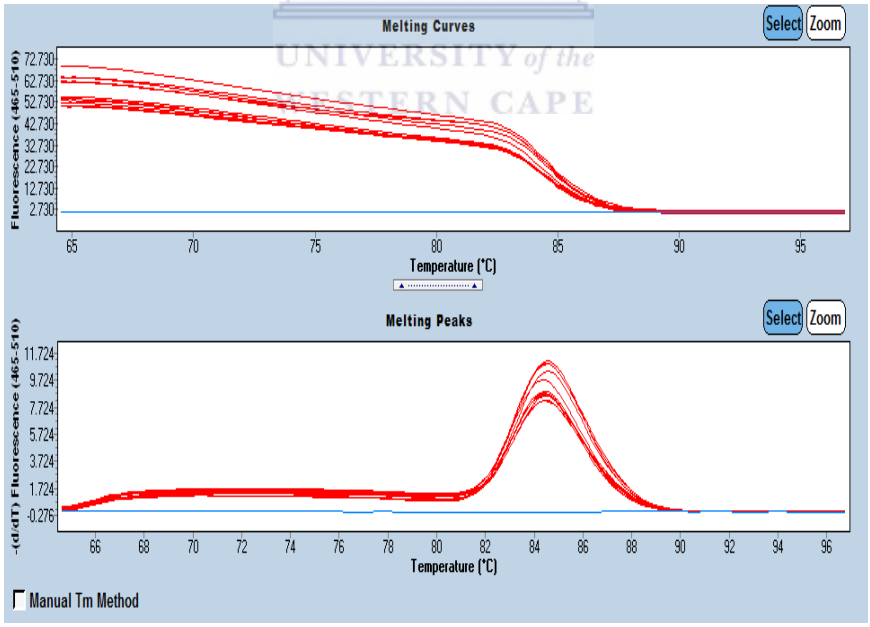
MiRNA 191A- melting peaks and curve



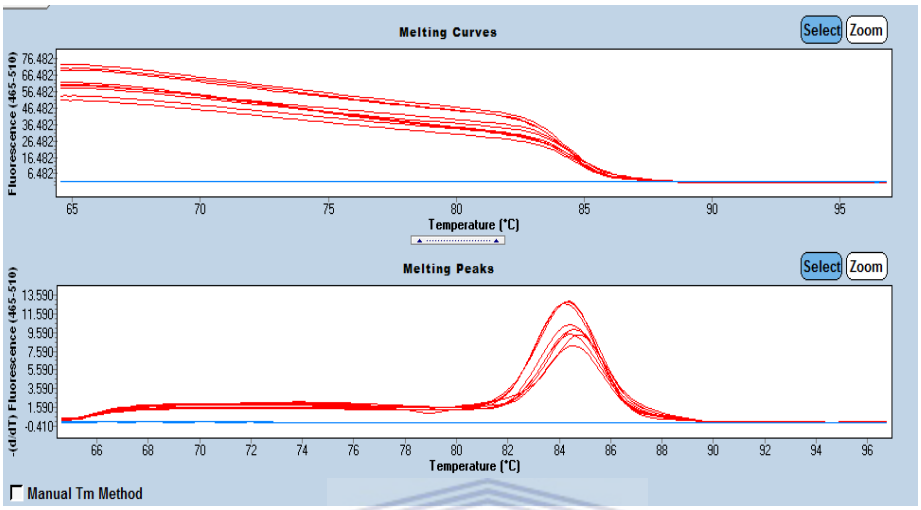
MiRNA 1 – Melting curve



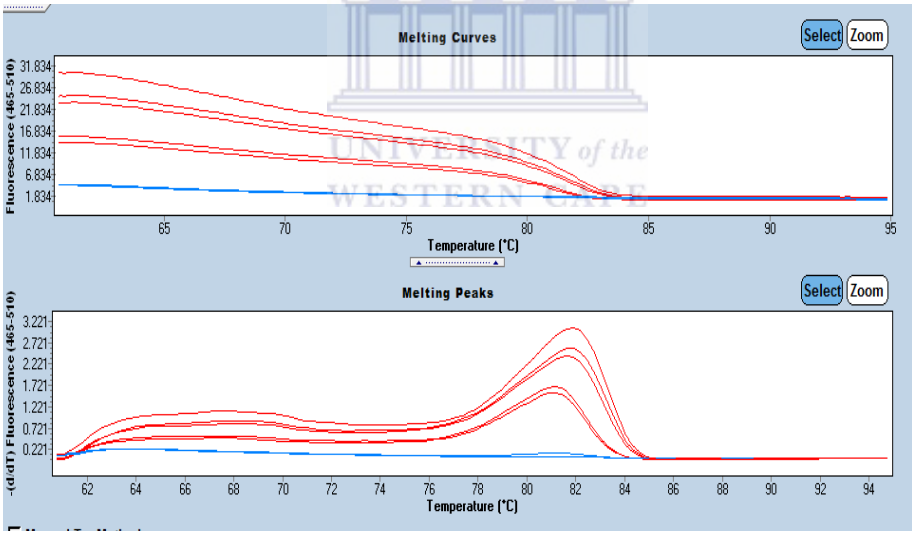
MiRNA 2- Melting curve



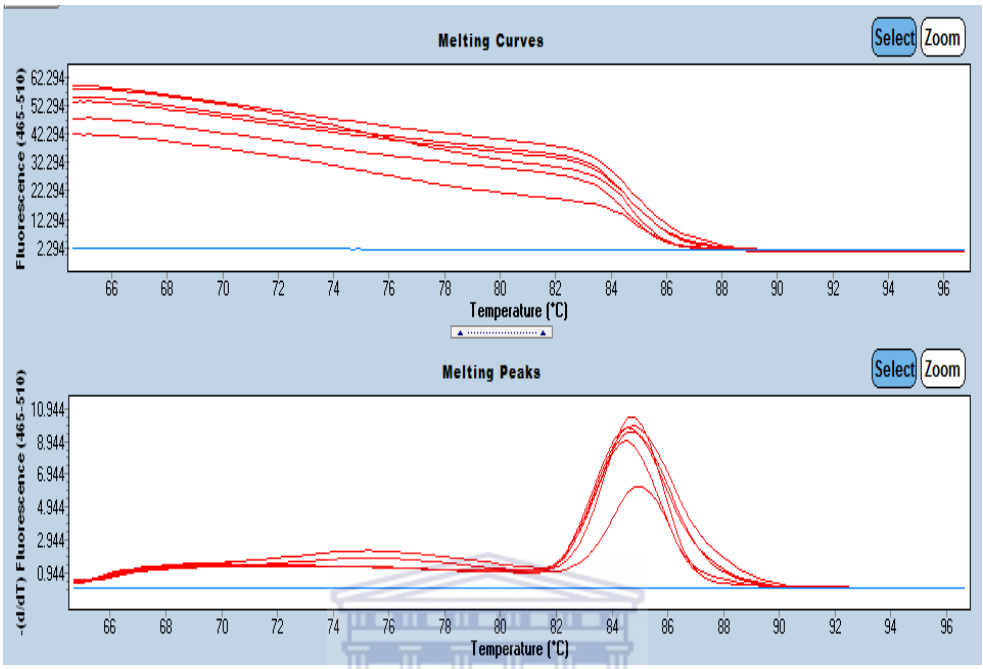
MiRNA 3 – Melting curve



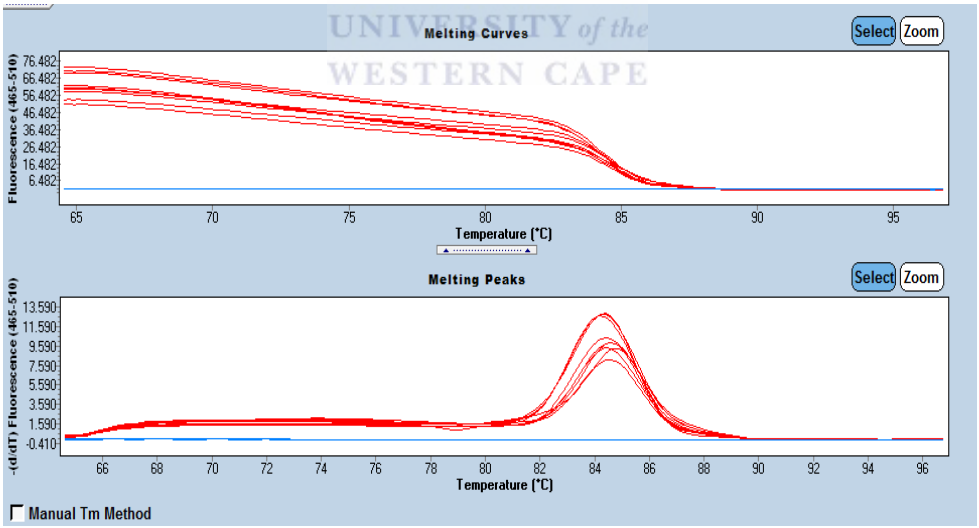
MiRNA 4 – Melting curve



MiRNA 6 – Melting Curve



MiRNA 7 – Melting Curve



MiRNA 8 – Melting Curve

UNIVERSITY of the
WESTERN CAPE

3. Descriptive Statistics

Table 3.7: Descriptive statistics used to determine the differential expression of miR1-8 in BPH1 in comparison to PNT2C2 cell lines.

control	[reference gene]	[target gene 1]	[target gene 2]	[target gene 3]	[target gene 4]	[target gene 5]	[target gene 6]	[target gene 7]	[target gene 8]	[target gene 9]	[target gene 10]	[target gene 11]
	GAPDH	miR-1	miR-2	miR-3	miR-4	miR-5	miR-6	miR-7	miR-8	internal control	HKG2	PTEN
		☐ *	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
n	3	3	3	3	3	3	3	3	3	3	3	3
mean	18.19	34.67	34.67	27.74	31.97	28.17	35.00	29.20	31.21	35.00	28.07	27.07
standard error	0.04	0.33	0.32	0.06	0.07	0.03	0.00	0.06	0.12	0.00	0.07	0.07
CV [%]	0.39	1.67	1.59	0.37	0.36	0.20	0.00	0.34	0.67	0.00	0.41	0.43
sample(s)	[reference gene]	[target gene 1]	[target gene 2]	[target gene 3]	[target gene 4]	[target gene 5]	[target gene 6]	[target gene 7]	[target gene 8]	[target gene 9]	[target gene 10]	[target gene 11]
	GAPDH	miR-1	miR-2	miR-3	miR-4	miR-5	miR-6	miR-7	miR-8	internal control	HKG2	PTEN
n	3	3	3	3	3	3	3	3	3	3	3	3
mean	18.05	34.67	34.47	25.24	31.93	26.61	35.00	29.03	31.84	35.00	27.93	27.10
standard error	0.11	0.33	0.27	0.21	0.03	0.18	0.00	0.03	0.19	0.00	0.07	0.11
CV [%]	1.01	1.67	1.37	1.42	0.18	1.16	0.00	0.20	1.05	0.00	0.41	0.70
	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
E(target)^CP	1.098	1.000	1.149	5.631	1.023	2.949	1.000	1.122	0.643	1.000	1.097	0.977
	1.098											
Normalization Factor **	1.098											
	1.000	0.911	1.046	5.129	0.932	2.686	0.911	1.022	0.586	0.911	0.999	0.890
	[reference gene]	[target gene 1]	[target gene 2]	[target gene 3]	[target gene 4]	[target gene 5]	[target gene 6]	[target gene 7]	[target gene 8]	[target gene 9]	[target gene 10]	[target gene 11]
Exression ratio(s):	GAPDH	miR-1	miR-2	miR-3	miR-4	miR-5	miR-6	miR-7	miR-8	internal control	HKG2	PTEN
Significant (randomization test)												
Target gene is UP-regulated by the factor:			1.046	5.129		2.686		1.022				
Target gene is DOWN-regulated by the factor:		-1.098			-1.073		-1.098		-1.707	-1.098	-1.001	-1.124
Absolute gene regulation:		0.911	1.046	5.129	0.932	2.686	0.911	1.022	0.586	0.911	0.999	0.890
Absolute gene regulation (standard error):		*±0.29945*	*±0.31435*	*±0.85987*	*±0.08637*	*±0.39462*	*±0.07006*	*±0.09174*	*±0.10271*	*±0.07006*	*±0.10083*	*±0.10486*
		-1.098	1.046	5.129	-1.073	2.686	-1.098	1.022	-1.707	-1.098	-1.001	-1.124
Absolute gene regulation (2-log):		-0.135	0.065	2.359	-0.101	1.425	-0.135	0.032	-0.771	-0.135	-0.001	-0.168
Absolute gene regulation (2-log standard error):		*±1.74*	*±1.67*	*±0.218*	*±3.533*	*±1.341*	*±3.835*	*±3.446*	*±3.283*	*±3.835*	*±3.31*	*±3.253*

