UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

PAN AFRICAN CANCER RESEARCH INSTITUTE
DEFEATING CANCER TOGETHER

Molecular profiling of long non-coding RNAs in prostate cancer cell lines

Dissertation for the degree **MSc Medical Oncology** of the Department of Medical Oncology, Faculty of Health Sciences, University of Pretoria

Mandisa Pearl Mbeje

20785102

Supervisors

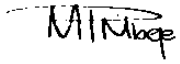
Dr Rahaba Marima
rahaba.marima@up.ac.za
Faculty of Health Sciences
Pan African Cancer Research Institute (PACRI)

Prof Zodwa Dlamini (co-supervisor)
zodwa.dlamini@up.ac.za
Faculty of Health Sciences
Pan African Cancer Research Institute (PACRI)

Dr Mmamoletla Kgoebane-Maseko (co-supervisor)
mw.kgoebane-maseko@up.ac.za
Faculty of Health Sciences
Pan African Cancer Research Institute (PACRI)

DECLARATION

I, the undersigned, hereby declare that the dissertation submitted herewith for the degree M.Sc. to the University of Pretoria, contains my own independent work and has not been submitted for any degree at any other university.



February 2023

DEDICATION

I dedicate this dissertation to:

God Almighty my creator, who is my source of knowledge and understanding.

To my parents, who have shown me endless love and support throughout my studies.

To my brother and sister, my friends, and my family, who have always encouraged me.

ACKNOWLEDGEMENTS

I would like to thank the following people and organizations for their assistance with this project:

- My supervisors, Dr Rahaba Marima, Prof Zodwa Dlamini and Dr Mmamoletla Kgoebane-Maseko for their invaluable guidance and insights.

- Ms Therese Dix-Peek and Prof. Clem Penny for their assistance and guidance.

- Demetra Demetriou and my fellow PACRI members for their support and encouragement.

- Julia and the FMG Programme for allowing me to use their QuantStudio™ 5 platform.

- Jeya, Elise and the Lasec and Bio-Rad team for their support.

- Dr Marima's UP Research Development Programme (RDP) grant, the National Research Foundation (NRF) of South Africa, and the South African Medical Research Council (SAMRC) for their assistance through grants.

PUBLICATIONS

Relevant to this study:

- Marima R, Hull R, Mbeje M, Molefi T, Mathabe K, Elbagory AM, Demetriou D, Dlamini Z. Role of Precision Oncology in Type II Endometrial and Prostate Cancers in the African Population: Global Cancer Genomics Disparities. International Journal of Molecular Sciences. 2022; 23(2):828. https://doi.org/10.3390/ijms23020628

- Marima R, Mbeje M, Hull R, Demetriou D, Mtshali N, Dlamini Z. Prostate Cancer Disparities and Management in Southern Africa: Insights into Practices, Norms and Values. Cancer Management and Research. 2022;14:3567-3579. https://doi.org/10.2147/CMAR.S382903

Additional publications:

- Khanyile, R., Marima, R., Mbeje, M., Mutambirwa, S., Montwedi, D., Dlamini, Z. (2023). AI Tools Offering Cancer Clinical Applications for Risk Predictor, Early Detection, Diagnosis, and Accurate Prognosis: Perspectives in Personalised Care. In: Dlamini, Z. (eds) Artificial Intelligence and Precision Oncology. Springer, Cham. https://doi.org/10.1007/978-3-031-21506-3_15

CONFERENCE POSTER PRESENTATIONS

- Published American Association for Cancer Research (AACR) abstract.
  Title: Telomerase RNA component (TERC) lncRNA hypermethylation regulates hsa-miR-320 family and promotes prostate cancer metastasis.
  Authors: Mandisa Mbeje, Jeyalakshmi Kandhavelu, Clement Penny, Zodwa Dlamini, Rahaba Marima.

- Presented and published PACRI Inaugural International Cancer Meeting abstract.
  Title: Long non-coding RNAs as prognostic markers in prostate cancer.
  Authors: Mandisa Mbeje, Jeyalakshmi Kandhavelu, Clement Penny, Zodwa Dlamini, Rahaba Marima.

Abstract

Prostate cancer (PCa) incidence rates continue to rise with an estimation of 375 000 deaths and 1.4 million new cases in the year 2020. It has become the second most commonly occurring cancer and the fifth most frequent cause of cancer-related deaths in men. Prostate cancer is also the leading cancer to occur in men in the sub-Saharan region. In PCa, the aberrant expression of certain RNA molecules has been linked to PCa progression. Non-coding RNAs (ncRNAs) are RNA transcripts that are known for their involvement in numerous cellular processes. They are important in tumorigenesis and their presence may indicate cancer progression stages. Long non-coding ribonucleic acids (lncRNAs) are transcripts with 200 or more nucleotides. The focus on these lncRNAs can be attributed to their involvement in each of the cancer "hallmarks". There have been multiple studies that have successfully identified lncRNAs in PCa. However, the landscape of the enriched molecular pathways targeted by these lncRNAs remains to be elucidated. Thus, this study aimed to profile aberrantly expressed lncRNAs and map the biological pathways associated with PCa progression. The differential expression patterns of 86 lncRNAs were compared between PC-3 and LNCaP PCa cell lines, which are highly metastatic and low metastatic (high grade and low grade), respectively. This was achieved by comparing the expression of lncRNAs between PC-3 and LNCaP cells using a 384-well plate of PCa lncRNA gene panel. LncRNA differential gene expression patterns were determined using the CFX Maestro software. The lncRNAs with a ±2 up or downregulation were considered to be differentially expressed. Annotation and enrichment analysis of lncRNA differential gene expression was performed using a human lncRNA sets database, LncSEA and DIANA-miRPath. Based on the PrimePCR array results and bioinformatics analysis, TERC lncRNA was selected for validation. The results of the array were validated with RT-qPCR using TERC primers. Thirty-six (36) of 86 lncRNAs were shown to be upregulated, while 12/86 of the lncRNA panel were downregulated. Bioinformatics analysis revealed that the upregulated lncRNAs are involved in various hallmarks of cancer. Interestingly, HOTAIR and TUG1 were demonstrated to be exosomal lncRNAs. In addition, the hypermethylation of HOTAIR and TERC was shown to sponge various miRNAs, promoting tumour progression. Furthermore, the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis and the Gene Ontology (GO) analysis revealed several cancer-related pathways, namely the signalling pathways regulating pluripotency of stem cells and the TGF-beta signalling pathway. Cytoscape ceRNA network analysis further illustrated the lncRNA/miRNA/mRNA axis in PCa progression. LncRNAs may be prognostic biomarkers as they have been reported to be stable in liquid and solid biopsies.

TABLE OF CONTENTS

List of Abbreviations

ADT	Androgen deprivation therapy
AFG3L1P	AFG3-like AAA ATPase 1, pseudogene
AJCC	American Joint Committee on Cancer
AR	Androgen receptor
arcRNA	Architectural RNA
B2M	Beta-2-microglobulin
BPH	Benign prostatic hyperplasia
cDNA	Complementary Deoxyribonucleic Acid
CDK6	Cyclin-dependent kinase 6
ceRNA	Competing endogenous RNA
CFLAR-AS1	CFLAR antisense RNA 1
COX6CP10	Cytochrome c oxidase subunit 6C pseudogene 1
Cq	Quantification cycle
CRPC	Castration-resistant prostate cancer
CTBP1-AS	Carboxyl terminal binding protein 1 antisense RNA
DBC1	Depleted in breast cancer 1
DEGs	Differentially expressed genes
DMEM	Dulbecco’s modified Eagle’s medium
DNA	Deoxyribonucleic acid
DRAIC	Downregulated RNA in cancer
DRE	Digital rectal examination
DUTP1	Deoxyuridine triphosphatase pseudogene 1
ECM	Extracellular matrix
EMT	Epithelial-mesenchymal transition
FAM83H-AS1	FAM83H antisense RNA 1 (head-to-head)
FBS	Fetal bovine serum
FDR	Fold discovery rate
FOXP4-AS1	FOXP4 antisense RNA 1
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
GAS5	Growth arrest-specific transcript 5

gDNA	Genomic DNA
GO	Gene Ontology
GOLPH3	Golgi phosphoprotein 3
GTF3AP2	General transcription factor IIIA pseudogene 2
GWAS	Genome-wide association studies
HICs	High-income countries
HMBS	Hydroxymethylbilane synthase
HOTAIR	HOX antisense intergenic RNA
KEGG	Kyoto Encyclopedia of Genes and Genomes
KRT81	Keratin 81
LBCS	LncRNA bladder and prostate cancer suppressor
LINC00339	Long intergenic non-protein coding RNA 339
LINC00493	Small integral membrane protein 26
LINC00657	Long intergenic non-protein coding RNA 657
LINC01024	Long intergenic non-protein coding RNA 1024/ mitosis associated long intergenic non-coding RNA 1
LINC01137	Long intergenic non-protein coding RNA 1137
lincRNA	Long intergenic non-coding RNA
LMICs	Lower middle-income countries
LNCaP	Lymph Node Carcinoma of the Prostate
lncRNA	Long non-coding RNA
LncSEA	Long non-coding RNA related Sets and Enrichment Analysis
LURAP1L-AS1	LURAP1L antisense RNA 1
MAGI1-IT1	MAGI1 intronic transcript 1
MALAT1	Metastasis-associated lung adenocarcinoma transcript 1
MEG3	Maternally expressed gene 3
miRNA	MicroRNA
mRNA	Messenger RNA
Orα	Oestrogen receptor alpha
NEAT1	Nuclear enriched abundant transcript 1
NICE	National Institute for Health and Care Excellence
PC-3	Prostate Cancer 3

PCa	Prostate cancer
PCA3	Prostate cancer antigen 3
PCAT1	Prostate cancer-associated transcript 1 (non-protein coding)
PCAT5	Prostate cancer-associated transcript 5 (non-protein coding)
PCAT29	Prostate cancer-associated transcript 29 (non-protein coding)
PCGEM1	Prostate-specific transcript 1(non-protein coding)
PRNCR1	Prostate cancer-associated non-coding RNA 1
PSA	Prostate-specific antigen
PTEN	Phosphatase and Tensin Homolog deleted on Chromosome 10
PTPRJ-AS1	PTPRJ antisense RNA 1
qPCR	Quantitative PCR
RAB14	RAB14, member RAS oncogene family
RAB23	RAB23, member RAS oncogene family
RNA	Ribonucleic acid
RNASEL	Ribonuclease L
RNF139-AS1	RNF139 antisense RNA 1 (head-to-head)
rRNA	Ribosomal RNA
RT	Reverse transcription
SChLAP1	Second chromosome locus associated with prostate 1
SEM	Standard error of the mean
siRNA	Small-interfering RNA
SNHG1	Small nucleolar RNA host gene 1
SNHG8	Small nucleolar RNA host gene 8
SNPs	Single nucleotide polymorphisms
SOCS2-AS1	SOCS2 antisense RNA 1
TBP	TATA box binding protein
TERC	Telomerase RNA component
TINCR	Terminal differentiation-induced non-coding RNA
TMEM106B	Transmembrane protein 106B
TNM	Tumour, node, and metastasis
TP53TG1	TP53 target 1 (non-protein coding)
tRNA	Transfer RNA

TUG1	Taurine upregulated 1 (non-protein coding)
USA	United States of America
Xist	X-inactive specific transcript

List of Figures

List of Tables

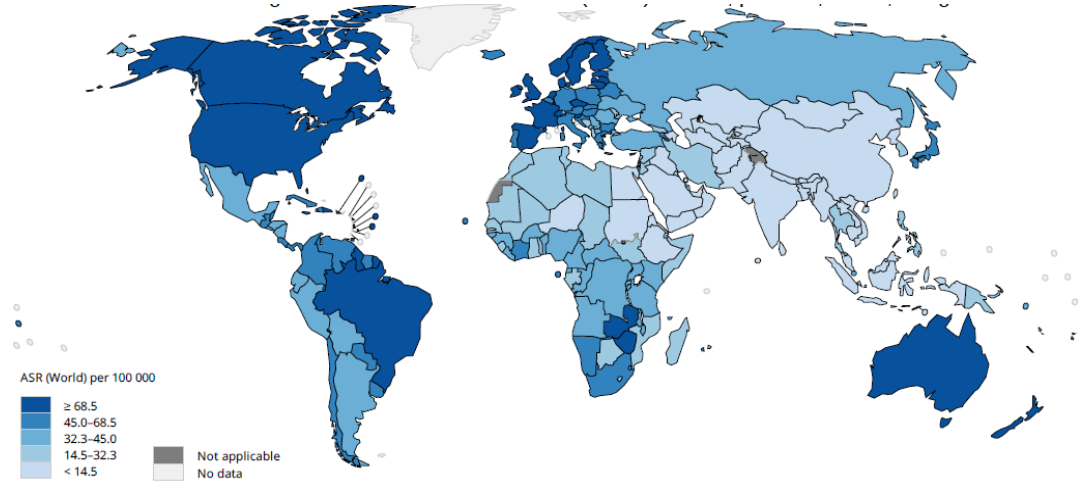
CHAPTER 1

INTRODUCTION

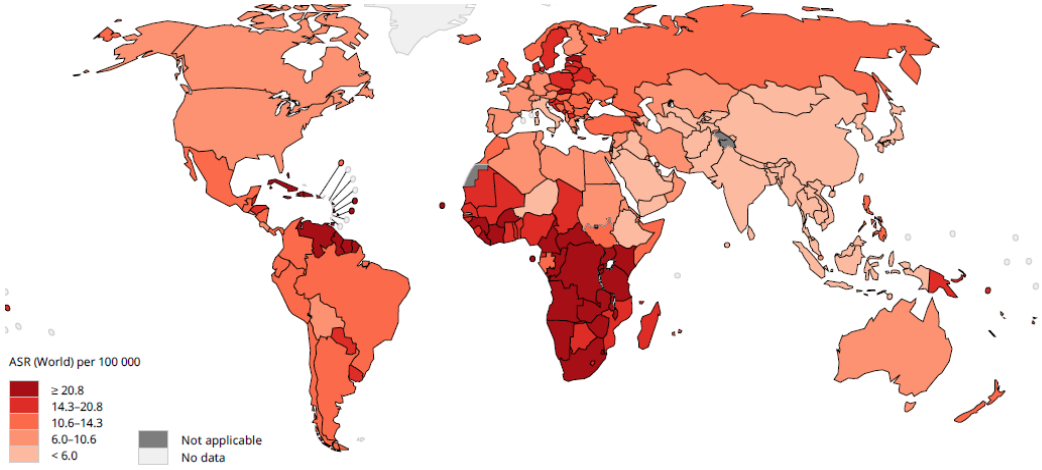
1.1. Introduction

In 2020, GLOBOCAN reported global prostate cancer (PCa) statistics. It was reported that there were an estimated 1.4 million new PCa cases, and 375 000 deaths occurred.[1] On a global scale, the epidemiology of long-term PCa shows great variability.[2] This is supported by the GLOBOCAN cancer statistics of 2020, which showed that incidence rates vary across different regions from 6.3 to 83.4 per 100 000 men. The incidence rate in developed countries is three times higher but death rates do not follow this same pattern of variability.[1] Some of the regions with the highest PCa incidence are Western and Northern Europe, North America, Southern Africa and the Caribbean (Figure 1.1). It is known that men of African descent are at a higher predisposition to developing PCa.[3] This, as well as other factors including medical care accessibility, are thought to be contributing factors to the trend seen in the epidemiology of PCa in sub-Saharan Africa.[2] Both genetic and non-genetic factors are associated with PCa such as genetic polymorphisms and age, respectively.[4]

A. Global estimated age standardized PCa incidence rates in 2020.



B. Global estimated age standardized PCa mortality rates in 2020.



C. Mortality and incidence rates of PCa in 2020 for selected countries.

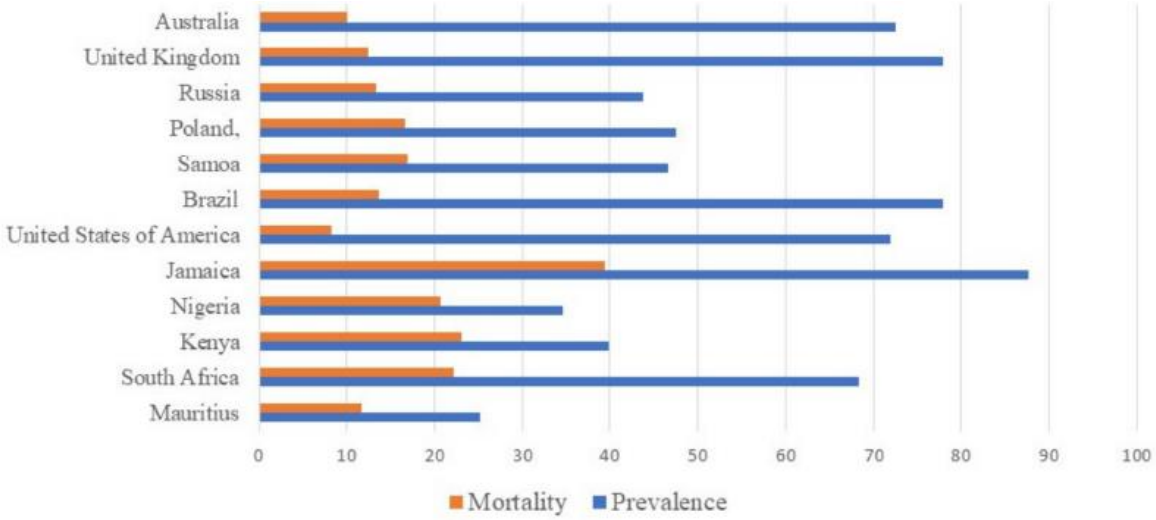


Figure 1.1: GLOBOCAN PCa epidemiology statistics for 2020. (A) global map for PCa showing the age-standardized PCa incidence rates. The highest rates were observed in North America, Northern and Western Europe, Southern Africa and the Caribbean. (B) Global map for PCa showing the world standardized mortality rates. The regions with the highest rates were the sub-Saharan African regions. (C) Bar graph showing the incidence and mortality rates in selected countries. All the data lies within the 95% confidence interval.[5]

Not more than 2% of the entire human genome has been shown to encode proteins.[6] Specifically, the non-coding RNAs (ncRNAs) have been implicated in various tumours, including PCa.[6] Non-coding RNAs can be classed into small and long non-coding RNAs, depending on their nucleotide length. These include short ncRNAs such as small-interfering RNA (siRNA) molecules, and translation-related ribosomal RNA (rRNA) and transfer RNA (tRNA).[7-8] They perform several functions including chromatic remodelling and regulation of signalling pathways. These molecules vary in length, but those that are >200 nucleotides long are considered as long non-coding RNAs (lncRNAs).[9-12] X-inactive specific transcript (Xist), which is necessary for X-chromosome inactivation was the first lncRNA to be discovered and it is a transcript that is 17 kb in length.[13] Since then, more than 100 000 novel lncRNAs have been discovered in the human genome alone.[14] In recent years, lncRNAs have seized the attention of researchers. This is because they have emerged as crucial factors in various cellular processes.[15]

Normal physiological functioning of lncRNAs includes development, cell differentiation, and the cell cycle.[16] They are also involved in the allelic expression regulation in processes such

as gene imprinting,[8] X-chromosome inactivation, and dosage compensation.[8,13] They have functions at all levels of gene expression[14] such as transcription, post-transcription and translation. The mechanisms in which they function include post-transcriptional modification, transcriptional co-activation, chromatin remodelling, and protein inhibition.[17] They are often found associated with histone proteins (H3K4me3 or H3K36me3), which are active chromatin epigenetic marks, confined to the promoter region or found throughout the gene respectively.[12] These RNA molecules are usually transcribed by RNA polymerase II,[12,18] they exhibit splicing of exons, and they have poly-A tails.[12]

LncRNAs can control gene expression and cancer cell fate through chromatin remodelling,[6] signal transduction, transcriptional and post-transcriptional modifications.[18-19] They do this by forming a network of interactions acting as tumour suppressors or oncogenes. LncRNAs also have tissue and cell-type specific expression[7,13] and are highly dysregulated in cancer. When compared to messenger RNA (mRNA), lncRNAs experience less expression (± 0.01% of total RNA).[7,19-20] These RNA molecules can be profiled with mRNAs since they share many common characteristics such as a 5' cap, a poly(A) tail, and RNA polymerase II transcription[19] and splicing.[6] LncRNAs can be sense (e.g. steroid receptor RNA activator) (SRA)),[21] antisense (e.g. HOX antisense intergenic RNA (HOTAIR)), bidirectional (e.g. Cytoplasmic Endogenous Regulator Of Oxidative Phosphorylation 1 (Cerox1)),[22] intronic (e.g. cold assisted intronic noncoding RNA (COLDAIR)),[21] and intergenic (e.g. XIST)[22] depending on their proximity to a protein-coding region. These molecules regulate gene expression via their interactions (both direct and indirect) with other RNAs, proteins, and chromatin.[23] Moreover, the aberrant expression of lncRNAs is implicated in several diseases, including PCa.