Table 3.8: Descriptive statistics used to determine the differential expression of miR1-8 in LNCaP in comparison to PNT2C2 cell lines.

control	[reference gene]	[target gene 1]	[target gene 2]	[target gene 3]	[target gene 4]	[target gene 5]	[target gene 6]	[target gene 7]	[target gene 8]	[target gene 9]	[target gene 10]	[target gene 11]
	GAPDH	miR-1	miR-2	miR-3	miR-4	miR-5	miR-6	miR-7	miR-8	internal control	HKG2	PTEN
		☐ *	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
n	3	3	3	3	3	3	3	3	3	3	3	3
mean	18.19	34.67	34.67	27.74	31.97	28.17	35.00	29.20	31.21	35.00	28.07	27.07
standard error	0.04	0.33	0.32	0.06	0.07	0.03	0.00	0.06	0.12	0.00	0.07	0.07
CV [%]	0.39	1.67	1.59	0.37	0.36	0.20	0.00	0.34	0.67	0.00	0.41	0.43
sample(s)	[reference gene]	[target gene 1]	[target gene 2]	[target gene 3]	[target gene 4]	[target gene 5]	[target gene 6]	[target gene 7]	[target gene 8]	[target gene 9]	[target gene 10]	[target gene 11]
	GAPDH	miR-1	miR-2	miR-3	miR-4	miR-5	miR-6	miR-7	miR-8	internal control	HKG2	PTEN
n	3	3	3	3	3	3	3	3	3	3	3	3
mean	18.05	35.00	34.57	26.30	32.03	25.94	35.00	29.07	28.07	35.00	27.93	27.10
standard error	0.11	0.17	0.09	0.21	0.09	0.04	0.00	0.07	0.12	0.00	0.07	0.11
CV [%]	1.01	0.86	0.44	1.37	0.48	0.28	0.00	0.40	0.74	0.00	0.41	0.70
	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
E(target)^CP	1.098	0.798	1.072	2.707	0.955	4.681	1.000	1.097	8.815	1.000	1.097	0.977
	1.098											
Normalization Factor **	1.098											
	1.000	0.727	0.976	2.466	0.870	4.263	0.911	0.999	8.030	0.911	0.999	0.890
	[reference gene]	[target gene 1]	[target gene 2]	[target gene 3]	[target gene 4]	[target gene 5]	[target gene 6]	[target gene 7]	[target gene 8]	[target gene 9]	[target gene 10]	[target gene 11]
Exression ratio(s):	GAPDH	miR-1	miR-2	miR-3	miR-4	miR-5	miR-6	miR-7	miR-8	internal control	HKG2	PTEN
Significant (randomization test)												
Target gene is UP-regulated by the factor:				2.466		4.263			8.030			
Target gene is DOWN-regulated by the factor:		-1.376	-1.024		-1.150		-1.098	-1.001		-1.098	-1.001	-1.124
Absolute gene regulation:		0.727	0.976	2.466	0.870	4.263	0.911	0.999	8.030	0.911	0.999	0.890
Absolute gene regulation (standard error):		*±0.19332*	*±0.23558*	*±0.41557*	*±0.09443*	*±0.36382*	*±0.07006*	*±0.09816*	*±1.13338*	*±0.07006*	*±0.10083*	*±0.10486*
		-1.376	-1.024	2.466	-1.150	4.263	-1.098	-1.001	8.030	-1.098	-1.001	-1.124
Absolute gene regulation (2-log):		-0.461	-0.035	1.302	-0.201	2.092	-0.135	-0.001	3.005	-0.135	-0.001	-0.168
Absolute gene regulation (2-log standard error):		*±2.371*	*±2.086*	*±1.267*	*±3.405*	*±1.459*	*±3.835*	*±3.349*	*±0.181*	*±3.835*	*±3.31*	*±3.253*
		0.193320000	0.235580000	0.415570000	0.094430000	0.363820000	0.070060000	0.098160000	1.133380000	0.070060000	0.100830000	0.104860000

Table 3.9: Descriptive statistics used to determine the differential expression of miR1-8 in MCF7 in comparison to MCF12A cell line.

control	[reference gene]	[target gene 1]	[target gene 2]	[target gene 3]	[target gene 4]	[target gene 5]	[target gene 6]	[target gene 7]	[target gene 8]	[target gene 9]	[target gene 10]	[target gene 11]
	GAPDH	miR-1	miR-2	miR-3	miR-4	miR-5	miR-6	miR-7	miR-8	internal control	HKG2	PTEN
		☐ *	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
n	3	3	3	3	3	3	3	3	3	3	3	3
mean	18.19	34.67	34.67	30.00	31.97	26.00	35.00	35.00	33.13	35.00	28.07	27.07
standard error	0.04	0.33	0.32	0.15	0.07	0.00	0.00	0.00	0.13	0.00	0.07	0.07
CV [%]	0.39	1.67	1.59	0.88	0.36	0.00	0.00	0.00	0.70	0.00	0.41	0.43
sample(s)	[reference gene]	[target gene 1]	[target gene 2]	[target gene 3]	[target gene 4]	[target gene 5]	[target gene 6]	[target gene 7]	[target gene 8]	[target gene 9]	[target gene 10]	[target gene 11]
	GAPDH	miR-1	miR-2	miR-3	miR-4	miR-5	miR-6	miR-7	miR-8	internal control	HKG2	PTEN
n	3	3	3	3	3	3	3	3	3	3	3	3
mean	18.05	34.67	34.47	30.07	31.93	26.23	35.00	32.37	33.10	35.00	27.93	26.50
standard error	0.11	0.33	0.27	0.07	0.03	0.09	0.00	0.09	0.15	0.00	0.07	0.06
CV [%]	1.01	1.67	1.37	0.38	0.18	0.58	0.00	0.47	0.80	0.00	0.41	0.38
	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
E(target)^CP	1.098	1.000	1.149	0.955	1.023	0.851	1.000	6.205	1.023	1.000	1.097	1.481
	1.098											
Normalization Factor **	1.098											
	1.000	0.911	1.046	0.870	0.932	0.775	0.911	5.652	0.932	0.911	0.999	1.349
	[reference gene]	[target gene 1]	[target gene 2]	[target gene 3]	[target gene 4]	[target gene 5]	[target gene 6]	[target gene 7]	[target gene 8]	[target gene 9]	[target gene 10]	[target gene 11]
Exression ratio(s):	GAPDH	miR-1	miR-2	miR-3	miR-4	miR-5	miR-6	miR-7	miR-8	internal control	HKG2	PTEN
Significant (randomization test)												
Target gene is UP-regulated by the factor:			1.046					5.652				1.349
Target gene is DOWN-regulated by the factor:		-1.098		-1.150	-1.073	-1.291	-1.098		-1.073	-1.098	-1.001	
Absolute gene regulation:		0.911	1.046	0.870	0.932	0.775	0.911	5.652	0.932	0.911	0.999	1.349
Absolute gene regulation (standard error):		*±0.29945*	*±0.31435*	*±0.12071*	*±0.08637*	*±0.07613*	*±0.07006*	*±0.55525*	*±0.14934*	*±0.07006*	*±0.10083*	*±0.13254*
		-1.098	1.046	-1.150	-1.073	-1.291	-1.098	5.652	-1.073	-1.098	-1.001	1.349
Absolute gene regulation (2-log):		-0.135	0.065	-0.201	-0.101	-0.368	-0.135	2.499	-0.101	-0.135	-0.001	0.432
Absolute gene regulation (2-log standard error):		*±1.74*	*±1.67*	*±3.05*	*±3.533*	*±3.715*	*±3.835*	*±0.849*	*±2.743*	*±3.835*	*±3.31*	*±2.915*

Table 3.10: Descriptive statistics used to determine the differential expression of miR1-8 in A549 in comparison to KMST1 cell line.

control	[reference gene]	[target gene 1]	[target gene 2]	[target gene 3]	[target gene 4]	[target gene 5]	[target gene 6]	[target gene 7]	[target gene 8]	[target gene 9]	[target gene 10]	[target gene 11]
	GAPDH	miR-1	miR-2	miR-3	miR-4	miR-5	miR-6	miR-7	miR-8	internal control	HKG2	PTEN
		☐ *	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
n	3	3	3	3	3	3	3	3	3	3	3	3
mean	18.19	34.67	32.47	30.33	31.97	26.00	35.00	35.00	33.13	35.00	28.07	27.07
standard error	0.04	0.33	0.07	0.15	0.07	0.00	0.00	0.00	0.13	0.00	0.07	0.07
CV [%]	0.39	1.67	0.36	0.83	0.36	0.00	0.00	0.00	0.70	0.00	0.41	0.43
sample(s)	[reference gene]	[target gene 1]	[target gene 2]	[target gene 3]	[target gene 4]	[target gene 5]	[target gene 6]	[target gene 7]	[target gene 8]	[target gene 9]	[target gene 10]	[target gene 11]
	GAPDH	miR-1	miR-2	miR-3	miR-4	miR-5	miR-6	miR-7	miR-8	internal control	HKG2	PTEN
n	3	3	3	3	3	3	3	3	3	3	3	3
mean	18.05	34.67	30.80	28.87	31.93	26.10	35.00	33.00	33.10	35.00	27.93	26.50
standard error	0.11	0.33	0.23	0.07	0.03	0.06	0.00	0.12	0.15	0.00	0.07	0.06
CV [%]	1.01	1.67	1.30	0.40	0.18	0.38	0.00	0.61	0.80	0.00	0.41	0.38
	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
E(target)^CP	1.098	1.000	3.175	2.764	1.023	0.933	1.000	4.000	1.023	1.000	1.097	1.481
	1.098											
Normalization Factor **	1.098											
	1.000	0.911	2.892	2.517	0.932	0.850	0.911	3.643	0.932	0.911	0.999	1.348
	[reference gene]	[target gene 1]	[target gene 2]	[target gene 3]	[target gene 4]	[target gene 5]	[target gene 6]	[target gene 7]	[target gene 8]	[target gene 9]	[target gene 10]	[target gene 11]
Exression ratio(s):	GAPDH	miR-1	miR-2	miR-3	miR-4	miR-5	miR-6	miR-7	miR-8	internal control	HKG2	PTEN
Significant (randomization test)												
Target gene is UP-regulated by the factor:			2.892	2.517				3.643				1.349
Target gene is DOWN-regulated by the factor:		-1.098			-1.073	-1.177	-1.098		-1.073	-1.098	-1.001	
Absolute gene regulation:		0.911	2.892	2.517	0.932	0.850	0.911	3.643	0.932	0.911	0.999	1.349
Absolute gene regulation (standard error):		*±0.29945*	*±0.53067*	*±0.33957*	*±0.08637*	*±0.07369*	*±0.07006*	*±0.40444*	*±0.14934*	*±0.07006*	*±0.10083*	*±0.13254*
		-1.098	2.892	2.517	-1.073	-1.177	-1.098	3.643	-1.073	-1.098	-1.001	1.349
Absolute gene regulation (2-log):		-0.135	1.532	1.332	-0.101	-0.235	-0.135	1.865	-0.101	-0.135	-0.001	0.432
Absolute gene regulation (2-log standard error):		*±1.74*	*±0.914*	*±1.558*	*±3.533*	*±3.762*	*±3.835*	*±1.306*	*±2.743*	*±3.835*	*±3.31*	*±2.915*

Table 3.11: Descriptive statistics used to determine the differential expression of miR1-8 in HT29 in comparison to KMST1 cell line.

control	[reference gene]	[target gene 1]	[target gene 2]	[target gene 3]	[target gene 4]	[target gene 5]	[target gene 6]	[target gene 7]	[target gene 8]	[target gene 9]	[target gene 10]	[target gene 11]
	GAPDH	miR-1	miR-2	miR-3	miR-4	miR-5	miR-6	miR-7	miR-8	internal control	HKG2	PTEN
		☐ *	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
n	3	3	3	3	3	3	3	3	3	3	3	3
mean	18.19	34.67	32.00	30.00	31.97	26.00	35.00	35.00	33.13	35.00	28.07	27.07
standard error	0.04	0.33	0.00	0.15	0.07	0.00	0.00	0.00	0.13	0.00	0.07	0.07
CV [%]	0.39	1.67	0.00	0.88	0.36	0.00	0.00	0.00	0.70	0.00	0.41	0.43
sample(s)	[reference gene]	[target gene 1]	[target gene 2]	[target gene 3]	[target gene 4]	[target gene 5]	[target gene 6]	[target gene 7]	[target gene 8]	[target gene 9]	[target gene 10]	[target gene 11]
	GAPDH	miR-1	miR-2	miR-3	miR-4	miR-5	miR-6	miR-7	miR-8	internal control	HKG2	PTEN
n	3	3	3	3	3	3	3	3	3	3	3	3
mean	18.05	34.67	32.00	28.87	30.00	26.10	35.00	33.00	33.10	35.00	27.93	25.70
standard error	0.11	0.33	0.00	0.07	0.12	0.06	0.00	0.12	0.15	0.00	0.07	0.06
CV [%]	1.01	1.67	0.00	0.40	0.67	0.38	0.00	0.61	0.80	0.00	0.41	0.39
	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
E(target)^CP	1.098	1.000	1.000	2.194	3.909	0.933	1.000	4.000	1.023	1.000	1.097	2.579
	1.000											
Normalization Factor **	1.098											
	1.000	0.911	0.911	1.998	3.560	0.850	0.911	3.643	0.932	0.911	0.999	2.348
	[reference gene]	[target gene 1]	[target gene 2]	[target gene 3]	[target gene 4]	[target gene 5]	[target gene 6]	[target gene 7]	[target gene 8]	[target gene 9]	[target gene 10]	[target gene 11]
Exression ratio(s):	GAPDH	miR-1	miR-2	miR-3	miR-4	miR-5	miR-6	miR-7	miR-8	internal control	HKG2	PTEN
Significant (randomization test)												
Target gene is UP-regulated by the factor:				1.998	3.560			3.643				2.349
Target gene is DOWN-regulated by the factor:		-1.098	-1.098			-1.177	-1.098		-1.073	-1.098	-1.001	
Absolute gene regulation:		0.911	0.911	1.998	3.560	0.850	0.911	3.643	0.932	0.911	0.999	2.349
Absolute gene regulation (standard error):		*±0.29945*	*±0.07006*	*±0.27731*	*±0.42808*	*±0.07369*	*±0.07006*	*±0.40444*	*±0.14934*	*±0.07006*	*±0.10083*	*±0.23077*
		-1.098	-1.098	1.998	3.560	-1.177	-1.098	3.643	-1.073	-1.098	-1.001	2.348
Absolute gene regulation (2-log):		-0.135	-0.135	0.999	1.832	-0.235	-0.135	1.865	-0.101	-0.135	-0.001	1.232
Absolute gene regulation (2-log standard error):		*±1.74*	*±3.835*	*±1.85*	*±1.224*	*±3.762*	*±3.835*	*±1.306*	*±2.743*	*±3.835*	*±3.31*	*±2.115*

Table 3.12: Descriptive statistics used to determine the differential expression of miR1-8 in HepG2 in comparison to KMST1 cell line.

control	[reference gene]	[target gene 1]	[target gene 2]	[target gene 3]	[target gene 4]	[target gene 5]	[target gene 6]	[target gene 7]	[target gene 8]	[target gene 9]	[target gene 10]	[target gene 11]
	GAPDH	miR-1	miR-2	miR-3	miR-4	miR-5	miR-6	miR-7	miR-8	internal control	HKG2	PTEN
		*										
n	3	3	3	3	3	3	3	3	3	3	3	3
mean	18.19	34.67	32.00	30.00	31.97	26.03	35.00	35.00	33.03	35.00	28.07	27.07
standard error	0.04	0.33	0.00	0.15	0.07	0.03	0.00	0.00	0.20	0.00	0.07	0.07
CV [%]	0.39	1.67	0.00	0.88	0.36	0.22	0.00	0.00	1.06	0.00	0.41	0.43
sample(s)	[reference gene]	[target gene 1]	[target gene 2]	[target gene 3]	[target gene 4]	[target gene 5]	[target gene 6]	[target gene 7]	[target gene 8]	[target gene 9]	[target gene 10]	[target gene 11]
	GAPDH	miR-1	miR-2	miR-3	miR-4	miR-5	miR-6	miR-7	miR-8	internal control	HKG2	PTEN
n	3	3	3	3	3	3	3	3	3	3	3	3
mean	18.05	34.50	31.53	26.93	32.00	25.97	33.83	31.97	33.13	35.00	27.93	26.50
standard error	0.11	0.25	0.29	0.07	0.06	0.15	0.12	0.19	0.15	0.00	0.07	0.06
CV [%]	1.01	1.26	1.60	0.43	0.31	0.97	0.62	1.01	0.76	0.00	0.41	0.38
	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
E(target)^CP	1.098	1.120	1.382	8.378	0.977	1.047	2.245	8.187	0.933	1.000	1.097	1.481
	1.098											
Normalization Factor **	1.098											
	1.000	1.020	1.259	7.632	0.890	0.954	2.045	7.457	0.850	0.911	0.999	1.349
	[reference gene]	[target gene 1]	[target gene 2]	[target gene 3]	[target gene 4]	[target gene 5]	[target gene 6]	[target gene 7]	[target gene 8]	[target gene 9]	[target gene 10]	[target gene 11]
Exression ratio(s):	GAPDH	miR-1	miR-2	miR-3	miR-4	miR-5	miR-6	miR-7	miR-8	internal control	HKG2	PTEN
Significant (randomization test)												
Target gene is UP-regulated by the factor:		1.020	1.259	7.632			2.045	7.457				1.349
Target gene is DOWN-regulated by the factor:					-1.124	-1.048			-1.177	-1.098	-1.001	
Absolute gene regulation:		1.020	1.259	7.632	0.890	0.954	2.045	7.457	0.850	0.911	0.999	1.349
Absolute gene regulation (standard error):		*±0.29927*	*±0.2714*	*±1.05916*	*±0.08745*	*±0.12288*	*±0.23185*	*±1.11771*	*±0.16083*	*±0.07006*	*±0.10083*	*±0.13254*

Table 3.13: Descriptive statistics used to determine the differential expression of miR1-8 in Caski in comparison to KMST1 cell lines.

control	[reference gene]	[target gene 1]	[target gene 2]	[target gene 3]	[target gene 4]	[target gene 5]	[target gene 6]	[target gene 7]	[target gene 8]	[target gene 9]	[target gene 10]	[target gene 11]
	GAPDH	miR-1	miR-2	miR-3	miR-4	miR-5	miR-6	miR-7	miR-8	internal control	HKG2	PTEN
		*										
n	3	3	3	3	3	3	3	3	3	3	3	3
mean	18.19	34.57	32.00	30.00	31.97	26.00	34.93	34.97	33.13	35.00	28.07	27.07
standard error	0.04	0.30	0.00	0.15	0.07	0.00	0.07	0.03	0.13	0.00	0.07	0.07
CV [%]	0.39	1.48	0.00	0.88	0.36	0.00	0.33	0.17	0.70	0.00	0.41	0.43
sample(s)	[reference gene]	[target gene 1]	[target gene 2]	[target gene 3]	[target gene 4]	[target gene 5]	[target gene 6]	[target gene 7]	[target gene 8]	[target gene 9]	[target gene 10]	[target gene 11]
	GAPDH	miR-1	miR-2	miR-3	miR-4	miR-5	miR-6	miR-7	miR-8	internal control	HKG2	PTEN
n	3	3	3	3	3	3	3	3	3	3	3	3
mean	18.05	34.67	29.43	28.87	30.00	23.10	35.17	35.13	33.03	35.00	27.93	25.70
standard error	0.11	0.33	0.09	0.07	0.12	0.06	0.09	0.13	0.20	0.00	0.07	0.06
CV [%]	1.01	1.67	0.52	0.40	0.67	0.43	0.43	0.66	1.06	0.00	0.41	0.39
	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
E(target)^CP	1.098	0.934	5.924	2.194	3.909	7.464	0.851	0.891	1.072	1.000	1.097	2.579
	1.098											
Normalization Factor **	1.098											
	1.000	0.851	5.396	1.998	3.560	6.799	0.775	0.811	0.976	0.911	0.999	2.349
	[reference gene]	[target gene 1]	[target gene 2]	[target gene 3]	[target gene 4]	[target gene 5]	[target gene 6]	[target gene 7]	[target gene 8]	[target gene 9]	[target gene 10]	[target gene 11]
Exression ratio(s):	GAPDH	miR-1	miR-2	miR-3	miR-4	miR-5	miR-6	miR-7	miR-8	internal control	HKG2	PTEN
Significant (randomization test)												
Target gene is UP-regulated by the factor:			5.396	1.998	3.560	6.799						2.349
Target gene is DOWN-regulated by the factor:		-1.175					-1.291	-1.232	-1.024	-1.098	-1.001	
Absolute gene regulation:		0.851	5.396	1.998	3.560	6.799	0.775	0.811	0.976	0.911	0.999	2.349
Absolute gene regulation (standard error):		±0.26557	±0.53018	±0.27731	±0.42808	±0.58948	±0.08413	±0.09936	±0.18056	±0.07006	±0.10083	±0.23077
		-1.175	5.396	1.998	3.560	6.799	-1.291	-1.232	-1.024	-1.098	-1.001	2.349
Absolute gene regulation (2-log):		-0.233	2.432	0.999	1.832	2.765	-0.368	-0.301	-0.035	-0.135	-0.001	1.232
Absolute gene regulation (2-log standard error):		±1.913	±0.915	±1.85	±1.224	±0.762	±3.571	±3.331	±2.469	±3.835	±3.31	±2.115