The second most diagnosed cancer is PCa, after lung cancer, among men.[24] It is, however, the leading cancer in the sub-Saharan region.[25] In normal tissues, lncRNAs are expressed at low levels, while in cancer cells, they exhibit dysregulated expression.[26] In the transcriptome, several lncRNAs were shown to be directly involved in PCa including SChLAP1. The overexpression of this lncRNA is associated with two-and-a-half times increased risk of metastatic progression in prostate cancer patients that have undergone a radical prostatectomy.[27]

In this project, an 86 lncRNA PCR gene array panel was used for profiling and identification of the differentially expressed genes in advanced metastasis (PC-3) cells and low metastasis (LNCaP) cells. PC-3 cells generally grow and spread at a faster rate, like high-grade tumours while LNCaP cells have low metastatic potential, similar to low-grade PCa tumours.[28] The lncRNA expression profiles of the PC-3 cell line were compared to that of the

LNCaP cell line. RT-qPCR was used to validate the PrimePCR Array results. These methods have been widely used in gene expression studies as they are reliable and cost-effective. The differentially expressed lncRNAs were further studied using *in silico* bioinformatics tools.

The progression of PCa modulated by protein-encoding genes is an advanced area due to the abundance of proteomics data, compared to PCa modulation by ncRNAs. There exist significant knowledge gaps in the field of gene regulation by ncRNAs, specifically lncRNAs. Although lncRNAs have little to no protein-coding potential, they hold great potential to be innovative diagnostic and prognostic tools.[29] Their versatile mechanisms of action in various biological processes make them key regulators of gene expression. These are yet to be fully elucidated and continuous efforts are required to achieve this.

1.2. Aim

To profile the lncRNAs in high-grade metastatic PC-3 cells compared to low-grade metastatic LNCaP cells.

1.3. Objectives

1. To determine the differential expression patterns of lncRNAs in advanced PCa PC-3 cells compared to low-grade metastatic PCa LNCaP cells achieved by using the PrimePCR™ 384 well plate Prostate Cancer lncRNA Array.
2. To map the molecular and biological pathways enriched by the differentially expressed lncRNAs identified in objective 1 using bioinformatics tools. Human lncRNA sets database, LncSEA and DIANA-miRPath v3.0 were used.
3. To validate the PrimePCR Array results, using RT-qPCR.

CHAPTER 2

LITERATURE REVIEW

2. Literature Review

2.1. Prostate cancer epidemiology, screening, and diagnosis

The third most common cancer to occur is PCa, worldwide.[1] GLOBOCAN stated in 2020 that new PCa cases reached 7.3% of all cancer sites. Of these sites, 3.8% of cancer deaths resulted from PCa. The variability of PCa epidemiology can be seen across different regions, with the sub-Saharan region showing some of the highest rates of PCa incidence and mortality (Figure 1.1).[1,30] Incidence rates were highest in the Western and Northern regions of Europe, the Northern regions of America, New Zealand, the Caribbean, Southern Africa, Australia, and the lowest rates occurred in the Northern regions of Africa and Asia. In 48 countries, including several sub-Saharan countries, PCa was the most common cause of cancer deaths. Between 1995 and 2018, incidence rates have increased dramatically in countries like South Africa, Mozambique and Kenya where an annual increase of between 2-10% has occurred.[25] It remains unclear what has caused this increase but it is believed to be due to improving healthcare and awareness of the disease in these countries, which has led to an increase in PSA tests being done.[25]

The average size of a prostate gland is 28-47 cm^3 and it is located in the lower pelvis, underneath the bladder, and around the beginning of the urethra.[31] The prostate is in the lower pelvis, underneath the bladder, and around the beginning of the urethra. It secretes an alkaline fluid which forms part of the seminal fluid that nourishes and assists sperm cells with their motility.[31-32] Anatomically, the prostate has four zones, namely the fibromuscular, transition, central and peripheral zones. The largest zone is the peripheral one which makes up 75% of the gland and about 75% of PCas develop in this zone. The transition zone is where benign prostate hyperplasia (BPH) usually develops (Figure 1.2).

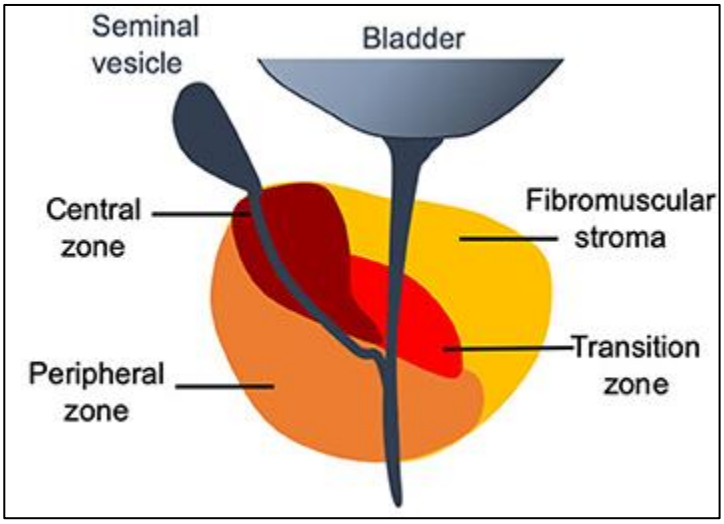


Figure 1.2: Diagram of the prostate showing the four zones of the gland. The central zone (dark red) is the smallest. BPH primarily originates in the transition zone (red) and the fibromuscular zone or stroma (yellow) is composed of non-glandular tissue. The peripheral zone (orange) is the most common site of PCa occurrence.[33]

Several men who are diagnosed with PCa do not show symptoms until the tumour had progressed to the point where it disrupts normal prostate function.[24] The common symptoms for PCa are erectile dysfunction, lower urinary tract symptoms (LUTS), visible haematuria incontinence, hesitancy and incomplete emptying of the bladder,[34] and are often also seen in benign conditions that affect the prostate such as prostatitis and BPH. This presents a challenge for clinicians to establish whether the patient is symptomatic for PCa or other conditions.[24]

In PCa, early detection and diagnosis may be a pivoting factor in the treatment and overall prognosis of the disease. This is because the cancerous tumour is easier to treat if detected early and before it has metastasized. The process of diagnosis begins with counselling of the patient.[24] Two main methods are then used by doctors to screen for PCa, which include the determination of prostate-specific antigen (PSA) levels in the blood and the use of the digital rectal exam (DRE). The PSA has been a widely used biomarker for the early detection of PCa since the 1980s.[35] It is produced by the prostate gland and it forms part of the seminal fluid.[36] This serine protease belongs to the Kallikrein family, and its testing has faced several controversies as it is known to yield false positive and negative results.[24] The doubt on the reliability of the PSA test is because other disorders may result in elevated levels of PSA such as BPH, prostatitis and trauma.[24,31] These levels may be lowered, for example, due to surgeries performed on the prostate, or androgen deprivation therapy (ADT).[31] Although it allows for PCa to be diagnosed early in asymptomatic patients, other innovative methods are still needed for PCa testing.[31] Normal PSA levels are between 0 to 4.0 ng/mL. The PSA values of $\geq$ 4 ng/mL are considered to be abnormal and are followed up by other tests.[37] The PSA guidelines have been revised to be more age-specific so that men at high risk for PCa development can be identified.[38] The National Institute for Health and Care Excellence (NICE) revised the PSA thresholds to: PSA= 3 ng/mL for men ages 50-59, PSA= 4 ng/mL for men aged 60-70, and PSA=5 ng/mL for men above the age of 70.[38-39]

The development of a mechanism that will differentiate between clinically significant and insignificant PCa is necessary. The current tests used such as the PSA and the DRE test need to be augmented by innovative approaches as they produce significant false and false positive results.[24] For the diagnosis of PCa, a prostate needle biopsy is done[31] after thorough communication between the patient and the practitioner. During the biopsy, a small

portion of the gland is removed using a transrectal or transperineal ultrasound-guided needle biopsy.[24,40] This biopsy is then viewed under the microscope and histologically determined to be cancerous or not. If the biopsy indicates PCa, this is then useful for prognosis. The tumour is graded using a Gleason score and the tumour, node, and metastasis (TNM) system, which was developed by the American Joint Committee on Cancer (AJCC).[31] Adenocarcinomas are the frequently occurring type of PCas accounting for 95% of the cases.[31] Treatment of PCa is heavily dependent on the disease stage. The age of the patient and co-morbidities are also considered. Guidelines do exist for the treatment and management of the disease. Treatment options for localized disease progression include radical prostatectomy where the seminal vesicles and the prostate are removed.[31] Radiation therapy, brachytherapy and chemotherapy (advanced PCa) are among other treatment options for PCa.[31] Figure 1.1 demonstrates the age-standardized PCa incidence rates and mortality rates for 2020.

2.2. Screening and diagnosis differences between lower-middle-income countries and high-income countries

The cost of healthcare has contributed immensely to the incidence and death rates of various cancers, including PCa. High-income countries (HICs) have numerous cancer screening programs in place because they have access to better resources.[41] Countries such as the United States of America (USA), Canada and Australia are considered to be HICs, and they experienced an influx of PCa incidence in the late 1980s through to the early 1990s. This was greatly due to the introduction of PSA testing.[42] A reduction in incidence rates was observed during the late 2000s[41] as changes in screening recommendations[42] and PSA testing decreased. Lower middle-income countries (LMICs) such as Mozambique, Kenya and Zimbabwe saw an opposing trend. In these countries, between the mid-1990s and 2018, incidence rates of PCa have been increasing, which can be attributed to more PSA tests being done due to improving healthcare systems.[25] Resource availability is linked to screening and treatment option availability.[43] The precise severity of the burden of PCa on the economies of LMICs cannot be determined accurately because of a lack of evidence.[44] These countries have limited funds dedicated to screening and diagnosis tools, the knowledge on PCa is limited, and it may be for these reasons that the mortality rates are high due to the late detection and treatment.

Certain states in the USA have ensured that health insurers (private) cover tests such as DRE and PSA by putting laws in place that make this obligatory. It is also claimed by several states that such tests are provided to public employees as part of the healthcare benefits. This is evidence that the public has more available resources for the early diagnosis and

treatment of PCa in HICs. On the contrary, many people in LMICs cannot afford health insurance, and neither do they receive health care benefits. They would have to cover the great costs of screening, detection and diagnosis costs themselves. This would explain the higher PCa mortality rates in these countries as early tumour detection and treatment option expenses are not readily available. Moreover, in LMICs such as South Africa, there is a multi-level referral system that PCa patients undergo before consulting with an oncologist. This system involves a consultation at a local clinic or primary healthcare centre, a consultation at a secondary healthcare facility such as a district hospital and then at a tertiary facility with oncology services.[38] This process often delays early diagnosis and treatment and limits treatment options which are all contributing factors to cancer treatment outcomes.

Men of African origin have a higher predisposition for PCa development.[3] Many of these men are at a greater disadvantage because they tend to reside in LMICs. The predisposition, with the limited screening and diagnosis resources, make men of African descent more vulnerable to PCa mortality than their HIC counterparts.

2.3. Prostate cancer staging

A critical aspect of PCa prognosis and treatment planning is accurate and uniform staging that allows for tumour behaviour and treatment selection to be predicted efficiently. First introduced in 1922, the TNM staging system has become the accepted and uniform staging system for PCa by the AJCC. Since then, it has been updated continuously.

To predict patient prognosis, both clinical and pathological staging is necessary to access the extent to which the tumour has spread.[45] Clinical staging is done based on the results from PSA testing, DRE and the Gleason score. Pathological staging, on the other hand, is based on the laboratory results from the surgically removed tissue and histology to identify the extent of the tumour within the prostate gland and the tissues surrounding it.[45]

The TNM staging system developed by the AJCC asks three questions. How large the tumour is and where it is located (T). If it has spread into the lymph nodes and the number of lymph nodes it has spread to (N) and if other parts of the body have been affected (M). The results are then combined, determining the stage. The five stages of cancer are stage 0, I to IV, and each stage gives doctors a common description of cancer.

The Gleason score is also used to grade PCa. It is based on how the tumour looks under a microscope and how similar it looks to normal healthy tissue. A score on a scale of 3 to 5 is given based on the arrangement of the cancerous cells in the prostate in two different locations. To determine the number, the pathologist looks for the main cell growth pattern

and the cancerous cells that are similar to normal cells get a low score while cells that look different to healthy cells receive a higher score. Thereafter, scores are combined to achieve an overall Gleason score between 6 and 10. A Gleason score of 6 or lower is low-grade cancer where cell differentiation is satisfactory, and the cells look similar to normal cells. A Gleason of 7 is given to the moderately differentiated cells and this is medium-grade cancer. A Gleason above 7 is defined as high-grade cancer where the cells are poorly or undifferentiated. In cancer stage grouping, the T, N, and M classifications are combined with the Grade Group and PSA levels. Table 1.1 summarizes the various stages of cancer and the description of what tumours fall under which stage.

Table 1.1: Description of the tumour stages and the characteristics that define them.

AJCC Stage	**Stage grouping**	**Stage description**
Stage I	Stage I	Early-stage cancer with cancer growing slowly. The tumour may not be felt and can be found on half of one side or less of the prostate, The cancerous cells look like healthy cells and the PSA level is low.
Stage II	Stage IIA	The tumour may not be felt and can be found on less than half of one side of the prostate gland. Cancerous cells are well differentiated but the PSA levels are below 20.
	Stage IIB	Tumour is found on one side or both sides of the prostate and it is large enough to be identified during DRE. There is moderate cell differentiation and PSA levels are medium.
	Stage IIC	Tumour on one side of the prostate. Tumour is detectable during DRE; PSA levels are medium but there is moderate to poor cell differentiation.
Stage III	Stage IIIA	Cancer has spread into neighbouring tissues. PSA level is high.
	Stage IIIB	PSA levels are high. The tumour has spread to neighbouring tissues like the rectum or bladder.
	Stage IIIC	Cancerous cells are poorly differentiated, and the tumour has spread even further.
Stage IV	Stage IVA	The tumour had spread into the lymph nodes near the prostate.
	Stage IVB	The tumour has spread into further lymph nodes and other body parts.

2.4. *In vitro* prostate cancer study models

Human cancer cell lines are often used in research to understand the biology of cancer and to test the effectiveness of certain anticancer therapeutic agents.[46] In PCa specifically, the most commonly used cell lines are Lymph Node Carcinoma of the Prostate (LNCaP) and Prostate Cancer 3 (PC-3).[47] LNCaP cells are low metastatic cells that were originally derived from a lymph node of an individual with metastatic PCa.[48] They are sensitive to androgens and they express androgen receptors (ARs) and PSA. Moreover, similar to human prostate adenocarcinoma, their growth can be inhibited by androgen withdrawal. PC-3 cells were derived from bone metastasis of human PCa.[47] These cells have the highest metastatic potential and they are androgen-independent.[49] Furthermore. PC-3 cells do not express AR or PSA.[28] An AR binds androgens which is essential for the development and maintenance of a normal prostate.[50] The AR signalling pathway is necessary for PCa cell growth and ADTs are the main way to treat PCa initially.[51] When PCa cells become androgen-independent, they no longer rely on androgens for growth and the implication of this is resistance to ADTs.

2.5. Risk factors

Age, ethnicity, family history, genetic mutations such as mutations in breast cancer genes 1 and 2 (*BRCA1* and *BRCA2)*, as well as certain conditions like Lynch syndrome are common factors that increase the risk of developing PCa.[1] Age, the primary risk factor, increases the risk of a man developing PCa rapidly after age 50. An estimated 75% of all PCas are diagnosed in males above the age of 65 and disease occurrence and diagnosis is rare in men under 40.[31] According to the Office for National Statistics data from the United Kingdom, from the age of 50, age-specific incidence rates increase and reach a peak in men who are 90 years old or older.[24] Since the average life expectancy has increased, more PCa cases will arise, which is a significant health concern.[31]

In a study by Jaratlerdsiri et al.[52], whole-genome sequencing of African against European treatment-naïve PCa samples was done. Their analysis identified around 2 million somatic variants and an increased tumour mutation burden and mutational signatures linked to African ancestry. The study revealed several driver genes including PCAT1, and also suggests that both genetic and environmental factors are contributors to the disparities observed in PCa between the ethnicities.[52] Race has been identified as the second highest risk factor. It is known that black men in the Caribbean and the USA, who are of Western African descent, are the most affected group with the highest incidence rates worldwide.[3] Although the significance of race in the development and incidence of PCa is understudied, there may be a correlation with dietary, environmental, and socioeconomic factors.[1,31] A

2022 study by Gong et al.[53] aimed to further elucidate the ethnicity-based PCa disparities. To do this, these authors studied the relationship and landscape of somatic structural variations driving ethnic disparity to identify the potential role of somatic structural variations that contribute to aggressive PCa in men of African ancestry.[53] In this study, they found several gene duplication events in the African-derived tumours. Moreover, a spectrum of other structural variations like deletions and hyper-translocations were identified in these tumours, which shows a mechanism for the ethnic disparities that are observed in aggressive PCa.

Having a family history of PCa dramatically increases the risk of disease development. In a study done by Kiciński et al.[54], it was estimated that a man with a first-degree relative (a father or a brother) who has been diagnosed with PCa has a risk of two times (95% confidence interval (CI): 2.25-2.74). Males with an affected brother, have three times (95% CI: 2.37-4.15) increased risk and for men with a father diagnosed, two times (95% CI: 2.02-2.72). In men with 2+ first-degree relatives who have been diagnosed with PCa; this risk rises to four (95% 2.61-7.39).

Evidence from previous studies has shown the relationship between PCa and germline variations.[24,55] It was hypothesized that these variations may influence the expression of genes in PCa tissue. Genome-wide association studies (GWAS) have shown >100 PCa risk-related single nucleotide polymorphisms (SNPs).[24] In a study by Alvarez-Cubero et al[55], the association between several SNPs, environmental factor exposure, expression patterns and PCa were investigated. They focused on genes that were associated with PCa and its severity such as ribonuclease L (RNASEL), macrophage scavenger receptor 1 (MSR1) and elaC ribonuclease Z 2 (ELAC2), which are involved in DNA damage response pathways.[55] RNASEL, for example, is involved in cell apoptosis, differentiation, and tumourigenesis as well as interferon-regulated antiviral and inflammatory responses. A study has shown that variants like rs486907 are linked to heritable PCa while rs635261 and rs627839 are linked to fatal PCa. The study also showed that tumour progression can be promoted by the interaction between RNASEL and AR.[55] Moreover, the interaction between RNASEL and AR has been further studied by Dayal et al.[56] The authors reported that cell migration and AR signalling were suppressed by RNASEL in PCa cell lines. In the study, various protein-based experiments such as immunoblotting followed by Western blotting were used to show that RNASEL negatively regulated AR signalling. Furthermore, the knockdown of RNASEL restored and upregulated AR signalling and cell migration.[56]

LncRNAs are multi-functioning molecules that are implicated in various cancer types including PCa. These include epigenetic and gene expression regulation functions.[57-58] The

aberrant expression of these molecules can signify stages of PCa progression and maintain signalling pathways that are tumour-related.[29,57] These non-coding RNA molecules, therefore, have tremendous potential as biomarkers. They may be used for diagnostic and prognostic purposes, they may also serve as potential targets for PCa therapies.[29]

2.6. Long non-coding RNAs (lncRNAs)

2.6.1. Characteristics

Only approximately 2% of the mammalian genome has protein-coding potential. The remaining 98% represents non-protein encoding genes, such as pseudogenes and transposons. Some of these genes are transcribed into RNA that is called ncRNA, which in the past was considered as “transcriptional noise” or “junk DNA”.[7-9,14] This noise comprises of pseudogenes, transposons, and other simple repeats.[59] NcRNAs are divided into two types based on how often they are expressed, namely constitutive and regulatory types.[60-61] Non-coding RNAs can be classed based on length into short, mid, and long non-coding RNAs, ranging from 19-31, 20-200 and >200 nucleotides, respectively.[19,62] The short non-coding RNAs include microRNA (miRNA), small nuclear RNA (snRNA), rRNA, tRNA, miscellaneous other RNA (miscRNA) small nucleolar RNA (snoRNA), and siRNA molecules.[7-8] LncRNAs account for the majority of non-coding transcripts, examples of these are antisense RNA (asRNA) and long intergenic non-coding RNA (lincRNA). Long non-coding RNAs can even be longer than 100 000 nucleotides and can be further divided into intronic, intergenic, sense, bidirectional, and antisense depending on their proximity to the closest protein-coding gene[14,16] (Figure 1.3). Intronic lncRNAs are those that exist within introns of protein-encoding genes. Intergenic lncRNAs are those that exist between two protein-coding gene sequences.[14] Sense lncRNAs are those over one or many exons in a protein-encoding sequence on the sense strand while antisense lncRNAs are those over one or more protein-encoding gene sequences on the non-coding (antisense) strand.[14] Bidirectional lncRNAs are those that are transcribed opposite to the transcriptional starting point of a transcript.[14,16,22] Figure 1.3 illustrates the different lncRNA types based on their proximity to coding regions.