Table 3.14: Descriptive statistics used to determine the differential expression of miR1-8 in Caski in comparison to KMST1 cell lines.

control	[reference gene]	[target gene 1]	[target gene 2]	[target gene 3]	[target gene 4]	[target gene 5]	[target gene 6]	[target gene 7]	[target gene 8]	[target gene 9]	[target gene 10]	[target gene 11]
	GAPDH	miR-1	miR-2	miR-3	miR-4	miR-5	miR-6	miR-7	miR-8	internal control	HKG2	PTEN
		*										
n	3	3	3	3	3	3	3	3	3	3	3	3
mean	18.19	34.87	32.27	30.17	32.07	26.00	34.97	35.00	33.13	35.00	28.07	27.07
standard error	0.04	0.13	0.15	0.07	0.09	0.00	0.09	0.00	0.13	0.00	0.07	0.07
CV [%]	0.39	0.66	0.78	0.38	0.48	0.00	0.44	0.00	0.70	0.00	0.41	0.43
sample(s)	[reference gene]	[target gene 1]	[target gene 2]	[target gene 3]	[target gene 4]	[target gene 5]	[target gene 6]	[target gene 7]	[target gene 8]	[target gene 9]	[target gene 10]	[target gene 11]
	GAPDH	miR-1	miR-2	miR-3	miR-4	miR-5	miR-6	miR-7	miR-8	internal control	HKG2	PTEN
n	3	3	3	3	3	3	3	3	3	3	3	3
mean	18.05	34.93	32.13	30.57	32.33	23.13	35.30	34.93	33.10	35.00	27.93	25.70
standard error	0.11	0.07	0.13	0.15	0.18	0.13	0.06	0.07	0.15	0.00	0.07	0.06
CV [%]	1.01	0.33	0.72	0.82	0.94	1.00	0.28	0.33	0.80	0.00	0.41	0.39
	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
E(target)^CP	1.098	0.956	1.097	0.758	0.831	7.294	0.794	1.047	1.023	1.000	1.097	2.579
	1.098											
Normalization Factor **	1.098											
	1.000	0.871	0.999	0.690	0.757	6.644	0.723	0.954	0.932	0.911	0.999	2.349
	[reference gene]	[target gene 1]	[target gene 2]	[target gene 3]	[target gene 4]	[target gene 5]	[target gene 6]	[target gene 7]	[target gene 8]	[target gene 9]	[target gene 10]	[target gene 11]
Exression ratio(s):	GAPDH	miR-1	miR-2	miR-3	miR-4	miR-5	miR-6	miR-7	miR-8	internal control	HKG2	PTEN
Significant (randomization test)												
Target gene is UP-regulated by the factor:						6.644						2.349
Target gene is DOWN-regulated by the factor:		-1.149	-1.001	-1.449	-1.321		-1.383	-1.048	-1.073	-1.098	-1.001	
Absolute gene regulation:		0.871	0.999	0.690	0.757	6.644	0.723	0.954	0.932	0.911	0.999	2.349
Absolute gene regulation (standard error):		*±0.11058*	*±0.1567*	*±0.09311*	*±0.11875*	*±0.79882*	*±0.0767*	*±0.0856*	*±0.14934*	*±0.07006*	*±0.10083*	*±0.23077*
		-1.149	-1.001	-1.449	-1.321	6.644	-1.383	-1.048	-1.073	-1.098	-1.001	2.349
Absolute gene regulation (2-log):		-0.200	-0.001	-0.535	-0.401	2.732	-0.468	-0.068	-0.101	-0.135	-0.001	1.232
Absolute gene regulation (2-log standard error):		*±3.177*	*±2.674*	*±3.425*	*±3.074*	*±0.324*	*±3.705*	*±3.546*	*±2.743*	*±3.835*	*±3.31*	*±2.115*

References

"ACS: What Is Prostate Cancer?" American Cancer Society: Information and Resources for Cancer: Breast, Colon, Prostate, Lung and Other Forms. Web. 9 October 2012.

"Talking With Your Patients About Screening for Prostate Cancer". Retrieved 2012-10-09

*Huang DW, Sherman BT, Lempicki RA, (2009). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res. 2009;37(1):1-13.

Aaltonen LA, Peltomaki P, Leach FS, Sistonen P, Pylkkanen L, Mecklin J-P, Jiirvinen H, Powell SM, Jen J, Hamilton SR (1993): Clues to the pathogenesis of familial colorectal cancer. Science 260:812-816 .

Abelev GI (1963) Production of embryonal alpha-globulin by transplantable mouse hepatomas.Transplantation 1: 174–180

Ahmad A, Aboukameel A, Kong D, Wang Z, Sethi S, Chen W (2011): Phosphoglucose isomerase/autocrine motility factor mediates epithelial-mesenchymal transition regulated by miR-200 in breast cancer cells. Cancer Res, 71:3400-3409.

Ahmad A, Ali AS, Ali S, Wang Z, Kong D, Sarkar FH (2011): MicroRNAs: Targets of Interest in Breast Cancer Research. In MicroRNA: Expression, Detection and Therapeutic Strategies. Edited by Mulligan JA. New York: Nova Publishers;:59-78.

Ali S, Ahmad A, Aboukameel A, Bao B, Padhye S, Philip PA (2012): Increased Ras GTPase activity is regulated by miRNAs that can be attenuated by CDF treatment in pancreatic cancer cells. Cancer Lett.

Ali S, Ahmad A, Banerjee S, Padhye S, Dominiak K, Schaffert JM (2010): Gemcitabine sensitivity can be induced in pancreatic cancer cells through modulation of miR-200 and miR-21 expression by curcumin or its analogue CDF. Cancer Res, 70:3606-3617.

Ali S, Almhanna K, Chen W, Philip PA, Sarkar FH (2010): Differentially expressed miRNAs in the plasma may provide a molecular signature for aggressive pancreatic cancer. Am J Transl Res, 3:28-47.

Ali S, Banerjee S, Logna F, Bao B, Philip PA, Korc M (2011): Inactivation of Ink4a/Arf leads to deregulated expression of miRNAs in K-Ras transgenic mouse model of pancreatic cancer. J Cell Physiol.

Alizadeh AA (2000) Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature 403: 503–511

Alizadeh AA (2001) Towards a novel classification of human malignancies based on gene expression patterns. J Pathol 195: 41–52

Ambros V (2003): MicroRNA pathways in flies and worms: growth, death, fat, stress, and timing. Cell , 113:673-676.

Ambs S, Prueitt RL, Yi M, Hudson RS, Howe TM, Petrocca F, Wallace TA, Liu CG, Volinia S, Calin GA, et al. Genomic profiling of microRNA and

messenger RNA reveals deregulated microRNA expression in prostate cancer. *Cancer Res* 68(15):6162-70, 2008.

Aqeilan RI, Calin GA, Croce CM (2010): miR-15a and miR-16-1 in cancer: discovery, function and future perspectives. Cell Death Differ, 17:215-220.

Arroyo, J.D.; Chevillet, J.R.; Kroh, E.M.; Ruf, I.K.; Pritchard, C.C.; Gibson, D.F.; Mitchell, P.S.; Bennett, C.F.; Pogosova-Agadjanyan, E.L.; Stirewalt, D.L.; (2011). Argonaute2 Complexes Carry a Population of Circulating microRNAs Independent of Vesicles in Human Plasma. Proc. Natl. Acad. Sci. USA , 108, 5003–5008.

Baltimore D, Boldin MP, O'Connell RM, Rao DS, Taganov KD (2008): MicroRNAs: new regulators of immune cell development and function. Nat Immunol , 9:839-845.

Bao B, Ali S, Kong D, Sarkar SH, Wang Z, Banerjee S (2011): Anti-tumour activity of a novel compound-CDF is mediated by regulating miR-21, miR-200, and PTEN in pancreatic cancer. PLoS One, 6:e17850.

Baranwal S, Alahari SK (2010): miRNA control of tumour cell invasion and metastasis. Int J Cancer , 126:1283-1290.

Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, Marshall KA, Phillippy KH, Sherman PM, Holko M, Yefanov A, Lee H, Zhang N, Robertson CL, Serova N, S Davis S, Soboleva A, (2013), NCBI GEO: archive for functional genomics data sets—update Nucleic Acids Res. ; 41(Database issue)

Barringhaus, K. G., & Zamore, P. D. (2009). MicroRNAs: Regulating a Change of Heart. Journal of the American Heart Association, 199, 2217-2224.

Barringhaus, K. G., & Zamore, P. D. (2009). MicroRNAs: Regulating a Change of Heart. Journal of the American Heart Association, 199, 2217-2224.

Bast RC Jr, (1981) Reactivity of a monoclonal antibody with human

ovarian carcinoma. J Clin Invest 68: 1331–1337

Bast RC Jr, Ravdin P, Hayes DF, Bates S, Fritshe H Jr, Jessup JM, Kemeny N, Locker GY, Mennel RG, Somerfield MR (2001) 2000 update of recommendations for the use of tumour markers in breast and colorectal cancer: clinical practice guidelines of the American Society of Clinical Oncology. J Clin Oncol 19: 1865–1878.

Berger PL, Frank SB2, Schulz VV1, Nollet EA2, Edick MJ1, Holly B1, Chang TT1, Hostetter G1, Kim S1, Miranti CK3 (2014):.Transient induction of ING4 by Myc drives prostate epithelial cell differentiation and its disruption drives prostate tumourigenesis. Cancer Res;74(12):3357-68

Bernstein E, Caudy AA, Hammond SM, Hannon GJ (2001): Role for a bidentate ribonuclease in the initiation step of RNA interference. Nature, 409:363-366.

Bhattacharyya NP, Skandalis A, Ganesh A, Groden J, Meuth M (1994): Mutator phenotypes in human colorectal carcinoma cell lines. Proc Natl Acad Sci USA 91:6319-6323.

Bohnsack MT, Czaplinski K, Gorlich D: (2004) Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. RNA, 10:185-191.

Bommer GT, Gerin I, Feng Y, Kaczorowski AJ, Kuick R, Love RE (2007): p53-mediated activation of miRNA34 candidate tumour-suppressor genes. Curr Biol, 17:1298-1307.

Bonci D, Coppola V, Musumeci M, Addario A, Giuffrida R, Memeo L, D'Urso L, Pagliuca A, Biffoni M, Labbaye C, et al. The miR-15a-miR-16-1 cluster controls prostate cancer by targeting multiple oncogenic activities. *Nat Med* 14(11):1271-7, 2008.

Bonekamp D, Jacobs MA, El-Khouli R, Stoianovici D, Macura KJ (**2011**). "Advancements in MR Imaging of the Prostate: From Diagnosis to Interventions". *Radiographics* **31** (3 Suppl): 677–703.

Bonekamp D, Jacobs MA, El-Khouli R, Stoianovici D, Macura KJ (2011). "Advancements in MR Imaging of the Prostate: From Diagnosis to Interventions". *Radiographics* 31 (3 Suppl): 677–703.

Bracken CP, Gregory PA, Khew-Goodall Y, Goodall GJ (2009): The role of microRNAs in metastasis and epithelial-mesenchymal transition. Cell Mol Life Sci, 66:1682-1699.

Brase JC, Johannes M, Schlomm T, Falth M, Haese A, Steuber T (2011): Circulating miRNAs are correlated with tumour progression in prostate cancer. Int J Cancer, 128:608-616.

Buckingham, S. (2003). The Major World of microRNAs. Understanding the RNAissance (pp. 1-3). Prouts Neck, Maine: Nature Publishing Group.

Bushati N, Cohen SM (2008): MicroRNAs in neurodegeneration. Curr Opin Neurobiol, 18:292-296.

Capaccio D, Ciccodicola A, Sabatino L et al., 2010. "A novel germline

mutation in peroxisome proliferator-activated receptor gamma gene associated with large intestine polyp formation and dyslipidemia". *Biochim. Biophys. Acta* **1802** (6): 572

Chang TC, Wentzel EA, Kent OA, Ramachandran K, Mullendore M, Lee KH (2007): Transactivation of miR-34a by p53 broadly influences gene expression and promotes apoptosis. Mol Cell, 26:745-752.

Chen HY, Yu Sl, Chen CH, Chang CG, Chen CY, Yuan A, Cheng CL, Wang CH, Terng HJ, Kao SF, Chan WK, Li HN, Liu CC, Chen JJ and Yang PC (2007) A five-gene signature and clinical outcome in non-small-cell lung cancer. N Engl J Med 356: 11–20

Chen, X.; Ba, Y.; Ma, L.; Cai, X.; Yin, Y.; Wang, K.; Guo, J.; Zhang, Y.; Chen, J.; Guo, X. (2008); Characterization of microRNAs in Serum: A Novel Class of Biomarkers for Diagnosis of Cancer and Other Diseases. Cell Res. , 18, 997–1006

Chu EC, Tarnawski AS (2004). "PTEN regulatory functions in tumour

suppression and cell biology". Medical Science Monitor 10 (10): RA235–41

Clape C, Fritz V, Henriquet C, Apparailly F, Fernandez PL, Iborra F, (2009): miR-143 interferes with ERK5 signaling, and abrogates prostate cancer progression in mice. PLoS One, 4:e7542.

Coffey RN, Watson RW, O'Neill AJ, McElenyK & Fitzpatrick JM (2002) Androgen-mediated resistance to apoptosis.Prostate 53 300–309.

Collino, F.; Deregibus, M.C.; Bruno, S.; Sterpone, L.; Aghemo, G.; Viltono, L.; Tetta, C.; Camussi, G. (2010) Microvesicles Derived from Adult Human Bone Marrow and Tissue Specific Mesenchymal Stem Cells Shuttle Selected Pattern of miRNAs. PLoS One , 5, e11803.

Collins AT, Berry PA, Hyde C, Stower MJ, Maitland NJ (2005): Prospective identification of tumourigenic prostate cancer stem cells. Cancer Res, 65:10946-10951.

Coppola V, De MR, Bonci D (2010): MicroRNAs and prostate cancer. Endocr Relat Cancer, 17:F1-F17.

Croce, C. M. Molecular origins of cancer: Oncogenes and Cancer. *N. Engl. J. Med.* **2008**, *358*, 502–511

Cullen BR (2009): Viral and cellular messenger RNA targets of viral microRNAs. Nature, 457:421-425.

Cuperus, J. T., Fahlgren, N., & Carrington, J. C. (2011). Evolution and Functional Diversification of MIRNA Genes. The Plant Cell, 23, 431-442.

Cuperus, J. T., Fahlgren, N., & Carrington, J. C. (2011). Evolution and Functional Diversification of miRNA Genes. The Plant Cell, 23, 431-442.

Damber JE, Aus G. Prostate cancer. *Lancet* 371:1710-21, 2008.

Davis, B. N., & Hata, A. (2009). Regulation of MicroRNA Biogenesis: A miRiad of mechanisms. Cell Communication and Signalling, 7(18).

de la Chapelle A, Peltomaki P (1995): Genetics of hereditary colon cancer. Annu Rev Genet 29:329-349.

Dedeoğlu, B.G. (2014). High-throughput approaches for microRNA expresson analysis. Methods in Molecular Biology, 1107: 91-103

Denli AM, Tops BB, Plasterk RH, Ketting RF, Hannon GJ (2004): Processing of primary microRNAs by the Microprocessor complex. Nature, 432:231-235.

DeSano JT, Xu L (2009): MicroRNA regulation of cancer stem cells and therapeutic implications. AAPS J, 11:682-692.

Deyati, A, Younesi, E, Hofmann-Apitius, M, Novac, N. (2012) Challenges and opportunities for oncology biomarker discovery. Drug Discov Today. 2012 Dec(epub).

Dick JE (1996): Normal and leukemic human stem cells assayed in SCID

mice. Semin Immunol, 8:197-206.

Djulbegovic M, Beyth RJ, Neuberger MM, Stoffs TL, Vieweg J, Djulbegovic B, Dahm P (**2010**). "Screening for prostate cancer: systematic review and meta-analysis of randomised controlled trials". *BMJ* **341**: c4543.

Doench JG, Sharp PA (2004): Specificity of microRNA target selection in translational repression. Genes Dev , 18:504-511.

Eisen MB (1998) Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci USA 95: 14863–14868

Erson, A. E., & Petty, E. M. (2008). MicroRNAs in development and disease. Clinical Genetics, 74, 396-306.

Essink-Bot ML, de Koning HJ, Nijs HG, Kirkels WJ, van der Maas PJ, Schröder FH (**1998**). "Short-term effects of population-based screening for

prostate cancer on health-related quality of life". *J. Natl. Cancer Inst.* **90** (12): 925–31.

Essink-Bot ML, de Koning HJ, Nijs HG, Kirkels WJ, van der Maas PJ, Schröder FH (1998). "Short-term effects of population-based screening for prostate cancer on health-related quality of life". *J. Natl. Cancer Inst.* 90 (12): 925–31.

Eulalio, A., Huntzinger, E., Nishihara, T., Rehwinkel, J., Fauser, M., & Izaurralde, E. (2009). Deadenylation is a widespread effect of miRNA regulation. RNA, 15, 21-32.

Eulalio, A., Huntzinger, E., Nishihara, T., Rehwinkel, J., Fauser, M., & Izaurralde, E. (2009). Deadenylation is a widespread effect of miRNA regulation. RNA, 15, 21-32.

Exploratorium. (2004, February). Cancer: Cells Behaving Badly. Retrieved April 2013, from http://www.exploratorium.edu/imaging-

station/research/cancer/story_cancer.pdf

Fabian, M. R., Sonenberg, N., & Filipcowicz, W. (2010). Regulation of mRNA Translation and Stability by microRNAs. Annual Review of Biochemistry, 351-379.

Fabian, M. R., Sonenberg, N., & Filipcowicz, W. (2010). Regulation of mRNA Translation and Stability by microRNAs. Annual Review of Biochemistry, 351-379.

Filipowicz, W. Post-transcriptional gene silencing by siRNAs and miRNAs. *Curr. Opin. Struct. Biol.* **2005**, *15*, 331–341.

Finne P, Finne R, Auvinen A, Juusela H, Aro J, Maattanen L, Hakama M, Rannikko S, Tammela TL, Stenman U (2000) Predicting the outcome of prostate biopsy in screen-positive men by a multilayer perceptron network. Urology 56: 418–422

Folini M, Gandellini P, Longoni N, Profumo V, Callari M, Pennati M, Colecchia M, Supino R, Veneroni S, Salvioni R, et al. miR-21: an oncomir on strike in prostate cancer. *Mol Cancer* 9(1):12, 2010.

Franceschini A, Szklarczyk D, Frankild S, Kuhn M, Simonovic M, Roth A, Lin J, Minguez P, Bork P, von Mering C, Jensen LJ. (2013), STRING v9.1: protein-protein interaction networks, with increased coverage and integration. Nucleic Acids Res. :D808-15.

Friedman RC, Farh KK, Burge CB, Barte DPl. (2009) Most Mammalian mRNAs Are Conserved Targets of MicroRNAs Genome Research, 19:92-105.

Fu X, Xue C, Huang Y, Xie Y, Li Y (2010): The activity and expression of microRNAs in prostate cancers. Mol Biosyst, 6:2561-2572.

Galardi S, Mercatelli N, Giorda E, Massalini S, Frajese GV, Ciafre SA (2007): miR-221 and miR-222 expression affects the proliferation potential of human prostate carcinoma cell lines by targeting p27Kip1. J Biol Chem, 282:23716-23724.

Gandellini P, Folini M, Longoni N, Pennati M, Binda M, Colecchia M, Salvioni R, Supino R, Moretti R, Limonta P, et al. miR-205 Exerts tumor-suppressive functions in human prostate through down-regulation of protein kinase Cepsilon. *Cancer Res* 69(6):2287-95, 2009b.

Gandellini P, Folini M, Zaffaroni N. Towards the definition of prostate cancer-related microRNAs: where are we now? *Trends Mol Med* 15(9):381-90, 2009a.

Gangaraju VK, Lin H (2009): MicroRNAs: key regulators of stem cells. Nat Rev Mol Cell Biol, 10:116-125.

Gao S, Lee P, Wang H, Gerald W, Adler M, Zhang L (2005). Theandrogen receptor directly targets the cellular Fas/FasL-associated death domain protein-like inhibitory protein gene to promote the androgen independent growth of prostate cancer cells. Mol Endocrinol;19:1792 802. 10.

Gao S, Wang H, Lee P, Melamed J, Li CX, Zhang F, (2006) Androgen receptor and prostate apoptosis response factor-4 target the CFLAR gene to determine survival and apoptosis in the prostate gland. J Mol Endocrinol ; 36:463 –83

García D M, Baek D, C Shin C, Bell G W, Grimson A, Bartel D P (2011): Weak Seed-Pairing Stability and High Target-Site Abundance Decrease the Proficiency of lsy-6 and Other miRNAs Nat Struct Mol Biol., 18:1139-1146.

Garzon R, Fabbri M, Cimmino A, Calin GA, Croce CM. MicroRNA expression and function in cancer. *Trends Mol Med* 12(12):580-7, 2006.

Gerstein, M. (1999). BIOINFORMATICS Introduction. Retrieved May 2013, from http://www.imb-jena.de/~rake/PDF/gerstein.pdf

Gerstein, M. (1999). BIOINFORMATICS Introduction. Retrieved May 2013, from http://www.imb-jena.de/~rake/PDF/gerstein.pdf

Gilad, S.; Meiri, E.; Yogev, Y.; Benjamin, S.; Lebanony, D.; Yerushalmi, N.;

Benjamin, H.; Kushnir, M.; Cholakh, H.; Melamed, N.; (2008). Serum microRNAs are Promising Novel Biomarkers. PLoS One , 3, e3148.

Gold P and Freedman SO (1965) Specific carcinoembryonic antigens of the human digestive system. J Exp Med 122: 467–81

Golub TR (1999) Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science 286: 531–537

Griffiths-Jones S, Grocock RJ, Van DS, Bateman A, Enright AJ (2006): miRBase: microRNA sequences, targets and gene nomenclature. Nucleic Acids Res, 34:D140-D144.

Griffiths-Jones S. (2004), The microRNA Registry.NAR 32:D109-D111

Grimson A, Farh K K, Johnston W K, Garrett-Engele P, Lim L P , Bartel D P (2007) MicroRNA Targeting Specificity in Mammals: Determinants beyond Seed Pairing Molecular Cell, 27:91-105.