LncRNAs often exhibit *cis* and *trans* gene regulation. Their mode of function is usually determined by the chromosomal region of the genes that they control. *Cis*-acting lncRNAs regulate genes on one chromosome or in the same region of a chromosome while *trans*-acting lncRNAs regulate genes at a distance, located on a different chromosome. In either case, it is their interaction with DNA elements that results in their ability to regulate these genes.[63] Based on their functions and molecular mechanisms, lncRNAs can be divided into four main archetypes; namely, signals, decoys, guides and scaffolds, which form the four

primary functions of lncRNAs.[18] These types are not mutually exclusive since several lncRNAs can perform more than one function.

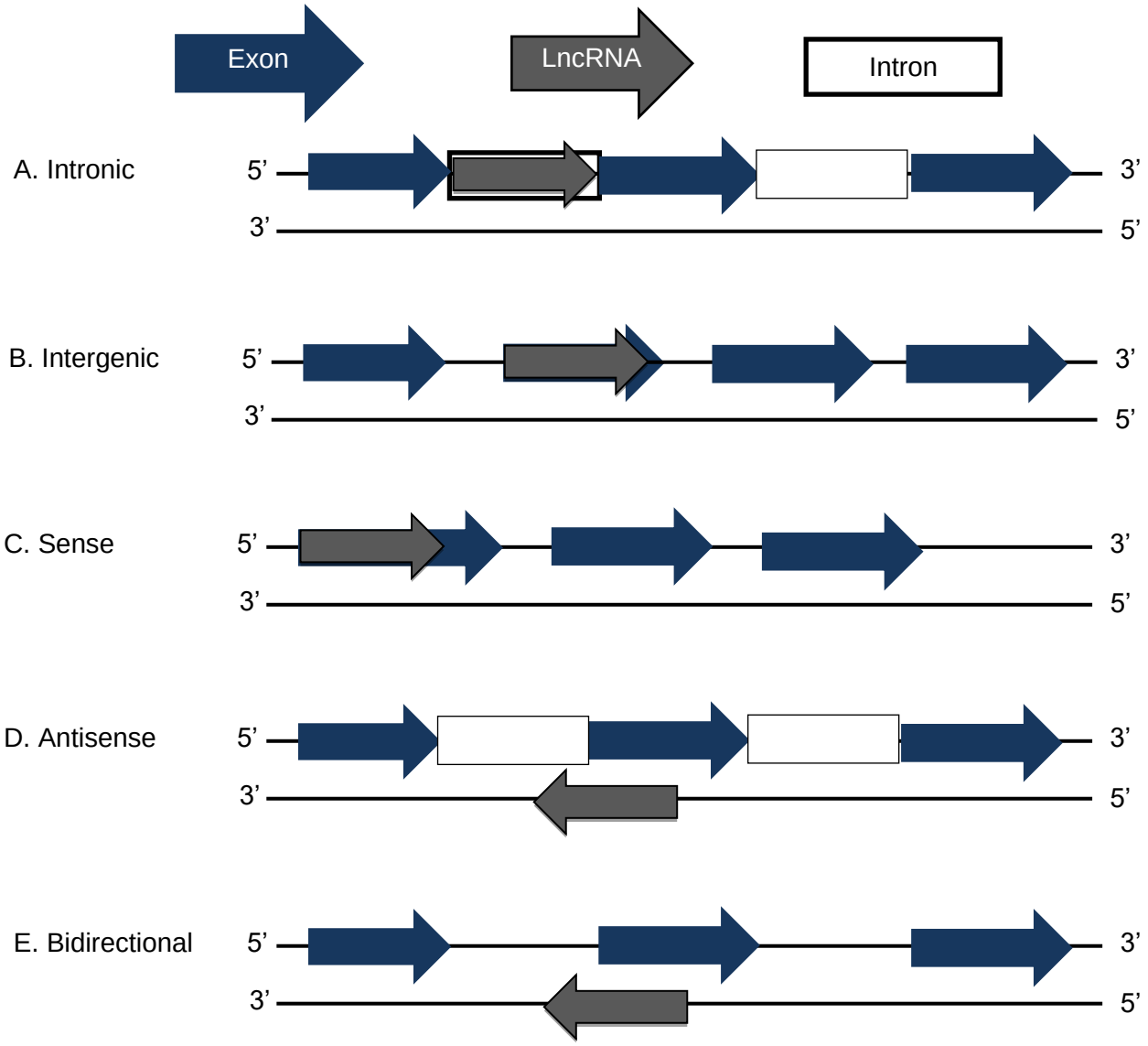


Figure 1.3: The classification of lncRNAs[16] Long non-coding RNAs can overlap protein-coding regions as seen in sense lncRNAs or they can overlap the antisense strand protein-coding regions. (Self-generated figure).

Like mRNA, lncRNAs have RNA polymerase II binding sites while some have 5' caps and poly(A) tails at the 3' end.[7,64] The lncRNAs that do not possess these 5'caps are abundant in the nucleus, however, they have lower levels of expression than mRNA molecules that aren't polyadenylated.[7] Contrary to mRNAs, lncRNAs contain fewer exons, are shorter in length, have a lower sequence conservation and are 10 times less abundant than mRNAs.[60] Unlike genes that encode proteins, lncRNA encoding genes do not follow the same evolutionary

rules, therefore they lack homologous sequences with other species.[7] They do not experience sequence conservation such as those in coding genes. However, they do have the ability to form secondary and tertiary structures to function, implying that the functional regions of these structures may be conserved.[7] It has also been found that lncRNA exons are less conserved than lncRNA promoters, which are similarly conserved as protein-encoding gene promoters.[7] This may mean that at the transcription level, selective constraints are acting on these lncRNAs.[22]

2.7. Biological functions of long non-coding RNAs

It is only recent that these non-coding RNAs have been identified to play significant roles in regulating biological processes.[9,15] These lncRNAs have regulatory functions in transcription, translation, and post-transcriptional modifications.[60,65] These lncRNAs regulate transcription through histone modifications such as HOTAIR, by interacting with transcription factors, and by chromatin remodelling. Some lncRNAs such as lncRNA bladder and prostate suppressor (LBCS) regulate AR translation through the AR signalling pathway to inhibit castration resistance.[66] Moreover, lncRNAs can bind to PCa-related proteins at a post-translational level, thus regulating their stability and function.[67] An example of this is lncRNA metastasis-associated lung adenocarcinoma transcript 1 (MALAT1) which competitively binds depleted in breast cancer 1 (DBC1) to inhibit p53 activity, thus promoting cell proliferation.[68] Evidence of this is that they are found in both the nucleus and the cytoplasm.[7,69] This also suggests that they are necessary for chromatin modification.[7] The regulation of lncRNAs can be controlled by transposable elements as seen in mouse short, interspersed element B2 and the human Alu RNAs, which under environmental stress, are transcribed by RNA polymerase III. They are repressors of transcription by targeting RNA polymerase II.[7] Some lncRNAs also function via interactions with RNA-binding proteins.[9,18,70-72] Their interaction with a variety of other proteins is also important in achieving their biological processes.[60] These interacting proteins include transcription factors and chromatin modifying complexes.[60] An important role of lncRNAs is scaffolding. Certain lncRNAs can be scaffolds for the assembly of certain molecular components.[18] A subtype of these scaffold lncRNAs is architectural RNAs (arcRNAs), which are necessary for the proper assembly of certain nuclear substructures.[73] One of these arcRNAs is nuclear enriched abundant transcript 1 (NEAT1), which is involved in the assembly of more than 60 proteins, some of which are transcription factors while some are RNA-binding proteins. These arcRNAs are rich in repeat sequences, which may be crucial for their functioning as scaffolds.[73]

The transcription of a lncRNA at a specific time and site indicates that these molecules serve as signals under certain stimuli.[18] They can therefore be used as markers for certain

biological events and developmental stages. Examples of these lncRNAs as signals include X chromosome inactivation (Xist and Xist antisense RNA (Tsix)), developmental differentiation (HOTAIR and terminal differentiation-induced non-coding RNA (TINCR)), and gene imprinting (H19 and Air).[7,60] There is evidence that lncRNAs act as decoys in the transcription of promoters and enhancers. Here, they can either positively or negatively regulate transcription.[18] They act as decoys for transcriptional factors,[17,65] chromatin modifiers and other regulatory factors by binding and titrating them.[18] Certain lncRNAs termed competing endogenous RNAs (ceRNAs) are known to interact with miRNAs. The miRNAs are known to be complementary to certain mRNA sequences and bind to them thus regulating gene expression.[18] This interaction results in the destabilization of the target mRNA and inhibits translation. With their ability to bind to miRNAs, ceRNAs prevent miRNAs from disrupting translation. Several miRNA and lncRNA relationships have been investigated recently, however, the extent of their functional relevance remains unclear.[16] The lncRNAs that are ceRNAs compete with miRNAs to regulate them in a process called miRNA sponging.[74] Figure 1.4 illustrates all of the mechanisms of lncRNAs including decoying, miRNA sponging, scaffolding, enhancing and signalling mechanisms.

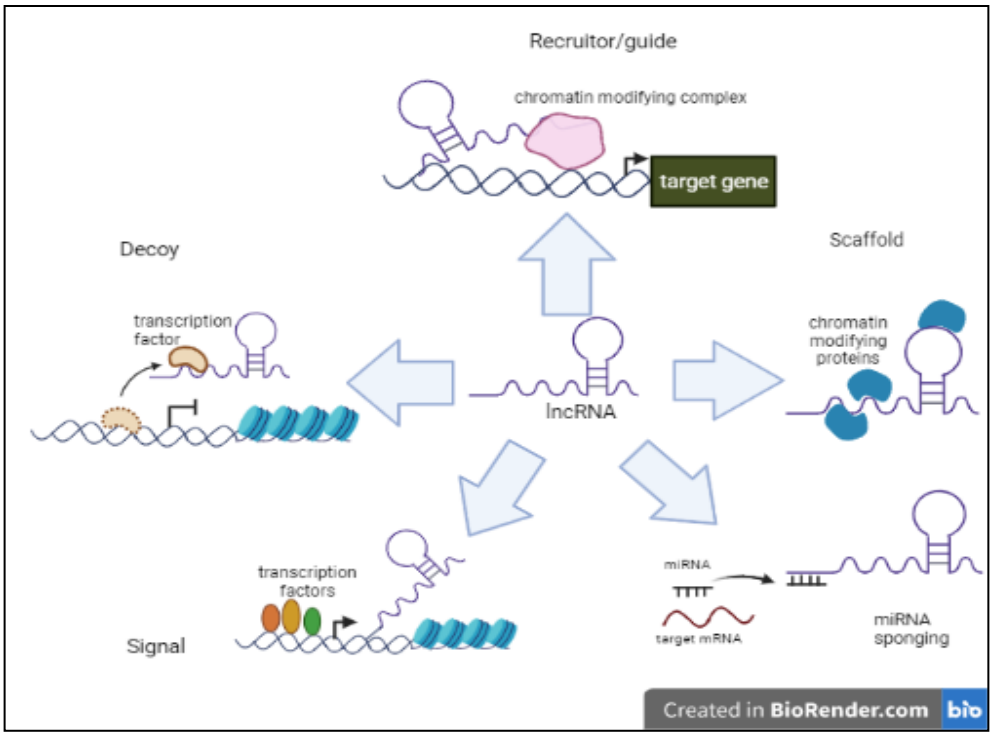


Figure 1.4: Mechanisms of lncRNAs for gene expression regulation. LncRNAs can function as a guide or a recruiter of proteins to specific regions of DNA. They can act as decoys for transcription factors and other proteins leading them away from targeted DNA regions. LncRNAs can be enhancers of gene expression and be a scaffold for the assembly of proteins such as the chromatin modifying complex. They function as miRNA sponges thus

titrating them away from their targeted mRNA and they can serve as signals or markers for certain biological processes. Image self-created with BioRender.com.

2.8 Roles of long non-coding RNAs in cancer

Studies have shown that the deregulation of lncRNAs has been implicated in various cancers as a contributing factor to malignant phenotypic changes.[10,26] Carcinogens, dysregulation of signalling pathways, or viral infections are some of the factors that can induce lncRNA aberrations.[10] The dysregulation of these lncRNAs can be used as a prognostic tool to establish patient treatment outcomes.[75-76] During the formation of a tumour, certain lncRNAs are either up- or down-regulated. These act like tumour suppressors or oncogenes.[16] Some examples of these lncRNAs are maternally expressed gene 3 (MEG3), a type of lncRNA, which behaves like a tumour suppressor in various cancers where its overexpression inhibits the growth of various cancer cell lines. The overexpression of HOTAIR lncRNA, on the other hand, is associated with tumour progression in aggressive endometrial, cervical and breast tumours, among others.[16]

In cancer cells, target genes can be abnormally expressed due to an irregular transcription of associated lncRNA genes. This results in tumour progression and poor prognosis.[17] Due to their tissue-specific expression, they have clinical significance in oncology.[12] Therapeutic techniques that are used against lncRNAs include the use of siRNAs, antisense oligonucleotides, and ribozymes which are destructive to lncRNAs.[77] Various lncRNAs have been reported in PCa, and some are discussed below.

In PCa transcriptomics, several lncRNAs have been identified; one of them is differential display 3 (DD3)/PCA3. This lncRNA is significantly overexpressed and is the most specific PCa molecule.[78-79] It has become a Food and Drug Approved (FDA) PCa biomarker,[78] with satisfactory specificity and sensitivity.[80] Its expression levels are independent of the age and trauma of the patient,[81] unlike PSA levels that can be elevated due to other clinical reasons other than PCa. The PCA3 is known as a regulator of cancer-related genes that have roles in cell adhesion, apoptosis, angiogenesis and signal transduction.[82] Studies on PCA3 knockdown have shown that PCA3 is an inhibitor of AR signalling and cell growth.[82-83] It was also shown that PCA3 knockdown resulted in the upregulation of cofactors such as SRCI and ARA54, which negatively control androgen-responsive genes responding to PCA3 silencing.[84] This evidence supports that PCA3 knockdown has potential as a therapeutic agent that could be used as an inhibitor of PCa cell growth[84] (Table 1.2).

A lncRNA that has been shown to localize genes involved in PCa tumour growth is NEAT1. This lncRNA is a component of the nuclear paraspeckle and is transported to the promoter regions of gene targets[29] thus promoting androgen-dependent PCa growth by increasing

chromatin marks for transcription.[85] It is associated with PCa progression with cells expressing this lncRNA being ADT resistant.[29] Oestrogen receptor alpha (Erα) expressed in PCa interacts with NEAT1 and this is thought to promote the development of castrate resistance. It has high expression in castration-resistant prostate cancer (CRPC) and the levels tend to increase after the use of ADTs in the long term. NEAT1 has the potential to be a target for therapies and a good prognostic marker because increased levels of it are linked to PCa cell metastasis and early biochemical recurrence.[29,85]

LncRNA HOTAIR was found to promote AR transcriptional regulation, thus driving CRPC. It does this by binding AR, preventing it from binding to E3 ubiquitin ligase mouse double minute 2 (MDM2) and undergoing ubiquitination and protein degradation.[29,86] In CRPC, HOTAIR is found to be upregulated after ADTs. This lncRNA promotes the growth and invasion of PCa cells. It has the ability to promote androgen-independent AR activity thus driving CRPC.[29]

The successful treatment of metastatic PCa is ADT, which improves the survival rate in PCa patients.[87] The mechanism of action of ADTs is inhibiting the androgen signalling pathway by the androgen receptor. This pathway promotes growth and its reactivation results in CRPC.[88] LncRNAs that are regulated by ARs are important for the control of gene expression. An example of such an RNA is carboxyl terminal binding protein 1 antisense RNA (CTBP1-AS). This lncRNA activates AR activity and thus promotes the growth of castrate-resistant tumours.[89] The aberrant expression of this lncRNA is therefore directly involved in the transformation of the cell towards CRPC.

In a study by Prensner et al.[27], SChLAP1 was shown to be the most highly expressed lncRNA in PCa patients with metastatic progression. SChLAP1 expression is prostate cancer-specific, and it is expressed minimally in other cancer tissue types. It is associated with a risk of biochemical recurrence, which is the rise of PSA levels in the blood after radical prostatectomy, and its abundance is associated with death from PCa. This lncRNA is used as a non-invasive biomarker[27] to identify patients who have a higher risk of developing CRPC.[90]

Telomerase enzyme functions to replace telomeric repeats on chromosomes that may compromise genetic stability.[91] The Telomerase RNA component (TERC) is an essential lncRNA for this process as it is a subunit for telomerase catalytic activity. It functions as a template for the synthesis and addition of a telomere repeat by telomerase reverse transcriptase (TERT) subunit.[18] Certain structures on the TERC molecule assist TERT in its catalytic and binding activity.[18,92] In 85% of cancer tumours, telomerase activity has been detected.[91,93] When its activity is inhibited, tumour cell apoptosis occurs due to telomere

shortening.[94] When it is upregulated, the telomeres fail to shorten, and the immortalization of cancer cells can occur.[91,95] In gastric cancer, TERC has been shown to promote the invasion and proliferation of tumour cells.[96] In PCa cell lines PC-3 and DU145, TERC has been shown to be overexpressed at all stages. *Myc* is a well-established PCa driver and in clinical samples, it was found to be co-expressed with TERC. When *Myc* expression was reduced, TERC expression was reduced as well. The overexpression of TERC was shown to be driven by *Myc* to aid PCa cell proliferation. This was successfully reduced by the knockdown of TERC .[97] Table 1.2 summarises lncRNAs that have been implicated in the tumorigenesis and prognosis of PCa.

Table 1.2: Summary of lncRNAs involved in PCa tumorigenesis and prognosis.

LncRNA	**Full lncRNA name**	**Expression in PCa**	**Involvement**
CTBP1-AS	Carboxyl terminal binding protein 1 antisense RNA	Upregulated	Promotes castrate-resistant tumour growth and activates AR activity.[89]
DRAIC	Downregulated RNA in cancer	Downregulated	Prognostic marker that is associated with the progression of PCa cells to castrate resistance.[98]
GAS5	Growth arrest-specific transcript 5	Downregulated	Aberrant expression promotes cell proliferation of PCa cells.[99] Associated with the progression of PCa to CRPC.[100]
HOTAIR	HOX antisense intergenic RNA	Upregulated	Binds AR preventing its ubiquitination and degradation. Promotes AR transcriptional regulation and drives CRPC.[86]
MEG3	Maternally expressed gene 3	Downregulated	Promotes proliferation, migration, and invasion. Decreases cell apoptosis rate.[101]
NEAT1	Nuclear enriched abundant transcript 1	Upregulated	Drives androgen-independent PCa cell growth and increases the

			risk of biochemical progression and metastasis of PCa.[85]
PCA-3	Prostate cancer antigen 3	Upregulated	Used for PCa diagnosis and monitoring. Regulates PCa cell survival by modulating AR signalling. Controls the expression of many cancer-related genes.[83]
PCAT1	Prostate cancer-associated transcript 1	Upregulated	Promotes PCa cell migration, invasion, and proliferation through c-Myc protein stabilization while repressing cellular apoptosis.[102]
PCAT5	Prostate cancer-associated transcript 5	Downregulated	Oncogenic and regulated by the ERG transcription factor. Affects cell growth migration and cell apoptosis.[103]
PCAT29	Prostate cancer-associated transcript 29	Downregulated	Tumour suppressor that inhibits cell migration and metastasis. Repressed by AR, associated with the progression of PCa cells.[98]
PCGEM1	Prostate-specific transcript 1	Upregulated	Binds AR and enhances gene activation mediated by AR. Promotes cell proliferation driving aggressive PCa.[104]
PRNCR1	Prostate cancer-associated non-coding RNA 1	Upregulated	Promotes the proliferation of PCa cells driving the progression of aggressive PCa.[104]
SChLAP1	Second chromosome locus	Upregulated	Promotes invasiveness and metastatic progression.

	associated with prostate 1		Involved in biochemical recurrence of PCa cells.[27]
SNHG1	Small nucleolar RNA host gene 1	Upregulated	Promotes cell proliferation of PCa cells.[105]
SOCS2-AS1	SOC2 antisense RNA 1	Upregulated	Promotes PCa cell growth while repressing apoptosis thus promoting CRPC development.[106]
TERC	Telomerase RNA component	Upregulated	Overexpressed in PCa and knockdown results in reduced cell proliferation.[97]

CHAPTER 3

MATERIALS AND METHODS

3. Methods

3.1. Study design

A comparative and analytical study design was used. The expression profiles of lncRNAs in LNCAP and PC-3 cells were compared and analysed to identify differentially expressed lncRNAs and their roles in PCa progression. Due to Bio-Rad's PrimePCR lncRNA PCR Array proprietary control, these arrays have high specificity and sensitivity for lncRNA detection. Bio-Rad has established great quality lncRNA arrays that are cost-effective and have an efficient pipeline for data analysis, compared to alternative lncRNA arrays and RNA-sequencing. The assays were run according to the manufacturer's instructions for the PrimePCR lncRNA H384 for the PC-3 and LNCaP PCa cells. Each of the 86 lncRNA genes was run in duplicate for each of the PC-3 and LNCaP cells, including three reference genes for both PC-3 and LNCaP samples. The genes that were covered in the array are all relevant to PCa. The array was followed by bioinformatics for annotation and enrichment analysis and then by RT-qPCR. Figure 3.1 demonstrates the experimental flow carried out in this study. Using the Bio-Rad's CFX Maestro software to calculate fold changes, up or downregulation of ±2 was considered as differentially expressed.