Gulati R, Tsodikov A, Wever EM. (2012) The impact of PLCO control arm

contamination on perceived PSA screening efficacy. Cancer Causes Control. 23: 827.

Hakimi JM, Rondinelli RH, Schoenberg MP & Barrack ER (1996) Androgen-receptor gene structure and function in prostate cancer. World Journal of Urology 14 329–337

Hammond SM, Bernstein E, Beach D, Hannon GJ (2000): An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells. Nature, 404:293-296.

Hanahan, D. & Weinberg, R. A. 2011 Hallmarks of cancer: the next generation. Cell 144, 646–674

Hao Y, Zhao Y, Zhao X, He C, Pang X, Wu TC, (2011): Improvement of prostate cancer detection by integrating the PSA test with miRNA expression profiling. Cancer Invest , 29:318-324.

He L, He X, Lim LP, De SE, Xuan Z, Liang Y (2007): A microRNA

component of the p53 tumour suppressor network. Nature, 447:1130-1134.

He L, Thomson JM, Hemann MT, *et al.* (**2005**). "A microRNA polycistron as a potential human oncogene". *Nature* **435** (7043): 828–33.

He L, Thomson JM, Hemann MT, *et al.* (2005). "A microRNA polycistron as a potential human oncogene". *Nature* 435 (7043): 828–33.

Hemminki A, Peltomaki P, Mecklin J-P, Jarvinen H, Salovaara R, Nystr6m-Lahti M, de la Chapelle A, Aaltonen LA (1994): Loss of the wild type MLH1 gene is a feature of hereditary nonpolyposis colorectal cancer. Nat Genet 8:405-410.

Huang DW, Sherman BT, Lempicki RA, (2009). Systematic and integrative analysis of large gene lists using DAVID Bioinformatics Resources. Nature Protoc. 2009;4(1):44-57

Iizuka N, Hamamoto Y, Oka M (2004) Predicting individual outcomes in

hepatocellular carcinoma. Lancet 364: 1837–1839

Ishiguro H, Akimoto K, Nagashima Y, Kojima Y, Sasaki T, Ishiguro-Imagawa Y, Nakaigawa N, Ohno S, Kubota Y, Uemura H (2009). aPKClambda/iota promotes growth of prostate cancer cells in an autocrine manner through transcriptional activation of interleukin-6. Proc Natl Acad Sci USA. 22;106(38):16369-74

Jin L, Hope KJ, Zhai Q, Smadja-Joffe F, Dick JE (2006): Targeting of CD44 eradicates human acute myeloid leukemic stem cells. Nat Med, 12:1167-1174.

Jones HB (1848) On a new substance occuring in the urine with mollities ossium. Phil Trans R Soc Lond 138: 55–62

Kalluri R, Weinberg RA (2009): The basics of epithelial-mesenchymal transition. J Clin Invest, 119:1420-1428.

Kanehisa M, Goto S (2000). "KEGG: Kyoto Encyclopedia of Genes and

Genomes". Nucleic Acids Res 28 (1): 27–30.

Ketting RF, Fischer SE, Bernstein E, Sijen T, Hannon GJ, Plasterk RH (2001): Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in C. elegans. Genes Dev, 15:2654-2659.
Kloosterman WP, Plasterk RH (2006): The diverse functions of microRNAs in animal development and disease. Dev Cell, 11:441-450.

Kong D, Banerjee S, Ahmad A, Li Y, Wang Z, Sethi S (2010): Epithelial to mesenchymal transition is mechanistically linked with stem cell signatures in prostate cancer cells. PLoS One, 5:e12445.

Kong D, Li Y, Wang Z, Banerjee S, Ahmad A, Kim HR (2009): miR-200 regulates PDGF-D-mediated epithelial-mesenchymal transition, adhesion, and invasion of prostate cancer cells. Stem Cells, 27:1712-1721.

Krutovskikh, V.A. and Herceg, Z. (2010) Oncogenic microRNAs (OncomiRs) as a new class of cancer biomarkers. Bioessays 32, 894–904
Latronico MV, Condorelli G (2009): MicroRNAs and cardiac pathology.

Nat Rev Cardiol , 6:419-429.

Lazar V, Grandjouan S, Bognel C, Couturier D, Rougier P, Bellet D, Bressac-de Paillerets B (1994): Accumulation of multiple mutations in tumour suppressor genes during colorectal tumourigenesis in HNPCC patients. Hum Mol Genet 3:2257-2260.

Lee RC, Feinbaum RL, Ambros V (1993): The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell, 75:843-854.

Lee Y, Ahn C, Han J, Choi H, Kim J, Yim J (2003) : The nuclear RNase III Drosha initiates microRNA processing. Nature, 425:415-419.

Lee Y, Kim M, Han J, Yeom KH, Lee S, Baek SH (2004) : MicroRNA genes are transcribed by RNA polymerase II. EMBO J, 23:4051-4060.

Lee YS, Kim HK, Chung S, Kim KS, Dutta A (2005): Depletion oPf human micro-RNA miR-125b reveals that it is critical for the proliferation of

differentiated cells but not for the down-regulation of putative targets during differentiation. J Biol Chem, 280:16635-16641.

Leshem O, Madar S, Kogan-Sakin I, Kamer I, Goldstein I, Brosh R (2011): TMPRSS2/ERG promotes epithelial to mesenchymal transition through the ZEB1/ZEB2 axis in a prostate cancer model. PLoS One, 6:e21650.

Levinson G, Gutman GA (1987): High frequencies of short frameshifts in poly-CA/TG tandem repeats borne by bacteriophage M13 in Escherichia coli K-12. Nucleic Acids Res 15:5313-5338 .

Lewis B P, Burge C B, Bartel D P (2005): Conserved Seed Pairing, Often Flanked by Adenosines, Indicates that Thousands of Human Genes are MicroRNA Targets Cell, 120:15-20.

Li J, Smyth P, Flavin R, Cahill S, Denning K, Aherne S, Guenther SM, O'Leary JJ, Sheils O. Comparison of miRNA expression patterns using total RNA extracted from matched samples of formalin-fixed paraffin-embedded (FFPE) cells and snap frozen cells. *BMC Biotechnol* 7:36, 2007.

Li T, Li D, Sha J, Sun P, Huang Y (2009): MicroRNA-21 directly targets MARCKS and promotes apoptosis resistance and invasion in prostate cancer cells. Biochem Biophys Res Commun, 383:280-285.

Li, J.; Smyth, P.; Flavin, R.; Cahill, S.; Denning, K.; Aherne, S.; Guenther, S.M.; O'Leary, J.J.; Sheils, O. (2007) Comparison of miRNA Expression Patterns using Total RNA Extracted from Matched Samples of Formalin-Fixed Paraffin-Embedded (FFPE) Cells and Snap Frozen Cells. BMC Biotechnol., 7, 36.

Liu B, Nicolaides NC, Markowitz S, Willson JKV, Parsons RE, Jen J, Papadopoulos N, Peltomaki P, de la Chapelle A, Hamilton SR (1995): Mismatch repair gene defects in sporadic colorectal cancers with microsatellite instability. Nat Genet 9:48-55 .

Liu B, Parsons R, Papadopoulos N, Nicolaides NC, Lynch HT, Watson P, Jass JR, Dunlop M, Wyllie A, Jessup JM (1996): Analysis of mismatch repair genes in hereditary non-polyposis colorectal cancer patients. Nat Med 2:69-174.

Liu C, Kelnar K, Liu B, Chen X, Calhoun-Davis T, Li H (2011): The microRNA miR-34a inhibits prostate cancer stem cells and metastasis by directly repressing CD44. Nat Med, 17:211-215.

Liu CG, Calin GA, Volinia S and Croce CM (2008) MicroRNA expression profiling using microarrays. Nat. Protoc. 3, 563–578

Livak, K.J., and Schmittgen, T.D., Analysis of relative gene expression data using real-time quantitative PCR and the 2^[-delta delta C(T)] Method. Methods., 25(4): 402-408, 2001.

Lodish, H., Berk, A., Matsudaira, P., Kaiser, C. A., Krieger, M., Scott, M. P., (2003). Cancer. In Molecular Cell Biology (p. Chapter 23). Gordonsville, Virginia: WHFreeman; 5th edition.

Lodish, H., Berk, A., Matsudaira, P., Kaiser, C. A., Krieger, M., Scott, M. P., et al. (2003). Cancer. In Molecular Cell Biology (p. Chapter 23). Gordonsville, Virginia: WHFreeman; 5th edition.

Loeb LA (1994): Microsatellite instability: marker of a mutator phenotype in cancer. Cancer Res 54:5059-5063.

Lu J, Getz G, Miska EA, Alvarez-Saavedra E, Lamb J, Peck D, Sweet-Cordero A, Ebert BL, Mak RH, Ferrando AA, Downing JR, et al. MicroRNA expression profiles classify human cancers. *Nature* 435(7043):834-8, 2005.

Lu Z, Liu M, Stribinskis V, Klinge CM, Ramos KS, Colburn NH (2008): MicroRNA-21 promotes cell transformation by targeting the programmed cell death 4 gene. Oncogene , 27:4373-4379.

Luscombe, N., Greenbaum, D., & Gerstein , M. (2001). What is Bioinformatics? A Proposed Definition and Overview of the Field. Methods of Information in Medicine, 40, 348-356.

Maragkakis M; Reczko M; Simossis VA; Alexiou P; Papadopoulos GL; Dalamagas T; Giannopoulos G; Goumas G; Koukis E; Kourtis K;

Vergoulis T; Koziris N; Sellis T; Tsanakas P; Hatzigeorgiou AG (2009). DIANA-microT web server: elucidating microRNA functions through target prediction. Nucleic Acids Research, 37(Web Server issue):W273-6.

Marcione, Marilyn (2011). "Prostate testing's dark side: Men who were harmed". Associated Press.

Markowitz S, Wang J, Myeroff L, Parsons R, Sun L, Lutterbaugh J, Fan RS, Zborowska E, Kinzler KW, Vogelstein B (1995): Inactivation of the type II TGF-beta receptor in colon cancer cells with microsatellite instability. Science 268:1336-1338.

Martinez-Sanchez , A., & Murphy, C. (2013). MicroRNA Target Identification-Experimental Approaches. Biology, 2, 189-205.

Martinez-Sanchez , A., & Murphy, C. (2013). MicroRNA Target Identification-Experimental Approaches. Biology, 2, 189-205.

Mascrenghe, M. (2004). How Bioinformatics can change your life . Retrieved May 2013, from www.oocities.org/mark_ai/bioinformatics.ppt

Mascrenghe, M. (2004). How Bioinformatics can change your life. available at: www.oocities.org/mark_ai/bioinformatics.ppt

Mattie MD, Benz CC, Bowers J, Sensinger K, Wong L, Scott GK (2006): Optimized high-throughput microRNA expression profiling provides novel biomarker assessment of clinical prostate and breast cancer biopsies. Mol Cancer, 5:24.

Mattie MD, Benz CC, Bowers J, Sensinger K, Wong L, Scott GK, Fedele V, Ginzinger D, Getts R, Haqq C. Optimized high-throughput microRNA expression profiling provides novel biomarker assessment of clinical prostate and breast cancer biopsies. *Mol Cancer* 5:24, 2006.

Melton C, Judson RL, Blelloch R (2010): Opposing microRNA families regulate self-renewal in mouse embryonic stem cells. Nature , 463:621-626.