3.2. Ethical approval

For this project, an ethical clearance was obtained from the Research Ethics Committee, University of Pretoria (Protocol number 561/2021) [Appendix A].

3.3. Research approach outline

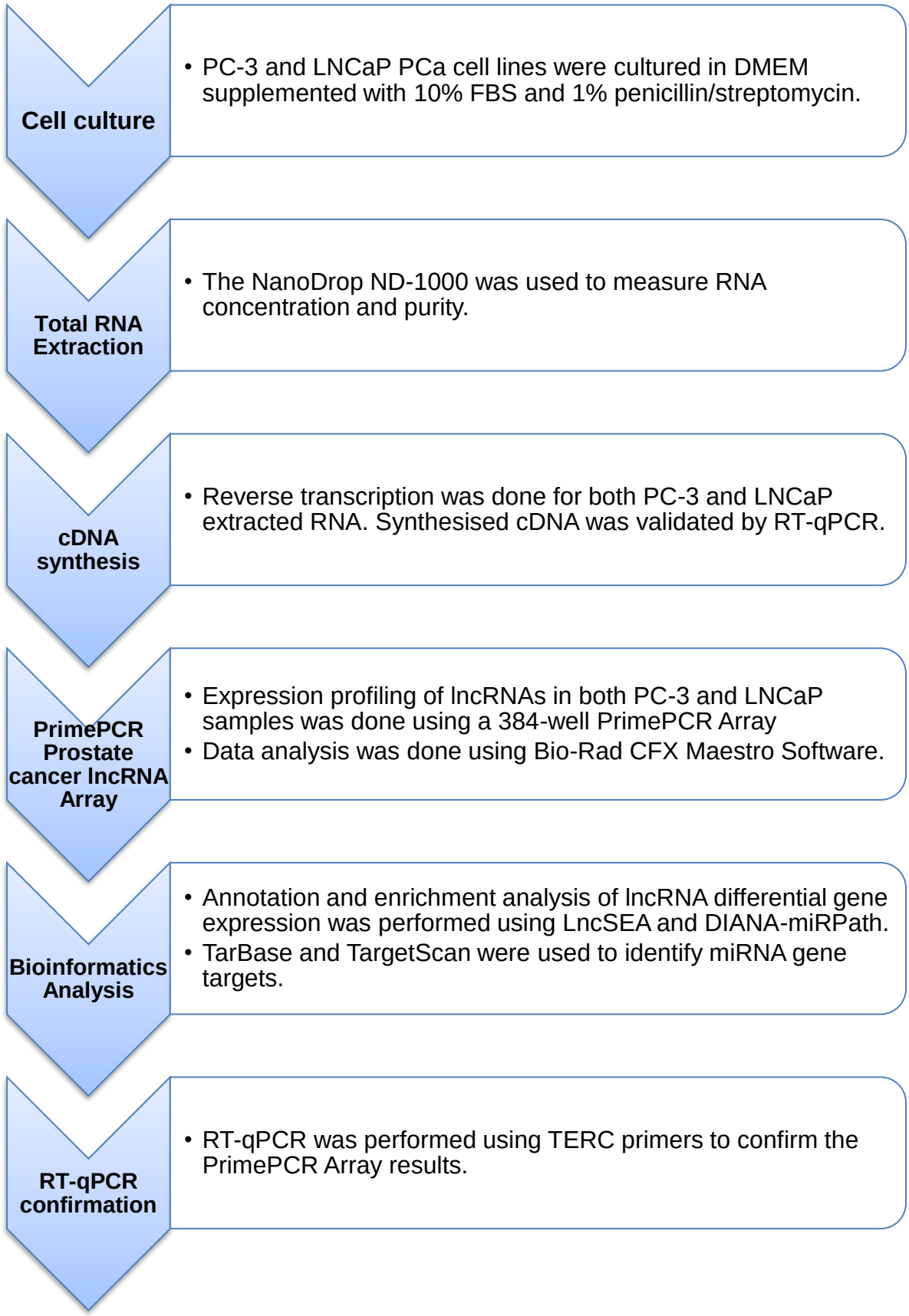


Figure 3.1: Research approach outline

3.4. Cell culture

PC-3 cells (ATTC Catalogue number, CRL-1435) and LNCaP (ATCC Catalogue number, CRL-1740) were purchased from the American Type Culture Collection (ATCC). The cell lines can be adapted to a single growth medium as previously demonstrated to show growth profiles that can be reproduced.[107] The PC-3 and LNCaP cells were grown in Dulbecco's modified Eagle's medium (DMEM):F12, 10% (v/v) fetal bovine serum (FBS), 1% (v/v) penicillin/streptomycin and 5% CO_2 until they reached 70-80% confluency at 37°C for 48 to 72 hours in a humid chamber. The cells were detached from the flask with trypsin, the pellet was centrifuged, and the supernatant was discarded before resuspending in 5 mL of DMEM full culture media. In a ratio of 1:2, the cells were split and sub-cultured. Media containing 10% FBS, and penicillin/streptomycin was replaced every 2-3 days until the desired 70-80% confluency was reached.

3.5. RNA Extraction

RNA was extracted from the cultured PC-3 and LNCaP cells using the RNeasy Mini Kit (Qiagen, Hilden, Germany) following the manufacturer's instructions. Aseptic techniques were used to maintain an RNase-free environment, as RNases degrade RNA. About 1 $x10^6$ cells were harvested from a 25 cm^2 flask. This was done by incubating the cells with 2 mL of trypsin for 3 min for cell detachment and then transferring the detached cells into 15 mL falcon tubes where an equal volume of FBS-containing media was added. Centrifugation was done for 3 min at room temperature at 800 x g and the supernatant was discarded. The cells were then lysed by adding 350 µL of Buffer RLT to the pellet. These tubes were vortexed briefly before an equal volume of 350 µL of 70% ethanol was added to the lysate and it was mixed by pipetting. Up to 700 µL of the lysate was transferred into a RNeasy Mini spin column, which was placed in a 2 mL collection tube. Each of the PC-3 and LNCaP collection tubes with mini spin columns were centrifuged for 15 sec at room temperature at 8000 x g once the lids were closed. The flow-through was then discarded and 350 µL of Buffer RW1 was added to the column and centrifuged at room temperature for 15 sec at 8000 x g. The flow-through was discarded and an on-column DNase digestion was performed. This was done by mixing 10 µL of DNase 1 stock solution with 70 µL of Buffer RDD for each column. This mix was inverted several times and centrifuged briefly. To the column membrane, 80 µL of the DNase solution was added, and this was incubated at room temperature for 15 min. Next, 350 µL of Buffer RW1 was added to the column, which was further centrifuged for 15 sec at 8000 x g. Once the flow through was discarded again, 500 µL of Buffer RPE was then added and the tubes were centrifuged at room temperature for 15 sec at 8000 x g once the lids were closed. The flow-through was discarded and an additional volume of 500 µL of Buffer RPE was added to the column. This was centrifuged at

room temperature for 2 min at 8000 x g. The flow-through was discarded and the RNeasy spin-column was placed on a new 2 mL collection tube and centrifuged at room temperature and at 8000 x g for 1 min to dry the membrane. The RNeasy spin column was inserted into a new 1.5 mL collection tube where 30 µL of RNase-free H_2O was added to the column membrane directly. This was incubated at room temperature for 1 min to allow the H_2O to cover the spin-column membrane completely. Total RNA was eluted by 30µL of RNase-free H_2O, centrifuging at 8000 x g for 1 min at room temperature. RNA purity and quantity were measured using the NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, Delaware, USA) where RNase-free H_2O was used as the blank. Highly pure RNA was extracted, which had an A260/280 absorption ratio of ±2, and an A260/230 absorption ratio of 2.0-2.2.

3.6. cDNA Synthesis

For cDNA synthesis, the ProtoScript® II First Strand cDNA Synthesis Kit was used (New England Biolabs, USA). The reactions were prepared according to the manufacturer's instructions on ice, once the individual tubes were thawed and mixed by inversion, this is shown in Table 3.1. The $d(T)_{23}VN$ primer is an anchored oligo-dT primer which ensures that the primer anneals to the template.

Table 3.1: cDNA synthesis reaction

Reagent	**Volume (µL)**
RNA template	1 µg/ µL *
Primer $d(T)_{23}$ VN (50 µM)	2
ProtoScript II Reaction Mix (2X)	10
ProtoScript II Enzyme Mix (10X)	2
Nuclease-free water to a final volume	x**
Total Volume	**20**

* PC-3 and LNCaP RNA concentrations were brought to 1 µg/µL.
**Dependent on the volume of the template

To each 20 µL reaction, 1 µL of the PrimePCR Reverse Transcription (RT) control template was added. The reactions were incubated in a Bio-Rad MJ Thermal Cycler (Hercules, California USA) under the following conditions: incubation at 42°C for 60 min and enzyme inactivation at 80°C for 5 min, according to the manufacturer's instructions.

To confirm that cDNA synthesis was successful, the synthesized cDNA samples were subjected to qPCR using Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) primers (accession number: NM_002046). The primer sequences are shown in Table 3.2.

Table 3.2: The forward and reverse primer sequences for GAPDH.

GAPDH-Forward	5’ – TGC ACC ACC AAC TGC TTA GC – 3’
GAPDH-Reverse	5’ – GGC ATG GAC TGT GGT CAT GAG – 3’

The running conditions shown in Table 3.3 were used to run RT-qPCR using the PikoReal 96 Real-Time PCR System (ThermoFisher Scientific, Inc., Waltham, USA).

Table 3.3: Real-time qPCR running conditions.

Cycle	**Stage**	**Temperature (°C)**	**Duration**
Hold	Initial denaturation	95	10 min
PCR 40 cycles	Denaturation	95	10 sec
	Annealing	60	30 sec
	Extension	60	30 sec
Melting Curve	Dissociation	60-95 (0.2°C increments)	1 sec hold time

3.7. PrimePCR lncRNA Array Principle and Procedure

For researchers, the detection of novel transcripts has always proven difficult, especially for those transcripts that have no protein-coding potential.[17] Long non-coding RNAs are highly tissue-specific, and they are expressed in low levels, which has made their discovery difficult.[7,13,20,108] Real-time qPCR (RT-qPCR) is used for gene expression analysis in multiple applications since it is highly specific and reliable. It can be used to validate microarray and RNA-sequencing results. Nonspecific amplification is one of the challenges presented during gene expression analyses and RT-qPCR overcomes this by using primers designed specifically for the amplification of the targeted genes.

Bio-Rad has developed a high quality, high-specificity PrimePCR lncRNA PCR Assay for lncRNA gene expression analyses. Due to guaranteed Bio-Rad’s Trademark quality control, the array was run once, each gene in duplicate for both PC-3 and LNCaP cells. The assays have been particularly designed for SYBR® Green, real-time PCR lncRNA detection. A PCa-specific 384-well plate lncRNA PCR Array was used (Bio-Rad, Catalogue number: 12004237). The Bio-Rad’s PCR Array is compatible with systems such as BioRad CFX and Thermo Fisher QuantStudio™ 5.

For the array, the cDNA samples and the PrimePCR component were thawed and mixed thoroughly by brief vortexing. The tubes were then briefly centrifuged so that the solutions could collect at the bottom of the tubes and then they were placed on ice. The PrimePCR Plate was brought to room temperature before the seal was removed. To the cDNA samples, 1 µL of RT (reverse transcription) control template was added to each 20 µL cDNA reaction. This RT control template had been resuspended in 200 µL nuclease-free water and kept on ice. The POWERUP SYBR Master Mix (Thermo Fisher, Catalogue number: A25741) was used and the reaction mixes were prepared as per the following Table 3.4 for the 384-well PCR array plate:

Table 3.4: PrimePCR Array reaction mixture

Component	**Volume per Reaction (µL)**	**Volume X420 (µL)**
20x PrimePCR Assay	Lyophilized in the well	Lyophilized in the well
2x POWERUP SYBR Master Mix	5	2100
cDNA sample	0.8	Loaded into the wells individually
Nuclease-Free Water	4.2	1764
Total Volume	**10**	**3884**

The reaction components excluding the cDNA samples were loaded into the wells (9.2 µL in total) and 0.8 µL of cDNA was loaded into the wells. This made a total of 10 µL of the PCR reaction mix into each well. In columns 1 to 12 from rows A to P, the wells were loaded with LNCaP cDNA, which served as the control sample. From columns 13 to 24 and rows A to P, the wells were loaded with PC-3 cDNA, which was the experimental sample. Next, 0.5 µL of the PrimePCR PCR Control Assay was added to the designated wells for PCR control (Table 3.6). The plate was sealed with an optical seal, then was briefly and gently vortexed and centrifuged before loading into the ABI QuantStudio™ 5 RT-qPCR system (Waltham, Massachusetts, USA). The software measures the amplification of the targeted genes and housekeeping genes. This highly sensitive system allows for accurate relative gene expression analysis. Table 3.5 shows the thermal cycling conditions for the array, while Table 3.6 shows the Prime PCR array layout.

Table 3.5: PrimePCR cycling conditions using the ABI QuantStudio™ 5 RT-qPCR system.

Stage	**Step**	**Temperature (°C)**	**Time**
Hold	Enzyme activation	50	2 min
		95	2 min

PCR	Denaturation	95	15 sec
	Annealing/Extension	95	1 min
Melt curve	Denaturation	95	15 sec
	Annealing	60	1 min
	Dissociation	95	15

3.8. Data Analysis using CFX Maestro

Bio-Rad's CFX Maestro analysis tool was used for the processing and analysis of the PrimePCR Array data of all the samples. The QuantStudio5 run file was exported to the CFX Maestro application.

A quantification cycle (Cq) of 35 was the gene expression cut-off value and genes with Cq values above 35 cycles were considered to not be detectable. A total of three housekeeping genes were included on the plate, Beta-2-microglobulin (B2M), Hydroxymethylbilane synthase (HMBS) and TATA box binding protein (TBP). These genes were used for the calculation of the change in Cq (ΔCq) value for each of the genes of interest, where ΔCq was calculated as: the Cq value of the gene of interest less than the Cq value of the reference gene. The difference between the ΔCq of the advanced metastasis group (PC-3) and that of the low metastasis group (LNCaP) is represented by ΔΔCq value, where the fold-change can be calculated by $2^{(-\Delta\Delta Cq)}$.[109] This value is a representation of the level of expression of the lncRNA in the advanced metastasis sample *versus* that in the low metastasis sample. The ±2 upregulation or downregulation fold change was used as a basis of target selection. The formulae are shown below.

ΔCq = Cq (Test targeted gene) – Cq (Test GAPDH Housekeeping gene)

ΔCq = Cq (Control target gene) – Cq (Control GAPDH housekeeping gene)

ΔΔCq = ΔCq (Test) - ΔCq control

Fold change = $2^{-[\Delta Cq\ (test)\ -\Delta Cq\ (control)]}$

Table 3.6: Representation of the Bio-Rad PrimePCR PCa lncRNA H384 Array. Each of the 86 lncRNAs is represented in four replicates. The 384 lncRNA panel ensures quality checks by the incorporated proprietary controls. Wells 1-12 in rows A-P were loaded with LNCaP cDNA (control sample) and wells 13-24 in rows A-P were loaded with PC-3 cDNA (experimental sample). The PCR control was loaded into the designated wells. Quality controls for gDNA contamination, PCR, RNA integrity and RT efficiency were included in the assay.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
A	PCGEM1	SOCS2-AS1	PCA3	PRNCR1	CBR3-AS1	PCAT29	PCGEM1	SOCS2-AS1	PCA3	PRNCR1	CBR3-AS1	PCAT29	PCGEM1	SOCS2-AS1	PCA3	PRNCR1	CBR3-AS1	PCAT29	PCGEM1	SOCS2-AS1	PCA3	PRNCR1	CBR3-AS1	PCAT29
B	RP1-30M3.6	GNG12-AS1	LINC00493	RP11-619J20.1	RP4-635E18.8	HMGN3-AS1	RP1-30M3.6	GNG12-AS1	LINC00493	RP11-619J20.1	RP4-635E18.8	HMGN3-AS1	RP1-30M3.6	GNG12-AS1	LINC00493	RP11-619J20.1	RP4-635E18.8	HMGN3-AS1	RP1-30M3.6	GNG12-AS1	LINC00493	RP11-619J20.1	RP4-635E18.8	HMGN3-AS1
C	RN7SL15P	AC003075.4	RP11-884K10.6	RP11-390P2.4	TP53TG1	RP11-382A20.4	RN7SL15P	AC003075.4	RP11-884K10.6	RP11-390P2.4	TP53TG1	RP11-382A20.4	RN7SL15P	AC003075.4	RP11-884K10.6	RP11-390P2.4	TP53TG1	RP11-382A20.4	RN7SL15P	AC003075.4	RP11-884K10.6	RP11-390P2.4	TP53TG1	RP11-382A20.4
D	KRT81	RP4-724E16.2	KB-1562D12.2	TERC	RP11-545E17.3	CFLAR-AS1	KRT81	RP4-724E16.2	KB-1562D12.2	TERC	RP11-545E17.3	CFLAR-AS1	KRT81	RP4-724E16.2	KB-1562D12.2	TERC	RP11-545E17.3	CFLAR-AS1	KRT81	RP4-724E16.2	KB-1562D12.2	TERC	RP11-545E17.3	CFLAR-AS1
E	RP3-443C4.2	FGD5-AS1	SNHG8	RP11-556O9.2	KCNQ1OT1	TINCR	RP3-443C4.2	FGD5-AS1	SNHG8	RP11-556O9.2	KCNQ1OT1	TINCR	RP3-443C4.2	FGD5-AS1	SNHG8	RP11-556O9.2	KCNQ1OT1	TINCR	RP3-443C4.2	FGD5-AS1	SNHG8	RP11-556O9.2	KCNQ1OT1	TINCR
F	SNHG19	RP11-379H18.1	RP11-759F5.1	lnc-FAM83G-3	XIST	RNF139-AS1	SNHG19	RP11-379H18.1	RP11-759F5.1	lnc-FAM83G-3	XIST	RNF139-AS1	SNHG19	RP11-379H18.1	RP11-759F5.1	lnc-FAM83G-3	XIST	RNF139-AS1	SNHG19	RP11-379H18.1	RP11-759F5.1	lnc-FAM83G-3	XIST	RNF139-AS1
G	AFG3L1P	LINC01024	RP11-46C24.7	LINC00339	RP11-48B3.3	B2M	AFG3L1P	LINC01024	RP11-46C24.7	LINC00339	RP11-48B3.3	B2M	AFG3L1P	LINC01024	RP11-46C24.7	LINC00339	RP11-48B3.3	B2M	AFG3L1P	LINC01024	RP11-46C24.7	LINC00339	RP11-48B3.3	B2M
H	AC006994.1	FOXP4-AS1	RP11-146E13.4	MAGI1-IT1	RP4-591B8.2	HMBS	AC006994.1	FOXP4-AS1	RP11-146E13.4	MAGI1-IT1	RP4-591B8.2	HMBS	AC006994.1	FOXP4-AS1	RP11-146E13.4	MAGI1-IT1	RP4-591B8.2	HMBS	AC006994.1	FOXP4-AS1	RP11-146E13.4	MAGI1-IT1	RP4-591B8.2	HMBS
I	PCAT1	HOTAIR	LINC00657	PTPRJ-AS1	SCHLAP1	TBP	PCAT1	HOTAIR	LINC00657	PTPRJ-AS1	SCHLAP1	TBP	PCAT1	HOTAIR	LINC00657	PTPRJ-AS1	SCHLAP1	TBP	PCAT1	HOTAIR	LINC00657	PTPRJ-AS1	SCHLAP1	TBP
J	RP3-523E19.2	COX6CP10	CTD-2562J17.7	CTA-363E19.2	WAC-AS1	PAQ1	RP3-523E19.2	COX6CP10	CTD-2562J17.7	CTA-363E19.2	WAC-AS1	PAQ1	RP3-523E19.2	COX6CP10	CTD-2562J17.7	CTA-363E19.2	WAC-AS1	PAQ1	RP3-523E19.2	COX6CP10	CTD-2562J17.7	CTA-363E19.2	WAC-AS1	PAQ1
K	RP11-121A14.2	RP11-226L15.5	RP11-174G6.5	RP11-473I1.5	RP11-156E6.1	PAQ2	RP11-121A14.2	RP11-226L15.5	RP11-174G6.5	RP11-473I1.5	RP11-156E6.1	PAQ2	RP11-121A14.2	RP11-226L15.5	RP11-174G6.5	RP11-473I1.5	RP11-156E6.1	PAQ2	RP11-121A14.2	RP11-226L15.5	RP11-174G6.5	RP11-473I1.5	RP11-156E6.1	PAQ2
L	RP11-196G18.22	BCDIN3D-AS1	GTF3AP2	RN7SL277P	RP11-338I21.1	gDNA	RP11-196G18.22	BCDIN3D-AS1	GTF3AP2	RN7SL277P	RP11-338I21.1	gDNA	RP11-196G18.22	BCDIN3D-AS1	GTF3AP2	RN7SL277P	RP11-338I21.1	gDNA	RP11-196G18.22	BCDIN3D-AS1	GTF3AP2	RN7SL277P	RP11-338I21.1	gDNA
M	RP11-119B16.2	LURAP1L-AS1	DUTP1	UBA6-AS1	CTD-2339M3.1	PCR	RP11-119B16.2	LURAP1L-AS1	DUTP1	UBA6-AS1	CTD-2339M3.1	PCR	RP11-119B16.2	LURAP1L-AS1	DUTP1	UBA6-AS1	CTD-2339M3.1	PCR	RP11-119B16.2	LURAP1L-AS1	DUTP1	UBA6-AS1	CTD-2339M3.1	PCR
N	RP11-67L2.2	RP11-363E7.4	FAM83H-AS1	LINC01137	CTD-3157E16.1	RQ1	RP11-67L2.2	RP11-363E7.4	FAM83H-AS1	LINC01137	CTD-3157E16.1	RQ1	RP11-67L2.2	RP11-363E7.4	FAM83H-AS1	LINC01137	CTD-3157E16.1	RQ1	RP11-67L2.2	RP11-363E7.4	FAM83H-AS1	LINC01137	CTD-3157E16.1	RQ1
O	NHS-AS1	AC009404.2	TUG1	AC005618.6	PSMD6-AS2	RQ2	NHS-AS1	AC009404.2	TUG1	AC005618.6	PSMD6-AS2	RQ2	NHS-AS1	AC009404.2	TUG1	AC005618.6	PSMD6-AS2	RQ2	NHS-AS1	AC009404.2	TUG1	AC005618.6	PSMD6-AS2	RQ2
P	CTD-2010I16.1	ZNF767P	RP11-98I9.4	RP11-244F12.3	NRSN2-AS1	RT	CTD-2010I16.1	ZNF767P	RP11-98I9.4	RP11-244F12.3	NRSN2-AS1	RT	CTD-2010I16.1	ZNF767P	RP11-98I9.4	RP11-244F12.3	NRSN2-AS1	RT	CTD-2010I16.1	ZNF767P	RP11-98I9.4	RP11-244F12.3	NRSN2-AS1	RT