Meltzer P. S. Cancer genomics: Small RNAs with big impacts. *Nature* **2005**, *435*, 745–746

Meltzer P. S. Cancer genomics: Small RNAs with big impacts. *Nature* **2005**, *435*, 745–746

Micalizzi DS, Farabaugh SM, Ford HL (2010): Epithelial-mesenchymal transition in cancer: parallels between normal development and tumour progression. J Mammary Gland Biol Neoplasia, 15:117-134.

Mitchell PS, Parkin RK, Kroh EM, Fritz BR, Wyman SK, Pogosova-Agadjanyan EL, Peterson A, Noteboom J, O'Briant KC, Allen A, et al. Circulating microRNAs as stable blood-based markers for cancer detection. *Proc Natl Acad Sci U S A* 105:10513-8, 2008.

Mitchell, P.S.; Parkin, R.K.; Kroh, E.M.; Fritz, B.R.; Wyman, S.K.; Pogosova-Agadjanyan, E.L.; Peterson, A.; Noteboom, J.; O'Briant, K.C.; Allen, A.; (2008). Circulating microRNAs as Stable Blood-Based Markers for Cancer Detection. Proc. Natl. Acad. Sci. USA , 105, 10513–10518.

Modrich P. (1991) Mechanisms and biological effects of mismatch repair. Annu Rev Genet 25:229-253 .

Mraz M, Pospisilova S, Malinova K, *et al.* (**2009**). "MicroRNAs in chronic

lymphocytic leukemia pathogenesis and disease subtypes". *Leuk Lymphoma* **50** (3): 506–9.

Mraz M, Pospisilova S, Malinova K, *et al.* (2009). "MicroRNAs in chronic lymphocytic leukemia pathogenesis and disease subtypes". *Leuk Lymphoma* 50 (3): 506–9.

Murata T, Takayama K, Urano T, Fujimura T, Ashikari D, Obinata D, Horie-Inoue K, Takahashi S, Ouchi Y, Homma Y, Inoue S. (2012) 14-3-3ζ, a novel androgen-responsive gene, is upregulated in prostate cancer and promotes prostate cancer cell proliferation and survival. Clin Cancer Res. 2012 Oct 15;18(20):5617-27

Musumeci M, Coppola V, Addario A, Patrizii M, Maugeri-Sacca M, Memeo L (2011): Control of tumour and microenvironment cross-talk by miR-15a and miR-16 in prostate cancer. Oncogene, 30:4231-4242.

Myeroff LL, Parsons R, Kim S-J, Hedrick L, ChoKR, Orth K, Mathis M,

Kinzler KW, Lutterbaugh J, Park K (1995): A transforming growth factor 3 receptor type II gene mutation common in colon and gastric but rare in endometrial cancers. Cancer Res 55:5545-5547

Nazarov, P., Reinsbach, S. E., Muller, A., Nicot, N., Philippidou, D., Vallar, L., (2013). Interplay of microRNAs, transcription factors and target genes: linking dynamic expression changes to function. Nucleic Acids Research, 1-15.

Nazarov, P., Reinsbach, S. E., Muller, A., Nicot, N., Philippidou, D., Vallar, L., et al. (2013). Interplay of microRNAs, transcription factors and target genes: linking dynamic expression changes to function. Nucleic Acids Research, 1-15.

Nguyen DP, Li J, Tewari AK. (2014). Inflammation and prostate cancer: the role of interleukin 6 (IL-6). 986-92.

Nilsen, T. W. (2007). Mechanisms of microRNA-mediated gene regulation in animal cells. TRENDS in Genetics, 23(5), 243-249.

Ozen M, Creighton CJ, Ozdemir M, Ittmann M (2008): Widespread deregulation of microRNA expression in human prostate cancer. Oncogene, 27:1788-1793.

Ozen M, Creighton CJ, Ozdemir M, Ittmann M. Widespread deregulation of microRNA expression in human prostate cancer. *Oncogene* 27(12):1788-93, 2008.

Paik S (2004) A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. N Engl J Med 351: 2817–2826

Pang Y, Young CY, Yuan H (2010): MicroRNAs and prostate cancer. cta Biochim Biophys Sin (Shanghai), 42:363-369.

Papagiannakopoulos T, Kosik KS (2008): MicroRNAs: regulators of oncogenesis and stemness. BMC Med, 6:15.

Papsidero LD (1980) A prostate antigen in sera of prostatic cancer patients. Cancer Res 40: 2428–2432

Park, N.J.; Zhou, H.; Elashoff, D.; Henson, B.S.; Kastratovic, D.A.; Abemayor, E.; Wong, D.T. (2009) Salivary microRNA: Discovery, Characterization, and Clinical Utility for Oral Cancer Detection. Clin. Cancer Res. , 15, 5473–5477.

Parsons R, Li G-M, Longley M, Modrich P, Liu B, Berk T, Hamiton SR, Kinzler KW, Vogelstein B (1995): Mismatch repair deficiency in phenotypically normal human cells. Science 268:738-740

Parsons R,Li G-M, Longley MJ, Fang W, Papadopoulos N, Jen J, de la Chapelle A, Kinzler KW, Vogelstein B, Modrich P (1993): Hypermutability and mismatch repair deficiency in RER+ tumour cells. Cell 75:1227-1236 .

Partin AW, Oesterling JE (1994): The clinical usefulness of prostate specific antigen: update. J Urol 152 (5 Pt 1): 1358-68, 1994

Patrawala L, Calhoun T, Schneider-Broussard R, Li H, Bhatia B, Tang S (2006): Highly purified CD44+ prostate cancer cells from xenograft human tumours are enriched in tumourigenic and metastatic progenitor cells.

Oncogene , 25:1696-1708.

Patrawala L, Calhoun-Davis T, Schneider-Broussard R, Tang DG (2007): Hierarchical organization of prostate cancer cells in xenograft tumours: the CD44+alpha2beta1+ cell population is enriched in tumour-initiating cells. Cancer Res, 67:6796-6805.

Peltomaki P, Aaltonen LA, Sistonen P, Pylkkanen L, Mecklin J-P, Jarvinen H, Green JS, Jass JR, Weber JL,Leach FS (1993): Genetic mapping of a locus predisposing to human colorectal cancer. Science 260:810-812.

Peng X, Guo W, Liu T, Wang X, Tu X, 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:e20341.

Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, Pollack JR, Ross DT, Johnsen H, Arslen LA, Fluge O, Pergamenschikov A, Williams C, Zhu SX, Lonning PE, Borrensen-Dale Al, Brown PO, Botstein D (2000) Molecular portraits of human breast tumours. Nature 406: 747–

752

Pfaffl MW, Hageleit M: Validities of mRNA quantification using recombinant RNA and recombinant DNA external calibration curves in real-time RT-PCR. Biotechn Lett2001, 23:275-282.

Pfaffl, MW., A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res., 29 (9): e452001.

Pfaffl, M.W., Horgan, G.W., and Dempfle, L ., Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. Nucleic Acids Res.,30(9):e36, 2002.

69.

Pfaffl, M.W., Lange, I.G., Daxenberger, A., and Meyer, H.H.D., Tissue-specific expression pattern of estrogen receptors (ER): quantification of ER alpha and ER beta mRNA with real-time RT-PCR. APMIS.,109(5): 345-355, 2001. (b)

Pollack JR (2007) A perspective on DNA microarrays in pathology research

and practice. Am J Pathol 171: 375–385

Pomeroy SL (2002) Prediction of central nervous system embryonal tumour outcome based on gene expression. Nature 415: 436–442

Porkka KP, Pfeiffer MJ, Waltering KK, Vessella RL, Tammela TL, Visakorpi T (2007): MicroRNA expression profiling in prostate cancer. Cancer Res, 67:6130-6135.

Porkka KP, Pfeiffer MJ, Waltering KK, Vessella RL, Tammela TL, Visakorpi T. MicroRNA expression profiling in prostate cancer. *Cancer Res* 67(13):6130-5, 2007.

Pritchard, C., Cheng, H., & Tewari, M. (2012). MicroRNA profiling: approaches and considerations. Nature Reviews Genetics, 13, 358-369.

Prueitt RL, Yi M, Hudson RS, Wallace TA, Howe TM, Yfantis HG, Lee DH, Stephens RM, Liu CG, Calin GA, Croce CM, Ambs S. Expression of microRNAs and protein-coding genes associated with perineural invasion

in prostate cancer. *Prostate* 68:1152-64, 2008.

Quackenbush J (2006) Microarray analysis and tumour classification. N Engl J Med 354: 2463–2472

Raver-Shapira N, Marciano E, Meiri E, Spector Y, Rosenfeld N, Moskovits N (2007): Transcriptional activation of miR-34a contributes to p53-mediated apoptosis. Mol Cell, 26:731-743.

Reinhart BJ, Slack FJ, Basson M, Pasquinelli AE, Bettinger JC, Rougvie AE, (2000): The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. Nature, 403:901-906.

Renan MJ (1993): How many mutations are required for tumourigenesis? Implications from human cancer data. Mol Carcinog 7:139-146.

Ribas J, Ni X, Haffner M, Wentzel EA, Salmasi AH, Chowdhury WH (2009): miR-21: an androgen receptor-regulated microRNA that promotes hormone-dependent and hormone-independent prostate cancer growth.

Cancer Res, 69:7165-7169.

Rosenwald A (2002) The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma. N Engl J Med 346: 1937–1947

Rüenauver K, Menon R, Svensson MA, Carlsson J, Vogel W, Andrén O, Nowak M, Perner S. (2014). Prognostic significance of YWHAZ expression in localized prostate cancer.Prostate Cancer Prostatic Dis. 310-4.

Sato, F., Tsuchiya, S., Meltzer, S., & Shimizu, K. (2011). MicroRNAs and epigenetics. Federation of European Biochemical Societies Journal, 278(10), 1598–1609.

Sato, F., Tsuchiya, S., Meltzer, S., & Shimizu, K. (2011). MicroRNAs and epigenetics. Federation of European Biochemical Societies Journal, 278(10), 1598–1609.

Schaefer A, Jung M, Mollenkopf HJ, Wagner I, Stephan C, Jentzmik F (2010): Diagnostic and prognostic implications of microRNA profiling in

prostate carcinoma. Int J Cancer 2010, 126:1166-1176.

Schaefer A, Jung M, Mollenkopf HJ, Wagner I, Stephan C, Jentzmik F, Miller K, Lein M, Kristiansen G, Jung K. Diagnostic and prognostic implications of microRNA profiling in prostate carcinoma. *Int J Cancer* 126(5):1166-76, 2010.

Schwarz DS, Hutvagner G, Haley B, Zamore PD (2002): Evidence that siRNAs function as guides, not primers, in the Drosophila and human RNAi pathways. Mol Cell, 10:537-548.

Selcuklu SD, Donoghue MT, Spillane C (2009): miR-21 as a key regulator of oncogenic processes. Biochem Soc Trans, 37:918-925.

Settles, B. (2008). Lecture 1 Introduction to Bioinformatics. Retrieved May 2013, from http://pages.cs.wisc.edu/~bsettles/ibs08/lectures/01-intro.pdf

Shi XB, Xue L, Yang J, Ma AH, Zhao J, Xu M (2007) : An androgen-

regulated miRNA suppresses Bak1 expression and induces androgen-independent growth of prostate cancer cells. Proc Natl Acad Sci USA , 104:19983-19988.