3.9. Bioinformatics pathway analysis

Using multiple bioinformatics tools and computational methods is an essential *in silico* technique to analyze genes, their biological products and their associated disease pathways.[110] Several valuable lncRNA tools and databases have been constructed for the analysis of lncRNAs.[111] In this study, various bioinformatics tools were utilized for annotation and enrichment analysis of the differentially expressed lncRNAs. These tools are described briefly below.

Annotation and enrichment analysis were performed using Long non-coding RNA related Sets and Enrichment Analysis (LncSEA).[111] LncSEA database (http://bio.liclab.net/LncSEA/Analysis.php) is a functional analysis tool that supports over 40 000 lncRNA reference sets and covers more than 50 000 lncRNAs.[111] LncSEA can accommodate multiple human lncRNA resources such as StarBase,[112] LncRNADisease[113] and LncBase.[114] Due to its inclusion of various bioinformatics tools, one does not need to use each tool individually. LncSEA performs enrichment and annotation analyses of lncRNAs and outputs information such as the disease associated with the lncRNAs are linked to and their interacting miRNAs. To run a LncSEA analysis, upregulated and downregulated lncRNAs, according to the fold change (Tables 4.4. and 4.5, Chapter 4), were obtained from the CFX Maestro analysis and used as input data in LncSEA. Parameters were set to include the hypergeometric test p-value of 0.01 and adjusted the p-value of 0.05.

The interaction of lncRNAs and miRNAs Is an important aspect of gene regulation because lncRNAs reduce the regulatory effects of miRNAs on their mRNA targets.[115] It is for this reason that the miRNA pathways were explored further with DIANA-miRPath v3.0. (http://www.microrna.gr/miRPathv3)[116], which is a web server that analyses the miRNA pathways[116] using miRNA targets entered by the user. This tool incorporates more than 600 000 experimentally supported miRNAs from DIANA-TarBase v7.0[117] and it allows for the visualization of the Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways of the miRNA targets.[116] GO is used to perform enrichment analysis on the input gene sets with the aim of identifying biological processes, molecular functions, and cellular locations that are affected by the studied genes. KEGG pathway enrichment analysis is used to map the pathways and molecular interactions associated with the input genes. For this analysis, the miRNAs that were shown to interact with the differentially expressed lncRNA genes using LncSEA (Supplementary Tables C3.1 and C3.2), were arranged in order of p-value, from lowest to highest. The top 30 miRNAs with

the lowest p-values were selected (≤1.05E-06). The default parameters, a microT threshold of 0.8 and a p-value threshold of 0,05 were applied.

For the GO analysis, three heatmaps were generated for each GO subcategory, namely, molecular function, biological process, and cellular component. The subcategory 'biological process' is presented in the results section below, while the two other subcategory heatmaps are shown in Appendix C4: Supplementary Figures C4.1 and C4.2.

Telomerase RNA component (TERC) lncRNA was selected for ceRNA network analysis as it interacted with the hsa-miR-320 family of miRNAs that are tumour suppressors and have been reported as potential PCa diagnostic biomarkers.[118] In LncSEA, TERC was shown to interact with four members of the miR-320 family namely hsa-miR-320a, hsa-miR-320b, hsa-miR-320c and hsa-miR-320d. It was also shown to interact with two other miRNAs, hsa-miR-4429 and hsa-miR-338-3p, which also have tumour suppressive qualities. To predict the gene targets for these miRNAs, TargetScan v8.0 (http://www.targetscan.org/)[119] and TarBase v8.0 (https://dianalab.e-ce.uth.gr/html/diana/web/index.php?r=tarbasev8)[117] were used. The database returned several hundreds of predicted gene targets. To narrow this down, only the gene targets with a stringent TarBase prediction score of > 0.9 and those that were common among these databases were selected for the ceRNA network. Network construction and visualization was done using Cytoscape v3.9.1.[120]

3.10. Validation of PrimePCR Array Data by RT-qPCR

RT-qPCR is a well-established and sensitive technique used for gene expression quantification.[121] Reference (housekeeping) genes are used as endogenous control because their expression should be constant.[122] To validate the PrimePCR Array data, RT-qPCR was performed. Each of the 3 biological replicates were run three independent times in three technical replicates. The representative gene from the PrimePCR Array that was selected was TERC, which was found to be upregulated (Table 4.4).

3.10.1 Primer Design

The sequence used to design the TERC primers was based on the NCBI accession numbers provided by Bio-Rad, following the PrimePCR Array analysis. The gene target was TERC (accession number NR_001566.1). Primer3 software (https://primer3.ut.ee/)[123] was used to design the forward and reverse primers for TERC. Primer specificity was ensured using the Basic Local Alignment Search Tool (BLAST).[124] The amplicon size was 117bp for TERC and 87bp for GAPDH, which was used as the reference gene (the primer sequences are shown in Table 3.2). The primer sequences for the TERC gene target are shown in Table 3.7. GADPH use as a reference gene in TERC lncRNA gene expression studies has been previously demonstrated.[125-126]

Table 3.7: The forward and reverse primer sequences of the TERC gene target used to perform RT-qPCR for the validation of the PrimePCR Array data.

TERC-Forward	5' – TGC TCT AGA ATG AAC GGT GGA – 3'
TERC-Reverse	5' – CCT AAC TGA GAA GGG CGT AGG – 3'

3.10.2. RT-qPCR

Total RNA from the LNCaP and PC-3 RNA samples was reverse transcribed to produce cDNA using the ProtoScript® II First Strand cDNA Synthesis Kit, as described in section 3.6. The cDNA generated was used immediately in RT-qPCR, as shown in Table 3.8 or stored at -20°C for future use.

Table 3.8: RT-qPCR reaction components:

Component	**Volume per Reaction (µL)**
2x SYBR Green Master Mix	5
TERC Forward primer*	0.4
TERC Reverse primer*	0.4
cDNA	0.8**
Nuclease-free water	3.4

* A final concentration of 4µM of each primer was used.

** No template reactions were included for the negative controls.

The reactions were prepared in strip PCR tubes with negative controls. The 10 µL reactions were loaded into the ABI 7500 Real-Time qPCR Instrument (ABI, Waltham, USA). The following thermal cycling conditions were used: 10 min at 95°C for denaturation, 40 cycles at 95°C for 15 sec, 1 min at 60°C for data collection, 15 sec at 95°C and 1 min at 60°C for the dissociation stage.

3.10.3 RT-qPCR data analysis

Once the runs were complete on the ABI 7500, the data was exported into an Excel file and fold change was calculated. The housekeeping gene GAPDH was used to calculate the change in the Cq (ΔCq) value for TERC where ΔCq was calculated as: the Cq value of the gene of interest minus the Cq value of the reference gene. The difference between the ΔCq of the advanced metastasis group (PC-3) and that of the low metastasis group (LNCaP) is represented by ΔΔCq value, where the fold-change was calculated by $2^{(-\Delta\Delta Cq)}$.[109]

3.11. Statistical analysis

The RT-qPCR data was collected for at least three independent experiments, and it was analysed using GraphPad Prism v9.5.0 and expressed as means ± standard error of mean (SEM). A one-sample t-test was performed to determine statistical significance, and a p-value ≤0.05 was considered statistically significant.

CHAPTER 4

RESULTS

4. Results

This chapter addresses the aim and objectives outlined in Chapter 1. It is divided into the PCR array analysis section, the *in silico* bioinformatics analysis section using various bioinformatics tools including LncSEA, DIANA-miRPath, TargetScan, and TarBase and lncRNA gene expression verification by RT-qPCR.

4.1. cDNA Verification

To verify the success of reverse transcription, the cDNA samples were subjected to real-time qPCR with GAPDH primers. The analysis is shown in Figure 4.1.

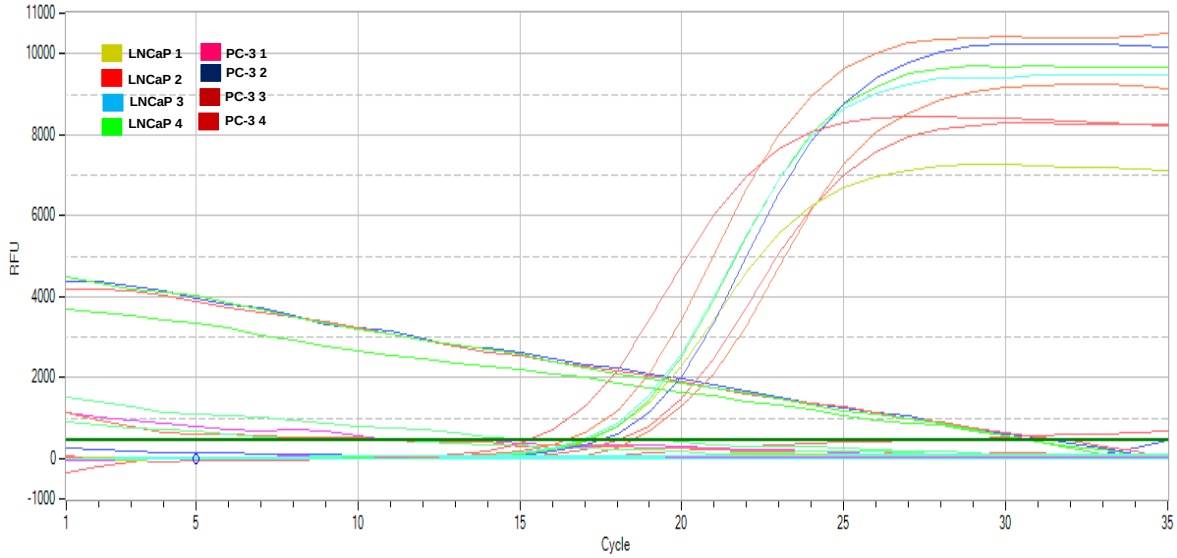


Figure 4.1: Real-time qPCR amplification curve. cDNA from LNCaP and PC-3 cell lines was subjected to qPCR using GAPDH primers. The cycle threshold was set to 35. All the samples were confirmed to have cDNA. The no template control showed no amplification.

The results shown in Figure 4.1 validate that GAPDH primers successfully annealed to the generated cDNAs, that were used for the downstream PCR array.

4.2. PrimePCR Array analysis using CFX Maestro

The Prostate Cancer lncRNA PrimePCR Array was run to obtain expression profiles of the various lncRNAs of interest. The aim was to identify the differentially expressed lncRNA genes (DEGs) between the control LNCaP sample and the experimental PC-3 sample. Table 4.1 refers to the quality control checks of the PCR array. These Bio-Rad-validated control assays are used to assess the quality of the experimental samples and how this may

affect qPCR performance. The positive PCR control targeted a synthetic DNA template which determines if samples have factors such as inhibitors that may affect the gene expression results. The RT control evaluated the RT performance by introducing a synthetic RNA template into the cDNA synthesis reaction. The genomic DNA (gDNA) control assay targeted a not-transcribed sequence of the genome. In the case of genomic DNA contamination, this control would have failed the quality check. The RNA quality control assessed whether the RNA quality could have affected the PCR results of the array. This array passed all four quality checks of the assay. This means that there were no factors that affected PCR negatively, RT performance was optimal, there was no gDNA contamination, and the RNA quality was optimal for PrimePCR.

Table 4.1: PrimePCR control analysis

Control	**Threshold Cq**	**LNCaP**	**PC-3**
Positive PCR	Cq ≤ 30	Passed	Passed
Reverse Transcription	Cq ≤ 30	Passed	Passed
DNA Contamination	Cq ≥ 35	Passed	Passed
RNA Quality	ΔCq ≤ 2	Passed	Passed

The PCR arrays measured the expression of 86 lncRNAs across two samples, the LNCaP and the PC-3, respectively. For this study, the LNCaP expression was set as the control for comparison. Using the CFX Maestro PrimePCR Analysis software, the array results were analysed. The gene expression changes are represented in various tables, graphs, scatter plots and a heatmap. The upregulated genes (Table 4.2), and the downregulated genes (Table 4.3) were further represented in a scatter plot (Figure 4.2), a volcano plot (Figure 4.3) and a heat map (Figure 4.4). The red dots/colours in these plots denote upregulation while green colours demonstrate downregulated genes. Thirty-six of 86 lncRNAs were shown to be upregulated, while 14% of the lncRNA panel was downregulated.

Table 4.2: Upregulated lncRNAs in PC-3 cells

LncRNA	**GenBank ID**	**Ensembl ID**	**Regulation/ Fold change**
KRT81	NC_000012.11, NG_008184.1, NT_029419.12	ENSG00000205426	1109964.44
AC003075.4	N/A	ENSG00000237773	165.03
CFLAR-AS1	NR_040030.1	ENSG00000226312	31.00
RP11-174G6.5	N/A	ENSG00000261324	19.77

LURAP1L-AS1	NR_125775	ENSG00000235448	17.21
CTD-2339M3.1	N/A	ENSG00000242602	14.68
RP11-244F12.3	N/A	ENSG00000259498	13.57
LINC00339	NR_110692.1 NR_023918.2 NR_109762.1 NR_023919.2 NR_109760.1 NR_109761.1 NR_109759.1	ENSG00000218510	8.77
CTD-3157E16.1	N/A	ENSG00000265519	8.73
RP3-443C4.2	N/A	ENSG00000233823	8.46
SOCS2-AS1	NR_038263.1	ENSG00000246985	7.82
TINCR	NM_001396408.1	ENSG00000223573	7.72
RNF139-AS1	NR_108047, NR_108048	ENSG00000245149	6.42
RN7SL15P	N/A	ENSG00000264573	5.53
HOTAIR	NR_047517, NR_047518, NR_003716	ENSG00000228630	5.19
COX6CP10	N/A	ENSG00000238008	4.63
AC009404.2	N/A	ENSG00000236255	4.61
RP1-30M3.6	N/A	ENSG00000272402	4.56
RP11-156E6.1	N/A	ENSG00000259623	4.28
RP4-635E18.8	N/A	ENSG00000271895	4.24
LINC00493	NM_001348957.2 NM_001348958.2 (transcript var. 1 and 2	ENSG00000232388	4.15
AFG3L1P	NR_003228, NR_003226, NR_003227	ENSG00000223959	3.77
MAGI1-IT1	NR_145422.1	ENSG00000272610	3.29
TUG1	NR_110492, NR_002323, NR_110493	ENSG00000253352	3.24
TERC	NR_001566.1	ENSG00000270141	3.16
CTA-363E19.2	N/A	ENSG00000272072	3.14
RP11-382A20.4	N/A	ENSG00000259986	3.01
RP11-226L15.5	N/A	ENSG00000260766	2.95
DUTP1	N/A	ENSG00000229048	2.90
GTF3AP2	N/A	ENSG00000258966	2.89
LINC01024	NR_102741.1 NR_102740.1 NR_102739.1	ENSG00000245146	2.78
LINC00657	NR_027451	ENSG00000260032	2.74
AC005618.6	N/A	ENSG00000272070	2.72
PTPRJ-AS1	N/A	ENSG00000254879	2.33

FOXP4-AS1	NR_126415, NR_126416, NR_126417	ENSG00000234753	2.31
RP11-338I21.1	N/A	ENSG00000271933	2.05

The fold change shown represents the mean of the replicates. The most upregulated lncRNA in this study was KRT81 with a fold change of 1109964.44. The least upregulated lncRNA was RP11-338I21.1 with a regulation reading of 2.05 as the cut-off value was set to 2.00. KRT81 is associated with metastatic lymph nodes of breast cancer. Amongst these upregulated lncRNAs are HOTAIR and TERC, which have been previously described to be upregulated in various cancers, including PCa.[91,127] In this study, the upregulation of these two lncRNAs was shown to be epigenetically regulated, as described in the next bioinformatics section.

Table 4.3: Downregulated lncRNAs in PC-3 cells

LncRNA	**GenBank ID**	**Ensembl ID**	**Regulation/Fold change**
PCAT29	NR_126437.1 NR_126438.1	ENSG00000259641	-349.40
PCAT1	NR_045262	ENSG00000253438	-110.80
RP11-884K10.6	N/A	ENSG00000271916	-71.40
KB-1562D12.2	N/A	ENSG00000253991	-21.46
LINC01137	NR_038842	ENSG00000233621	-7.37
TP53TG1	NR_015381	ENSG00000182165	-5.03
PRNCR1	NR_109833.1	ENSG00000282961	-4.85
FAM83H-AS1	NR_033849	ENSG00000203499	-2.66
PCGEM1	NR_002769.1	ENSG00000227418	-2.32
SNHG8	NR_003584.3, NR_034010.1, NR_034011.1	ENSG00000269893	-2.29
CTD-2562J17.7	N/A	ENSG00000254429	-2.11
RP11-48B3.3	N/A	ENSG00000254162	-2.10

PCAT29 was the most downregulated lncRNA of the 86 lncRNAs that were included in this study. This lncRNA had a fold change of -349.40 and it is a tumour suppressor gene that reduces tumour migration. The least downregulated lncRNA, RP11-48B3.3, had a regulation of -2.10.