Shibata D, Peinado MA, Ionov Y, Malkhosyan S, Perucho M (1994): Genomic instability in repeated sequences is an early somatic event in colorectal tumourigenesis that persists after transformation. Nat Genet 6:273-281.

Shimono Y, Zabala M, Cho RW, Lobo N, Dalerba P, Qian D (2009): Downregulation of miRNA-200c links breast cancer stem cells with normal stem cells. Cell, 138:592-603.

Si ML, Zhu S, Wu H, Lu Z, Wu F, Mo YY (2007): miR-21-mediated tumour growth. Oncogene, 26:2799-2803.

Sikand K, Slaibi JE, Singh R, Slane SD, Shukla GC (2010): miR 488* inhibits androgen receptor expression in prostate carcinoma cells. Int J Cancer.

Souaze, F., Ntodou-Thome, A., Tran, C.Y., Rostene, W., and Forgez, P., Quantitative RT-PCR: limits and accuracy. Biotechniques., 21(2): 280-285, 1996.

Sotiropoulou, G., Pampalakis, G., Lianidou, E., & Mourelatos, Z. (2009). Emerging roles of microRNAs as molecular swtiches in the integrated circuit of the cancer cell. RNA journal, 15(8), 1443-61.

Sotiropoulou, G., Pampalakis, G., Lianidou, E., & Mourelatos, Z. (2009). Emerging roles of microRNAs as molecular swtiches in the integrated circuit of the cancer cell. RNA journal, 15(8), 1443-61.

Spahn M, Kneitz S, Scholz CJ, Nico S, Rüdiger T, Ströbel P, Riedmiller H, Kneitz B. Expression of microRNA-221 is progressively reduced in aggressive prostate cancer and metastasis and predicts clinical recurrence. *Int J Cancer*, in press, doi: 10.1002/ijc.24715.

Stefani G, Slack FJ (2008): Small non-coding RNAs in animal development. Nat Rev Mol Cell Biol, 9:219-230.

Stephan C, Cammann H, Semjonow A, Diamandis EP, Wymenga LF, Lein M, Sinha P, Loening SA, Jung K (2002) Multicenter evaluation of an artificial neural network to increase the prostate cancer detection rate and reduce unnecessary biopsies. Clin Chem 48: 1279–1287

Stokowy, T., Eszlinger, M., Świerniak, M., Fujarewicz, K., Jarząb, B., Paschke, R. and Krohn, K. (2014). Analysis options for high-throughput sequencing in miRNA expression profiling. BMC Research Notes, 7(1): 144.

Strand M, Prolla TA, Liskay RM, Petes TD (1993): Destabilization of tracts of simple repetitive DNA in yeast by mutations affecting DNA mismatch repair. Nature 365:274-276.

Suh MR, Lee Y, Kim JY, Kim SK, Moon SH, Lee JY, (2004): Human embryonic stem cells express a unique set of microRNAs. Dev Biol,

270:488-498.

Sun T, Wang Q, Balk S, Brown M, Lee GS, Kantoff P (2009): The role of microRNA-221 and microRNA-222 in androgen-independent prostate cancer cell lines. Cancer Res, 69:3356-3363.

Szafranska-Schwarzbach AE, Adai AT, Lee LS, Conwell DL, Andruss BF (2011) Development of a miRNA-based diagnostic assay for pancreatic ductal adenocarcinoma. Expert Rev Mol Diagn 11(3): 249–257

Takeshita F, Patrawala L, Osaki M, Takahashi RU, Yamamoto Y, Kosaka N, Kawamata M, Kelnar K, Bader AG, Brown D, et al. Systemic delivery of synthetic microRNA-16 inhibits the growth of metastatic prostate tumors via down-regulation of multiple cell-cycle genes. *Mol Ther* 18(1):181-7, 2010.

Tanzer, A., Riester, M., Hertel, J., Bermudez-Santana, C. I., Gorodkin, J., Hofacker, I., (2010). Evolutionary Genomics of microRNAs and Their Relatives. In Evolutionary Genomics of microRNAs and Their Relatives, in

Evolutionary Genomics and Systems Biology (p. Chapter 15). Hoboken, NJ, USA.: John Wiley & Sons, Inc.

Tanzer, A., Riester, M., Hertel, J., Bermudez-Santana, C. I., Gorodkin, J., Hofacker, I., et al. (2010). Evolutionary Genomics of microRNAs and Their Relatives. In Evolutionary Genomics of microRNAs and Their Relatives, in Evolutionary Genomics and Systems Biology (p. Chapter 15). Hoboken, NJ, USA.: John Wiley & Sons, Inc.

Tarasov V, Jung P, Verdoodt B, Lodygin D, Epanchintsev A, Menssen A (2007): Differential regulation of microRNAs by p53 revealed by massively parallel sequencing: miR-34a is a p53 target that induces apoptosis and G1-arrest. Cell Cycle , 6:1586-1593.

Thiery JP, Acloque H, Huang RY, Nieto MA (2009): Epithelial-mesenchymal transitions in development and disease. Cell, 139:871-890.

Tong AW, Fulgham P, Jay C, Chen P, Khalil I, Liu S, Senzer N, Eklund AC, Han J, Nemunaitis J. MicroRNA profile analysis of human prostate cancers.

Cancer Gene Ther 16(3):206-16, 2009.

Usó, M., Jantus-Lewintre, E., Sirera, R., Bremnes, R. and Camps, C. (2014). miRNA detection methods and clinical implications in lung cancer. Future Oncology, 10(14): 2279-2292.

Valadi, H.; Ekstrom, K.; Bossios, A.; Sjostrand, M.; Lee, J.J.; Lotvall, J.O. (2007) Exosome-Mediated Transfer of mRNAs and microRNAs is a Novel Mechanism of Genetic Exchange between Cells. Nat. Cell Biol. , 9, 654–659.

van de Vijver MJ (2002) A gene-expression signature as a predictor of survival in breast cancer. N Engl J Med 347: 1999–2009

van Rooij, E. (2011) The art of microRNA research. Circulation Research, 108: 219-234.

Vandenboom Ii TG, Li Y, Philip PA, Sarkar FH (2008): MicroRNA and Cancer: Tiny molecules with major implications. Curr Genomics , 9:97-109.

Vella MC, Choi EY, Lin SY, Reinert K, Slack FJ (2004): The C. elegans microRNA let-7 binds to imperfect let-7 complementary sites from the lin-41 3'UTR. Genes Dev , 18:132-137.

Vergoulis T, Vlachos IS, Alexiou P, Georgakilas G, Maragkakis M, Reczko M, Gerangelos S, Koziris N, Dalamagas T, Hatzigeorgiou AG. (2012) TarBase 6.0: capturing the exponential growth of miRNA targets with experimental support., Nucleic Acids Res. ,40(Database issue) ,D222-9.

Vickers, K.C.; Palmisano, B.T.; Shoucri, B.M.; Shamburek, R.D.; Remaley, A.T. (2011) MicroRNAs are Transported in Plasma and Delivered to Recipient Cells by High-Density Lipoproteins. Nat. Cell Biol. , 13, 423–433.

Visvader JE, Lindeman GJ (2008): Cancer stem cells in solid tumours: accumulating evidence and unresolved questions. Nat Rev Cancer, 8:755-768.

Vlachos IS, Kostoulas N, Vergoulis T, Georgakilas G, Reczko M, Maragkakis M, Paraskevopoulou MD, Prionidis K, Dalamagas T, Hatzigeorgiou AG. (2012), DIANA miRPath v.2.0: investigating the combinatorial effect of microRNAs in pathways. Nucleic Acids Res. 40(Web Server issue):W498-504.

Volinia S, Calin GA, Liu CG, Ambs S, Cimmino A, Petrocca F, Visone R, Iorio M, Roldo C, Ferracin M, et al. A microRNA expression signature of human solid tumors defines cancer gene targets. *Proc Natl Acad Sci U S A* 103(7):2257-61, 2006.

Vrba L, Jensen TJ, Garbe JC, Heimark RL, Cress AE, Dickinson S (2010): Role for DNA methylation in the regulation of miR-200c and miR-141 expression in normal and cancer cells. PLoS One, 5:e8697.

Wach S, Nolte E, Szczyrba J, Stohr R, Hartmann A, Orntoft T, (2011): MicroRNA profiles of prostate carcinoma detected by multiplatform microRNA screening. Int J Cancer, 130:611-621.

Wang, E. (2008). An Overview of MicroRNA. In V. A. Erdmann, W. Poller, & J. Barciszewski, RNA Technologies in Cardiovascular Medicine and Research (pp. Chapter 1: 3-15). Berlin Heidelberg: Springer.

Wang, E. (2008). An Overview of MicroRNA. In V. A. Erdmann, W. Poller, & J. Barciszewski, RNA Technologies in Cardiovascular Medicine and Research (pp. Chapter 1: 3-15). Berlin Heidelberg: Springer.

Weigelt B, Hu Z, Livasy C, Carey LA, Ewend MG, Glas AM, Perou CM, Vant Veer LJ (2005) Molecular portraits and 70-gene prognosis signature are preserved throughout the metastatic process of breast cancer. Cancer Res 65: 9155–9158

Wightman B, Ha I, Ruvkun G (1993): Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans. Cell , 75:855-862.

Williams RJP, (2014) Binding, Transport and Storage of Metal Ions in Biological Cells, CHAPTER 1, Overview pg 1-5

Wong N and Wang X, (2015) miRDB: an online resource for microRNA target prediction and functional annotations. Nucleic Acids Research. 43(D1):D146-152.

Wu H, Mo YY. Targeting miR-205 in breast cancer. *Expert Opin Ther Targets* 13(12):1439-48, 2009.

Xu B, Niu X, Zhang X, Tao J, Wu D, Wang Z, (2011): miR-143 decreases prostate cancer cells proliferation and migration and enhances their sensitivity to docetaxel through suppression of KRAS. Mol Cell Biochem, 350:207-213.

Yi R, Qin Y, Macara IG, Cullen BR (2003): Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. Genes Dev, 17:3011-3016.

Yu F, Yao H, Zhu P, Zhang X, Pan Q, Gong C (2007): let-7 regulates self renewal and tumourigenicity of breast cancer cells. Cell, 131:1109-1123.

Yu Y, Kanwar SS, Patel BB, Oh PS, Nautiyal J, Sarkar FH (2012): MicroRNA-21 induces stemness by downregulating transforming growth factor beta receptor 2 (TGFbetaR2) in colon cancer cells. Carcinogenesis, 33:68-76.

Zeisberg M, Neilson EG (2009): Biomarkers for epithelial-mesenchymal transitions. J Clin Invest, 119:1429-1437.

Zernecke, A.; Bidzhekov, K.; Noels, H.; Shagdarsuren, E.; Gan, L.; Denecke, B.; Hristov, M.; Koppel, T.; Jahantigh, M.N.; Lutgens, E.; (2009). Delivery of microRNA-126 by Apoptotic Bodies Induces CXCL12-Dependent Vascular Protection. Sci. Signal. , 2, ra81.

Zhao, Y., & Srivastava, D. (2007). A developmental view of microRNA function. Trends in Biochemical Sciences, 32(4), 189-197.

Zhao, Y., & Srivastava, D. (2007). A developmental view of microRNA function. Trends in Biochemical Sciences, 32(4), 189-197.

Zheng C, Yinghao S, Li J (2011): MiR-221 expression affects invasion potential of human prostate carcinoma cell lines by targeting DVL2. Med Oncol.