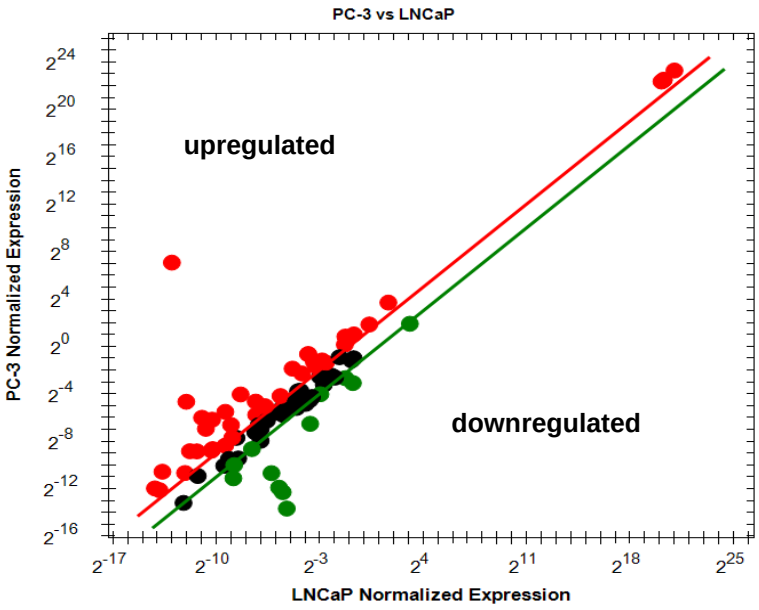


Figure 4.2: Scatter Plot showing the normalized expression between PC-3 and LNCaP lncRNAs. The regulation threshold was set to ±2. Expression of upregulated genes is shown in red; expression of downregulated genes is shown in green. The lncRNAs without significant gene expression changes are shown in black.

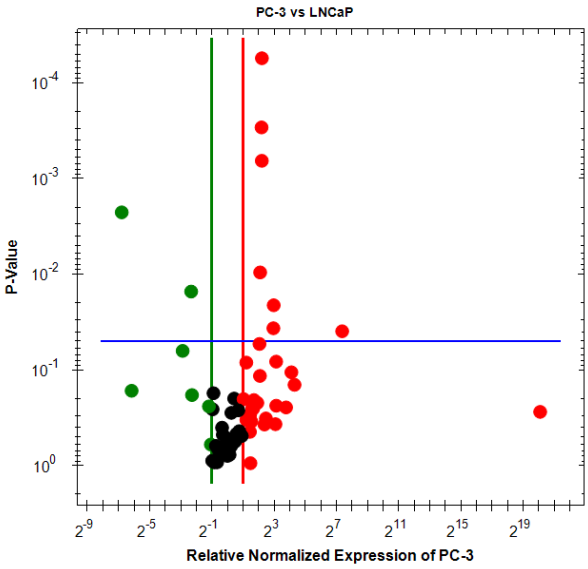


Figure 4.3: Volcano Plot showing the relative normalized expression between PC-3 and LNCaP lncRNAs. The regulation threshold was set to 2.00 and the p-value threshold was set to 0.05. The upregulated genes are represented in red, while green dots show downregulated lncRNA genes. The black dots represent genes that were not differentially expressed.

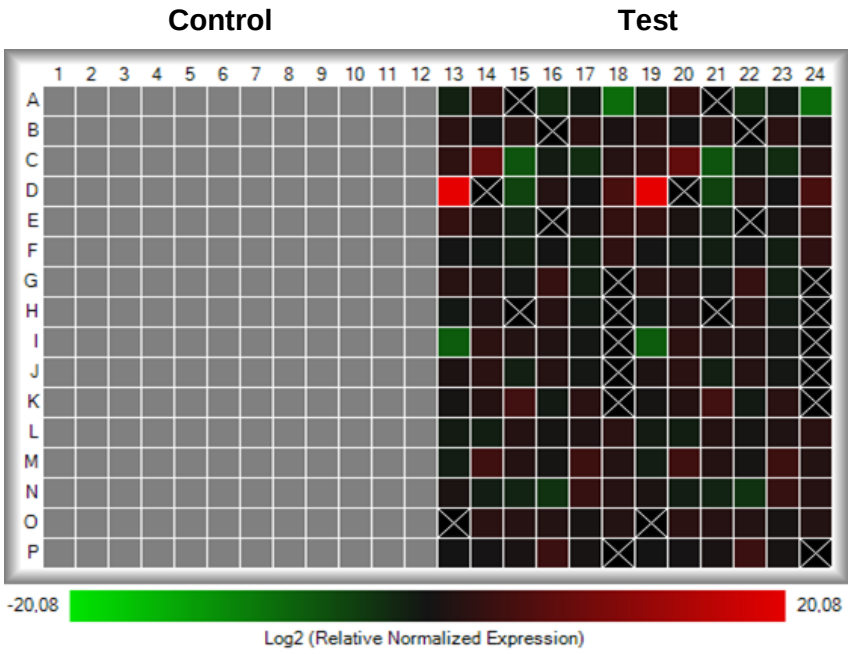


Figure 4.4: Heatmap showing the relative normalized expression of lncRNAs in the PC-3 experimental (test) group vs the LNCaP control group. Upregulated genes are shown in red while the downregulated genes are shown in green. The colour intensity is a reflection of the degree of upregulation or downregulation, while ± 20.08 represents the highest and the lowest value for up/down regulation.

In this study, PCAT29, wells 18 and 24, are shown to be the most downregulated lncRNA with a fold change of -349.44412 (Table 4.3). PCAT29 is a tumour suppressor, which suppresses the metastasis of PCa cells.[98] The most upregulated lncRNA was KRT81 with a fold change of 1109964.44. However, this fold change may need to be further validated in future studies using RT-qPCR. This lncRNA is part of the multigene keratin family that is expressed in multiple types of epithelia.[128] KRT81 is expressed in the hair cortex, and it is one of the main hair proteins. It has also been found to be upregulated in breast cancer cell lines and the lymph nodes of metastatic breast carcinoma.[128]

4.3. *In silico* bioinformatics analysis of differentially expressed lncRNAs in PCa cell lines

To determine the interactions of the differentially expressed lncRNAs from the PrimePCR Array analysis, bioinformatics tools specific to lncRNAs were used. These programmes were Long non-coding RNA related Sets and Enrichment Analysis (LncSEA) for annotation and enrichment analysis and DIANA-miRPath for the visualization of the miRNAs that belong to GO categories and regulate specific pathways. Cytoscape was used to visualize the ceRNA

network of TERC-associated miRNAs and mRNAs, which were identified using TargetScan and TarBase. The findings of these analyses are presented below.

4.3.1. LncSEA analysis of differentially expressed lncRNAs

The LncSEA database was used to determine the lncRNA annotations and enrichment. The lncRNAs selected for this analysis were the 36 upregulated and 12 downregulated lncRNAs that were profiled in section 3.8 using the PrimePCR Array. The LncSEA database showed the enrichment of a total of 20 upregulated lncRNAs and nine downregulated genes. The separate LncSEA results for upregulated and downregulated lncRNAs were combined as DEGs, those in the different classes (cancer hallmarks, disease, methylation patterns, exosome, drugs, subcellular localization, and miRNAs) are shown in Figures 4.5-4.11. For the full table of LncSEA results, please see Supplementary Tables C3.1 and C3.2. Figure 4.5 refers to the lncRNAs involved in the cancer hallmark class.

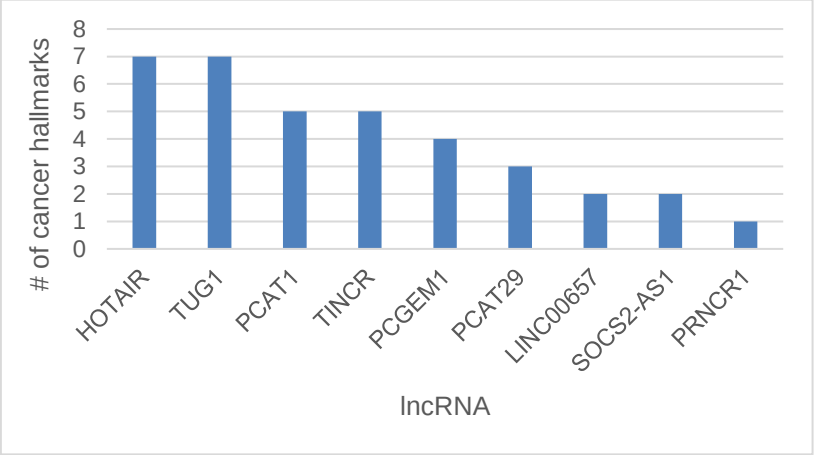


Figure 4.5: Graph showing the differentially expressed lncRNAs in PC-3 cells. Only lncRNAs shown to have an involvement in the cancer hallmarks are represented.

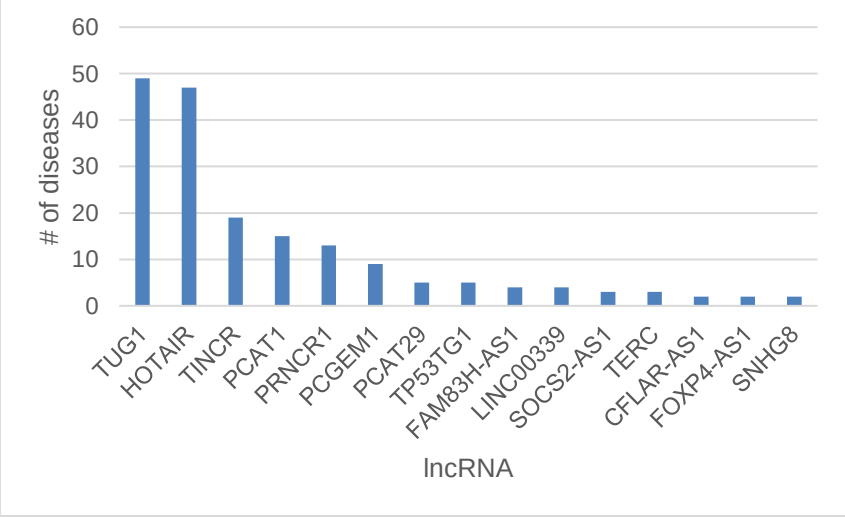


Figure 4.6: Graph showing the differentially expressed lncRNAs in PC-3 cells in the disease class of LncSEA. Some of the diseases included PCa, pre-eclampsia and glioma.

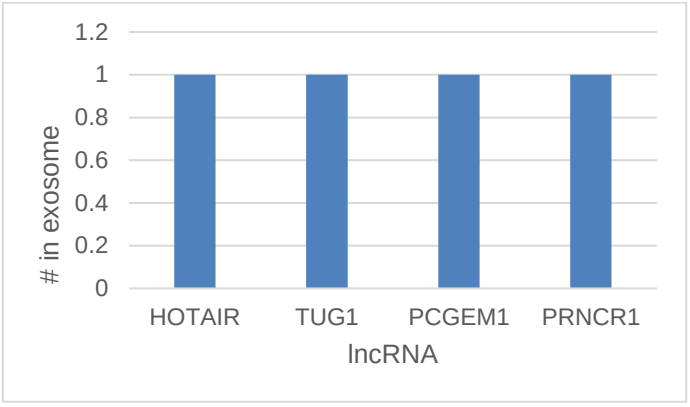


Figure 4.7: Graph showing the differentially expressed lncRNAs in PC-3 cells in the exosome class of LncSEA. These were the lncRNAs associated with the exosomal functions.

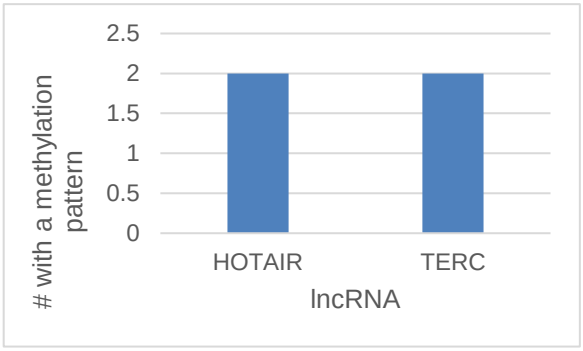


Figure 4.8: Graph showing the differentially expressed lncRNAs in PC-3 cells in the methylation pattern class of LncSEA. These were the lncRNAs that were hypermethylated.

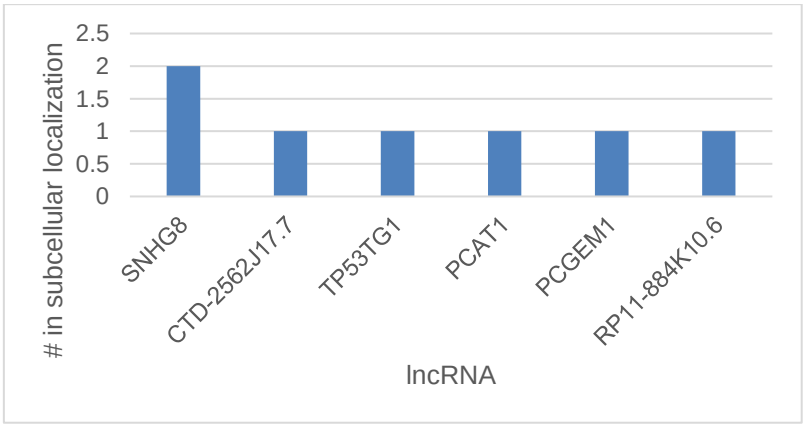


Figure 4.9: Graph showing the differentially expressed lncRNAs in PC-3 cells in the subcellular localization class of LncSEA. These lncRNAs were shown to be present in the exosome and nucleus. The number of times each lncRNA was shown to be subcellular localized is shown on the y-axis.

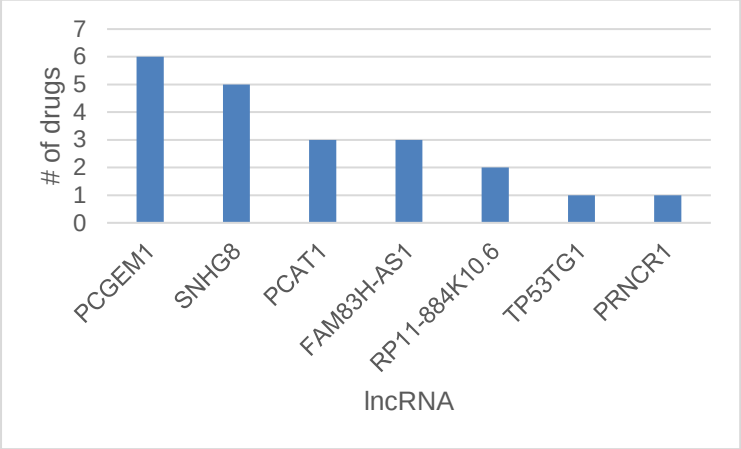


Figure 4.10: Graph showing the differentially expressed lncRNAs in PC-3 cells in the cancer treatment drugs class of LncSEA. These lncRNAs, including PCAT1, were shown to be drug targets in chemotherapy drugs such as Irinotecan.

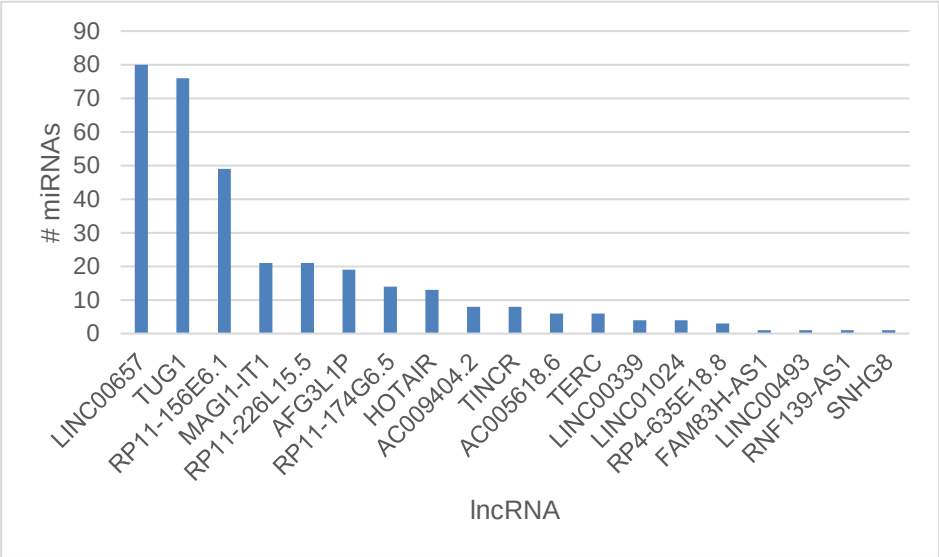


Figure 4.11: Graph showing the differentially expressed lncRNAs in PC-3 cells. Only lncRNAs shown to interact with miRNAs are represented.

The LncSEA platform analysed seven classes for the dataset including cancer hallmarks, disease, exosome, methylation pattern, subcellular localization, drug treatment, and miRNA. In Figure 4.5, the lncRNAs HOTAIR and TUG1 were associated with the most cancer hallmarks: namely, proliferation, apoptosis, metastasis, invasion, migration, prognosis, and EMT. This was followed by PCAT1 and TINCR which are involved in six hallmarks each. TUG1 and HOTAIR were also associated with the most diseases (49 and 47 diseases, respectively) (Figure 4.6). Most of these diseases were various types of cancers. Figure 4.7 shows exosomal lncRNAs. These lncRNAs may have regulatory functions of cell proliferation and metastasis. They may also be useful as cancer biomarkers due to their

association with clinicopathological cancer characteristics.[129] TERC and HOTAIR were the only two lncRNAs shown to have a methylation pattern (Figure 4.8), while lncRNAs in the subcellular localization class included CTD-2562J17.7, SNHG8, TP53TG1 (Figure 4.9). Figure 4.10 showed the lncRNAs that are drug targets. These drugs include chemotherapy drugs Lapatinib and Irinotecan, The miRNA class was associated with the most lncRNAs from the input dataset. These lncRNAs included TUG1, RP11-156E6.1, LINC00657, HOTAIR, and TERC (Figure 4.11). This analysis indicates that lncRNAs have important interactions with miRNAs, thus forming ceRNA networks. Multiple lncRNAs are linked to the same miRNA, such as hsa-miR-320a. This suggests that lncRNAs can be regulated by several lncRNAs, and visa versa. The cancer hallmark class is further differentiated into subclasses and enrichment sets such as epithelial-mesenchymal transition (EMT), proliferation, and invasion. The association between the cancer hallmark subclasses and the up and downregulated lncRNAs are shown in Figure 4.12.

These findings show that lncRNAs that were analysed were either up or downregulated and found to be important in PCa progression. In Figure 4.12, the proliferation cancer hallmark has the highest -log10 p-value. In (A), the lowest -log10 p-value was found in the EMT cancer hallmark. In the downregulated genes (B), the lowest -log10 p-value is prognosis. Figure 4.13 shows a combined graph of all the lncRNAs from the LncSEA analysis arranged in their several classes, to provide an overview of the lncRNAs across all the classes.

In LncSEA, the classes with the most lncRNAs were miRNA and disease. In these categories, LINC00657, HOTAIR, RP11-156E6.1 and TUG1 are the lncRNAs that are most represented. This finding suggests that they have significance in PCa progression towards advanced metastasis. Furthermore, the precise role of TERC lncRNA in PCa progression remains to be largely elucidated.

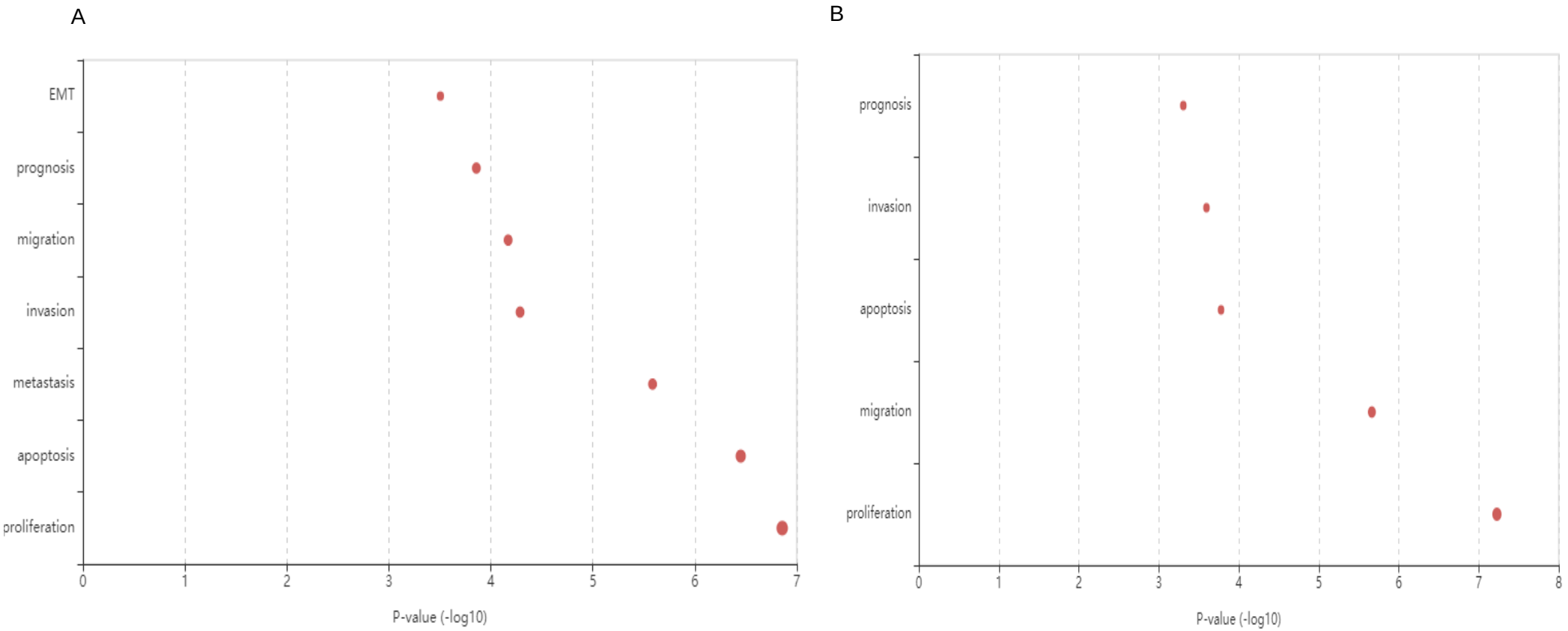


Figure 4.12: Graphs generated by LncSEA showing the p-value (-log10) and the corresponding enrichment sets for (A) the upregulated lncRNAs and the (B) downregulated lncRNAs that are in cancer hallmarks. The p-value expressed as -log10, allows for better visualization of the p-value, which is a very small value. In the upregulated lncRNAs, the proliferation hallmark is associated with the most lncRNAs. The downregulated lncRNAs are not involved in metastasis and EMT, compared to the upregulated lncRNAs.

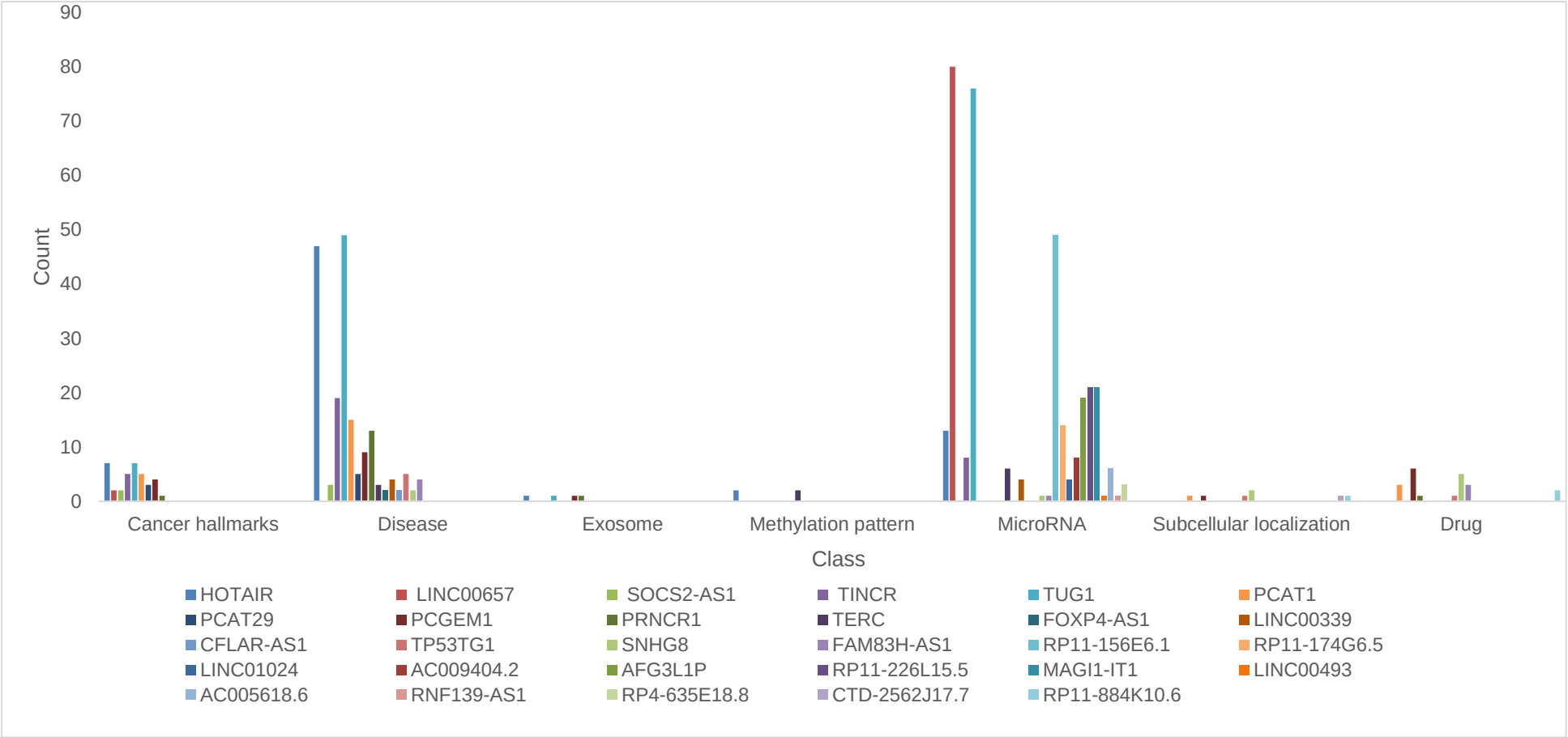


Figure 4.13: Graph showing the differentially expressed lncRNAs in PC-3 cells in all LncSEA classes. A total of 29 lncRNAs were found including TERC (dark purple), HOTAIR (blue) and PCAT1 (orange). The lncRNAs with the highest values are HOTAIR in disease (blue), LINC00657 (red) and RP11-156E6.1 in microRNA (light blue) and TUG1 in disease and microRNA classes (turquoise). Very few lncRNAs were identified in the exosome, and only two were identified to have a methylation pattern. In the drug class, only a few lncRNAs were identified as well as in subcellular localization.

4.3.2. DIANA-miRPath analysis of the differentially expressed lncRNAs and their associated miRNA interactions

MiRNA and lncRNA interactions were further analysed using DIANA-miRPath v3.0. This web server provides miRNA pathway analysis, and it allows for the visualization of the miRNA pathways including the KEGG and GO pathways. The input into this database were those miRNAs that were shown to interact with the differentially expressed lncRNAs in LncSEA (Supplementary Tables C3.1 and C3.2). From the total of 103 miRNAs which were found by LncSEA, a subset of 30 miRNAs were selected based on the p-value (≤1.05E-06). The miRNAs with significant KEGG pathways and GO terms are shown in Tables 4.4 and 4.5, Figure 4.14 and Figure 4.15.

4.3.2.1. Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment pathway analysis

The two categories that are linked to the most miRNAs in the KEGG analysis are summarized in Table 4.4. These include hsa-miR-106a-5p, hsa-miR-320b and hsa-let-7b-5p which are implicated in various cancers. The entire KEGG analysis table is shown in Appendix C: Table C4.1.

Table 4.4: Top two KEGG pathways associated with the most miRNAs. The KEGG pathway ID is shown in brackets.

Number of miRNAs	KEGG Pathway	Associated miRNAs
15	Signalling pathways regulating pluripotency of stem cells (hsa04550)	hsa-miR-4698, hsa-let-7b-5p, hsa-miR-106b-5p, hsa-miR-20a-5p, hsa-miR-548n, hsa-let-7d-5p, hsa-miR-320a, hsa-miR-186-5p, hsa-miR-320b, hsa-miR-320d, hsa-let-7a-3p, hsa-miR-548d-5p, hsa-let-7f-5p, hsa-let-7i-5p, hsa-let-7f-2-3p
11	TGF-beta signalling pathway (hsa04350)	hsa-miR-4698, hsa-miR-106a-5p, hsa-miR-106b-5p, hsa-miR-20a-5p, hsa-miR-519d-3p, hsa-miR-548n, hsa-miR-101-3p, hsa-miR-186-5p, hsa-let-7a-3p. hsa-miR-548d-5p, hsa-let-7f-2-3p

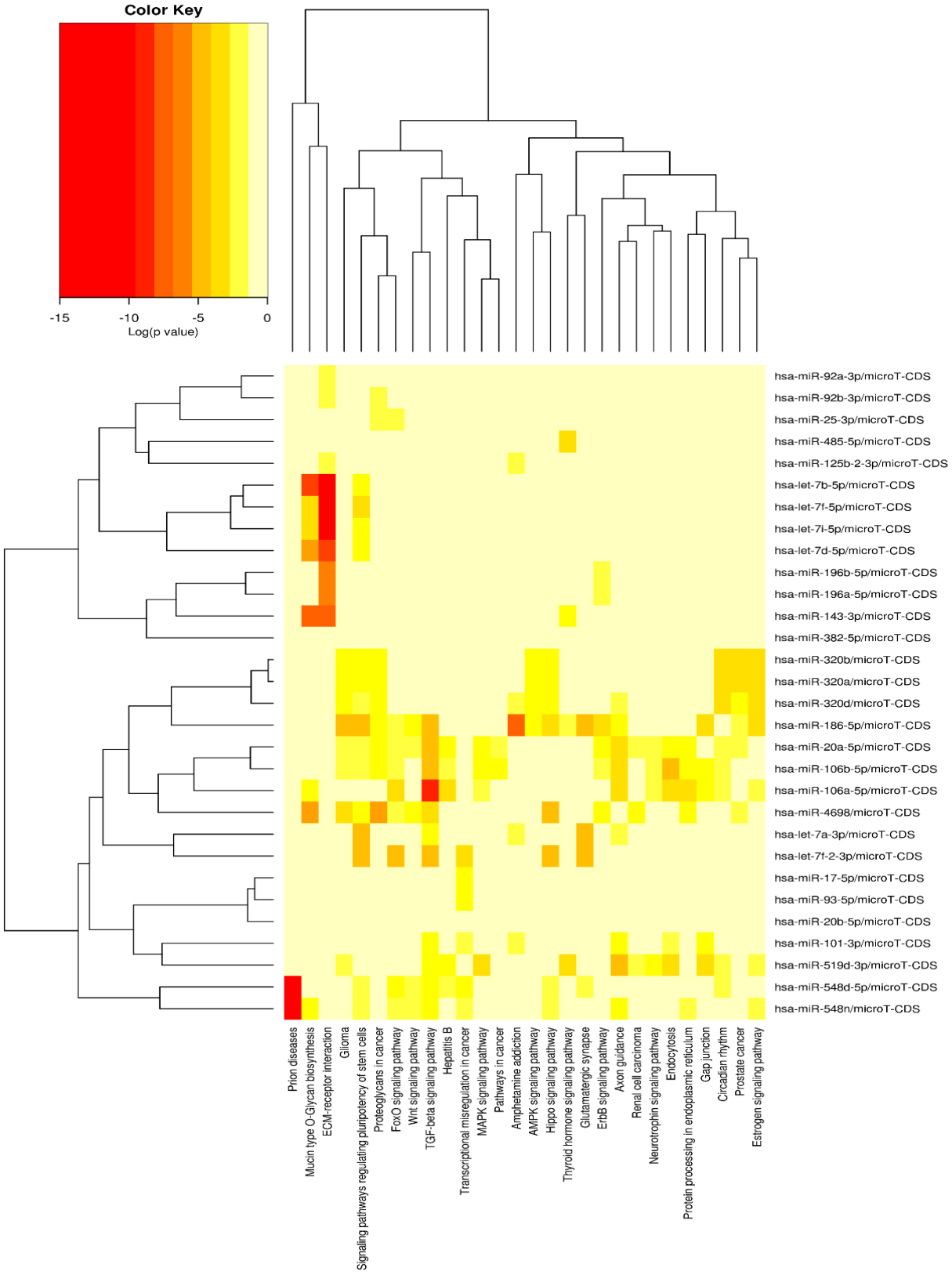


Figure 4.14: Heatmap of the miRNA versus KEGG pathway analysis. This heatmap shows the various miRNAs that are linked to the differentially expressed lncRNAs and their levels of enrichment in KEGG pathways. p-value<0.05; microT<0.8; Fischer's Test analysis method with fold discovery rate (FDR) correction. Lower log(p-values) is associated with higher

enrichment and a deeper red colour. The miRNAs are shown to be associated with various disease pathways, including PCa which is significantly enriched by the hsa-miR-320 family mainly. The hsa-miR-320 family is linked to TERC lncRNA, which is overexpression in PCa.

The KEGG pathway analysis and biological process GO analysis showed 28 enriched pathways. In Figure 4.14, the KEGG pathway extracellular matrix (ECM)-receptor interaction, is most enriched by the hsa-miR-let7 family. Prion disease shows the most enrichment by hsa-miR-548d-5p and hsa-miR-458n. Hsa-miR-106a-5p shows higher enrichment of the TGF-β-signalling pathway which is involved in tumorigenesis. Prostate cancer and other pathways related to PCa were also identified. KEGG/GO analysis outliers would be biological processes or pathways that are unrelated to cancer.

4.3.2.2. Gene Ontology (GO) enrichment analysis

The two categories that are linked to the most miRNAs in the GO analysis are summarized in Table 4.5. These also include hsa-miR-106a-5p and has-miR-320b and hsa-let-7b-5p. The full GO analysis table is shown in Appendix C: Table C4.2.

Table 4.5: Top two GO categories associated with the most miRNAs in the biological process subcategory, and the miRNA names. These miRNAs are associated with the differentially expressed lncRNAs in OCa. The GO category IDs are shown in brackets.

Number of miRNAs	GO Category	Associated miRNAs
30	Biosynthetic process (GO:0009058) and cellular nitrogen compound metabolic process (GO:0034641)	hsa-miR-4698, hsa-let-7b-5p, hsa-miR-485-5p, hsa-miR-106a-5p, hsa-miR-106b-5p, hsa-miR-17-5p, hsa-miR-20a-5p, hsa-miR-20b-5p, hsa-miR-519d-3p, hsa-miR-93-5p, hsa-miR-92b-3p, hsa-miR-548n, hsa-miR-143-3p, hsa-miR-196a-5p, hsa-miR-196b-5p, hsa-let-7d-5p, hsa-miR-320a, hsa-miR-101-3p, hsa-miR-125b-2-3p, hsa-miR-92a-3p, hsa-miR-186-5p, hsa-miR-382-5p, hsa-miR-320b, hsa-miR-320d, hsa-let-7a-3p, hsa-miR-548d-5p, hsa-miR-25-3p, hsa-let-7f-5p, hsa-let-7i-5p, hsa-let-7f-2-3p

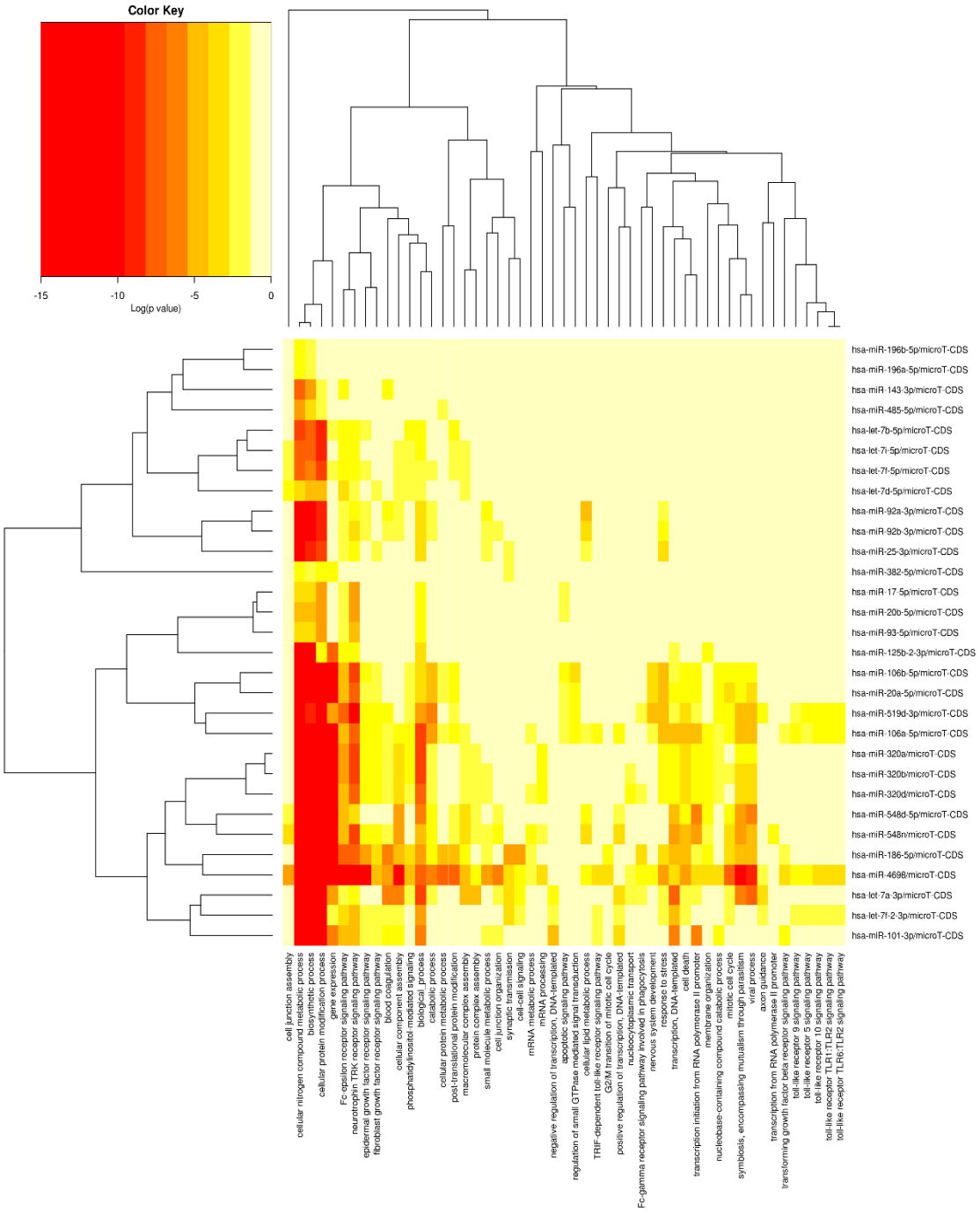


Figure 4.15: Heatmap showing the miRNAs that are associated with the differentially expressed lncRNAs versus the biological process GO subcategory (generated by DIANA-miRPath v3.0). p-value<0.05; microT<0.8; Fischer's Test analysis method with FDR correction. The intensity of the red colour indicates higher enrichment.

The GO analysis showed 51 enriched processes. The most enriched GO categories are the biosynthetic process, cellular nitrogen compound metabolic process, and cellular protein modification process (Figure 4.15). These processes are enriched by various miRNAs including the hsa-miR-106 a and b family and the hsa-miR-320 family which are involved in various cancers, including PCa.

4.3.3. LncRNA-miRNA-mRNA competing endogenous RNA (ceRNA) network

For the construction of the ceRNA network, TERC was selected as the lncRNA of interest because it can be hypermethylated and it interacts with the tumour suppressive miR-320 family and other miRNAs that have tumour-suppressive activity namely hsa-miR-338-3p and hsa-miR-4429 (Table 4.6). Further analysis and visualization of the TERC-associated miRNA-mRNA interactions were done using various miRNA gene target databases and Cytoscape v3.9.1. A total of 45 genes were predicted to interact with the TERC-associated miRNAs. Of these 45 targets, 24 of them are involved in tumour progression in various cancers (Figure 4.16), including disintegrin and metalloproteinase domain-containing protein 10 (ADAM10) which enhances PCa metastasis.

Table 4.6: ceRNA network basis showing the miRNAs associated with TERC based on p-value.

LncRNA	MiRNA	P-value
TERC	hsa-miR-320a	2.59E-07
	hsa-miR-320b	8.25E-07
	hsa-miR-320d	8.25E-07
	hsa-miR-320c	1.64E-05
	hsa-miR-4429	1.97E-05
	hsa-miR-338-3p	3.61E-05

The ceRNA network shows that some genes are targeted by more than one miRNA in the network. The nuclear fragile X mental retardation protein interacting protein 2 (NUFIP2) for example, interacts with both hsa-miR-320a and hsa-miR-338-3p. Several of these TERC-associated mRNA targets from the analysis are linked to the progression of PCa, such as ADAM10, Musashi RNA binding protein 2 (MSI2), and cyclin-dependent kinase 6 (CDK6) amongst others. Unlike in PCa, TMEM106B and RAB14 targets have both been characterized in the literature as drivers of cell proliferation in lung cancer.[130-131] Their presence in the ceRNA network of this current study may show that they are linked to PCa cell proliferation. The analysis also showed that hsa-miR-320c is the only miRNA that does not have a common mRNA target with another miRNA in this ceRNA network.[130-131]The collection of mRNA targets in the databases is dependent on experimental data that has been uploaded into the databases. The hsa-miR-320c miRNA may not be as well studied as the other miRNAs in the ceRNA network, therefore, the common mRNA targets with the other members may not have been established as yet.

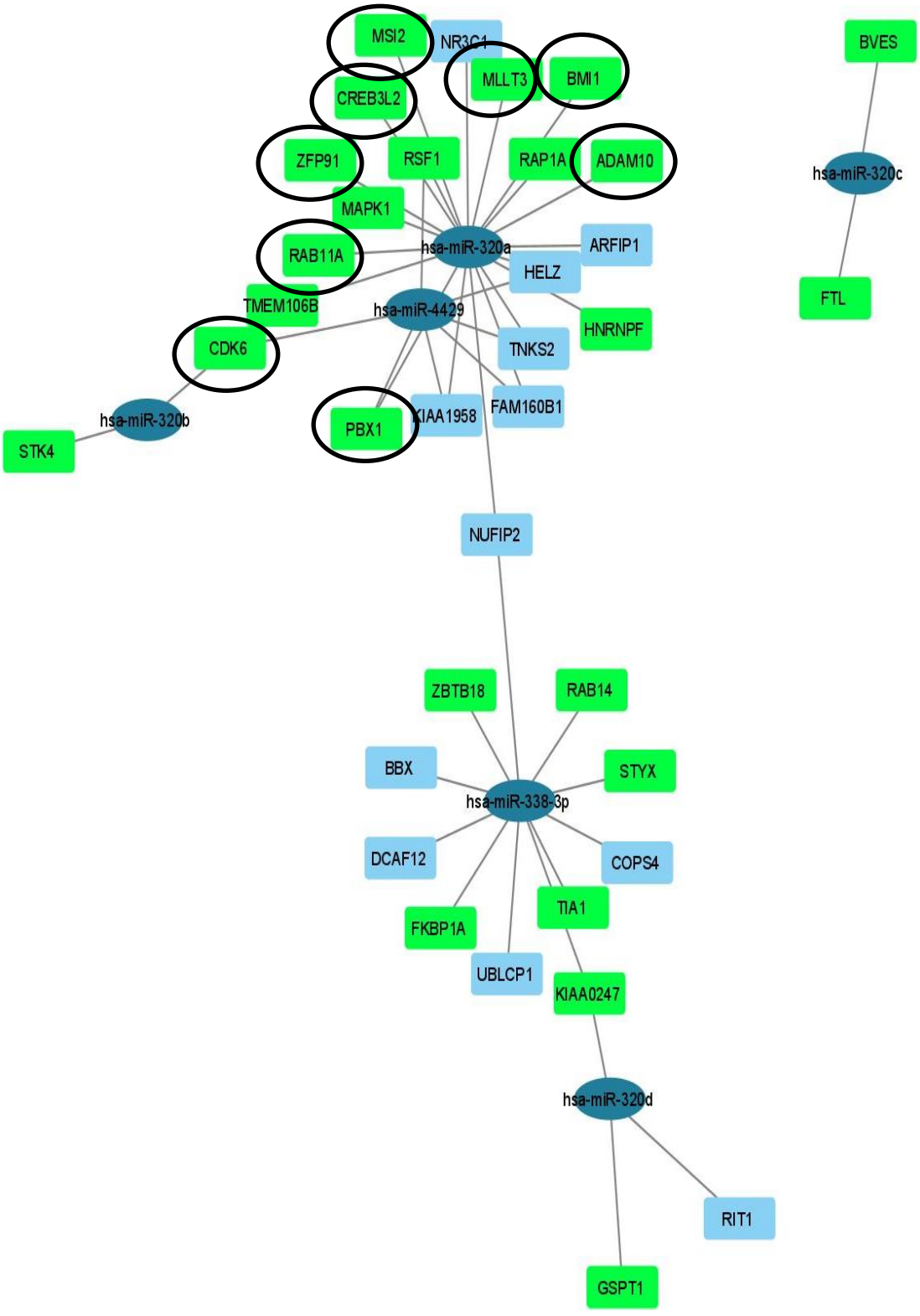


Figure 4.16: The competing endogenous RNA (ceRNA) network of TERC lncRNA using TERC-associated miRNAs. The miRNAs are represented in a dark blue ellipse, and the gene targets are represented in blue rectangles. The targets that are linked to cancer

progression are represented in green rectangles. The circled green targets are linked to PCa progression specifically.

4.4. Validation of the selected lncRNA using Real-Time quantitative PCR (RT-qPCR)

Based on the PrimePCR data and bioinformatics analysis, one of the upregulated lncRNAs, TERC, was selected for downstream RT-qPCR. TERC has been reported to be overexpressed in various tumours including PCa.[97] However, precise TERC oncogenic mechanisms remain to be largely elucidated. RT-qPCR gene expression method is highly sensitive but inexpensive when compared to RNA-sequencing, and it was used for the confirmation of the relative gene expression levels of TERC.

RT-qPCR was done using the AB 7500 RT-qPCR system. The Cq values were obtained, and fold change calculations were done as described in section 3.10.3 (Chapter 3). Figure 4.17 shows the melt curve of TERC from the RT-qPCR analysis. Following this TERC primer specificity run, TERC was then quantified to verify its upregulation observed in the Prime PCR Array data. This RT-qPCR run confirmed the TERC upregulation of 4.29-fold in high-grade PCa PC-3 cells compared to low-grade PCa LNCaP cells, Figure 4.18.

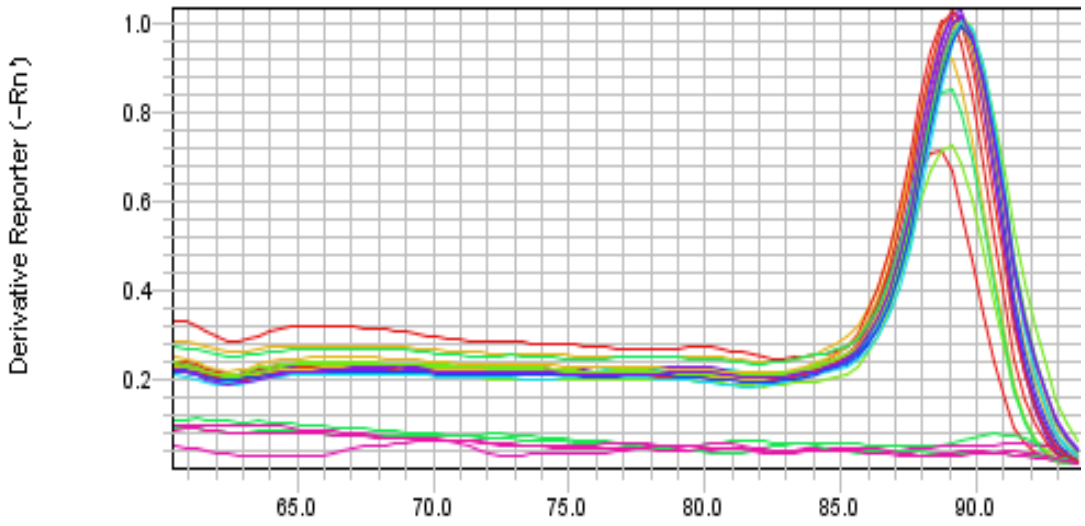


Figure 4.17: Melt curve generated using TERC primers and the cDNA samples from PC-3 and LNCaP cell lines. The melt curve shows that the samples have similar melting temperatures of about 89°C, this indicates TERC primer specificity. Non-template samples were used as negative controls.

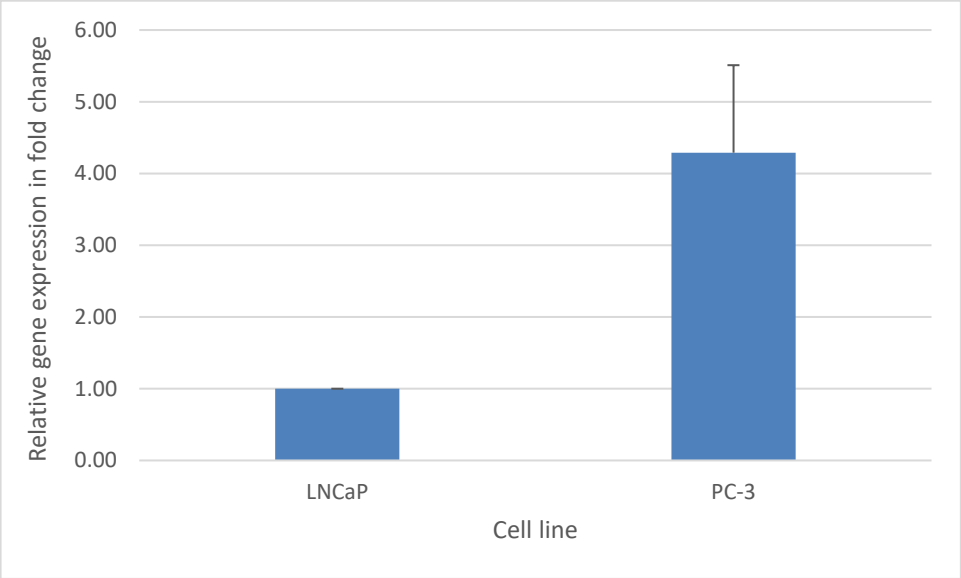


Figure 4.18: Graph showing the TERC lncRNA gene expression levels in the PC-3 test compared to LNCaP control cells. A 4.29 TERC upregulation is observed in PC-3 cells compared to LNCaP. The error bars denote the SEM of 0 and 1.22 for LNCaP and PC-3, respectively.

TERC expression, although upregulated, was not statistically significant (p-value >0.05). Both cell lines are metastatic, however, they have different tumour grades. Further investigation is required using normal and benign cells, as this implies that TERC may have a role in PCa metastasis, irrespective of the tumour grade.

CHAPTER 5

DISCUSSION AND CONCLUSION

5.1. Discussion

This chapter discusses the key findings of this study. While differential expression of the profiled lncRNAs was subject to *in silico* bioinformatics analysis to map the biological pathways enriched by these lncRNA targets, the involvement of these lncRNA gene targets and pathways in PCa progression is discussed. TERC, the lncRNA of interest in this study and its mechanisms of action in PCa metastasis is further discussed. In addition, the limitations of this study and future perspectives are also outlined.

5.1.1. The effect of differential expression patterns of lncRNAs in PCa progression

The differential expression of lncRNAs has been reported in various cancers, including PCa. Most of the lncRNAs implicated in PCa prognosis have also been identified in this study, although some inconsistencies with the literature were revealed. The upregulation of lncRNAs PRNCR1 and PCGEM1 are known to drive the progression of aggressive PCa.[104] PCGEM1 does this by enhancing the gene activation that is mediated by AR. However, this is not the case in this study as downregulation of these two lncRNAs was observed (Table 4.3). The most upregulated lncRNA is KRT81 (Table 4.2) with a fold change of 1109964.44. This lncRNA is not very well described in PCa. In a study by Zhang et al.[132], it was shown that KRT81 knockdown regulates interleukin-8 to inhibit melanoma metastatic progression.[132] In breast cancer, it has also been reported to aid in breast cancer cell migration and invasion, as KRT81 knockdown inhibited these hallmarks.[128] Although reports are lacking in oncogenic lncRNAs in PCa, KRT81 upregulation in this study is shown to be associated with PCa metastasis, although this warrants further elucidation of this oncogenic lncRNA in PCa.

TUG1 and LINC00657 were both upregulated in PC-3 cells with a fold change of 3.24 and 2.74, respectively. The upregulation of TUG1 is consistent with the literature. TUG1 is a lncRNA that has been proven to be upregulated in various malignant cancers[133] and the overexpression of this lncRNA is closely associated with cancer prognosis.[134] TUG1 has been shown to competitively bind miRNAs, regulating cyclin-dependent kinase inhibitors, thus regulating cancer proliferation. LINC00657 is a lncRNA that is associated with DNA damage[135] which is carcinogenic in gastric cancer.[136] In a breast cancer study, Golgi phosphoprotein 3 (GOLPH3) protein levels were shown to be decreased upon transfection with si-LINC00657. The protein GOLPH3 is associated with a poor prognosis of breast cancer. Furthermore, transfection with si-LINC00657 resulted in reduced migration, invasion, and proliferation of breast cancer cells.[137] On the contrary, it was also shown to have tumour suppression properties in glioblastoma by miRNA sponging.[138] This indicates that the same

lncRNA may serve as a double-edged sword in different cancers, either promoting or inhibiting tumour progression. Studying the precise mechanisms of unique lncRNAs in different cancers and tumour grades is necessary.

5.1.2. Differential lncRNA methylation patterns and lncRNA/miRNA interactions in PCa progression

Differential methylation patterns play a significant role in lncRNA regulation. These lncRNA epigenetic regulatory mechanisms have been implicated in various cancers, including PCa.[139] Hypermethylation can drive oncogenesis[140] and generally suppresses gene expression. In this study, the upregulated lncRNAs, HOTAIR and TERC were demonstrated by *in silico* analysis to be hypermethylated, and these hypermethylation patterns are in conjunction with the literature. In gastric cancer, HOTAIR has been shown to promote EMT. It induces inhibition of E-cadherin transcription by mediating the switching of histone H3 lysine 27 from being acetylated to being methylated.[141]. In this study, LncSEA revealed that HOTAIR is hypermethylated and it is an exosomal lncRNA.

MiRNAs are essential noncoding molecules regulating gene expression. They do this mainly by their interaction with the 3' untranslated region of mRNA molecules to inhibit translation and thus gene expression. However, they can also interact with promoters, the mRNA sequence and other regions to promote gene expression. Alternatively, their interaction with the coding sequence or the 5' UTR can also silence gene expression.[142] LncRNA-miRNA-mRNA regulatory networks are critical for the development and progression of cancer.[143] MiRNA sponging is the main mechanism in which lncRNAs interact with miRNAs. This involves microRNA response elements (MREs) and competing endogenous RNAs (ceRNAs).[144] CeRNAs function by competing with miRNAs to release gene inhibition and the sponging effect is the basis of a network of various biochemical processes that are mediated by lncRNAs, miRNAs and their related molecules. TUG 1 is an example of a lncRNA that regulates cyclin-dependent kinase inhibitors, competes with miRNAs and is therefore implicated in carcinogenesis.[145] In terms of regulation, miRNAs can regulate lncRNAs and lncRNAs can also regulate miRNAs. MiRNAs can degrade lncRNAs and affect their expression via DNA methylation as well.[144] This can be through hypermethylation of the promoter region of the lncRNA as shown in an article by Dong et al.[146] where aberrantly methylated MEG3 promoter silenced its expression. The silencing of this lncRNA resulted in reduced oesophageal cancer proliferation. It is for these reasons that lncRNA-miRNA interactions are important for normal development and function, and how the aberrant expression of lncRNAs can negatively affect its miRNA interactions thus leading to disease phenotypes such as cancer. Enormous amounts of data in this study revealed various

lncRNA/miRNA interactions, where a single lncRNA can regulate multiple miRNAs and vice versa. This lncRNA/miRNA relationship has been reported in various cancers,[147-148] including PCa[101,149] and is likely to play an important role in PCa progression. In gallbladder cancer, lncRNA HOTAIR has been shown to directly target c-*Myc*, which is a driver of cancer, through its interaction with hsa-miR-130a. The negative regulation of hsa-miR-130a thus promoted cancer cell proliferation and invasion.[147]

Furthermore, HOTAIR/hsa-miR-106-5p a and b family members interactions seem to have tumour-promoting properties.[150] In various studies, hsa-miR-106a-5p and hsa-miR-106b-5p promote PCa proliferation through Phosphatase and Tensin Homolog deleted on Chromosome 10 (PTEN).[151-152] When overexpressed, hsa-miR-106a promoted PC-3 cell growth by suppressing PTEN. The overexpression of PTEN inhibited cell growth, which is the same effect as the loss of hsa-miR-106a.[151] Additionally, HOTAIR/ hsa-miR-20a-5p and hsa-miR-20b-5p have also been revealed to play an important role in PCa progression. Hsa-miR-20b-5p has been reported to affect the expression of two tumour suppressor genes BIM and PTEN. The miRNA hsa-miR-20b-5p was shown to post-transcriptionally repress these genes resulting in the T-cell acute lymphoblastic leukaemia cells gaining protection from apoptosis. This caused the enhanced proliferation of these cells.[153] The miRNA miR-106b has been shown to be aberrantly expressed in multiple tumours and cancer cell lines including PCa. It works by directly affecting the genes that are involved in cancer tumorigenesis, migration, proliferation and metastasis.[154] This miRNA has shown the potential to be a biomarker in various cancers for the detection of a tumour and its prognostic assessment.[155]

Interestingly, TERC is shown to interact with several miRNAs including the hsa-miR-320 family, hsa-miR-338-3p, and hsa-miR-4429. The bioinformatic analysis revealed that TERC lncRNA undergoes hypermethylation and regulates the hsa-miR-320 family. The decreased hsa-miR-320 gene family expression has been reported in various tumours including PCa.[118] However, the precise mechanisms of the hsa-miR-320 gene family in PCa remain to be elucidated. This indicates a correlation between TERC hypermethylation and downregulation of hsa-miR-320 gene family expression in PCa. TERC lncRNA promotes PCa cell proliferation, migration, invasion, and metastasis by sponging the hsa-miR-320 family. This TERC/hsa-miR-320 regulatory axis holds the potential to understand PCa prognostic mechanisms and may also be targeted for novel therapeutics. The miR-320 family of miRNAs has been shown to be involved in PCa prognosis and this shows that it has diagnostic potential.[118] Furthermore, it has been reported that the serum levels of these miRNAs were significantly different between patients with a low tumour stage and a high tumour stage.[118] The tumour-suppressive quality of the miR-320 family makes it a potential

biomarker for cancer prognosis and cancer treatment. The members of this family are linked to inhibited cell proliferation, metastasis and EMT.[156] The miR-320 family has been reported to be downregulated in the process of tumorigenesis in numerous studies.[157-158] Furthermore, both hsa-miR-338-3p and hsa-miR-4429 have been shown to have tumour suppressive qualities in glioblastoma,[159] cervical cancer[160] and PCa.[161] The high expression of hsa-miR-17-5p, the miRNA linked to HOTAIR, has also been implicated in PCa.[162] It is specifically associated with the poor prognosis of PCa and can be used to predict disease progression. Furthermore, LINC00657 lncRNA, upregulated in this study, was reported to be directly associated with oesophageal cancer progression through LINC00657/miR-26a-5p/CKS2 ceRNA.[143] LINC00657 knockdown successfully induced apoptosis and suppressed the invasion and proliferation of the oesophageal cancer cells.[143] These findings substantiated by existing reports demonstrate that lncRNAs can be used to predict the path of disease progression by studying their differential expression, as well as the molecules they interact with.

5.1.3. Pathway enrichment analysis by the differentially expressed lncRNAs

PCAT1, PCGEM1, and PRNCR1 were all shown to be downregulated in PC-3 cells. In the LncSEA analysis, these lncRNAs were revealed to be associated with PCa pathogenesis. Interestingly, they were also associated with the proliferation cancer hallmark while PCAT1 and PCGEM1 were further implicated in cancer cell invasion, migration, and apoptosis. Figure 4.14 showed that the let-7 family of miRNAs is implicated in ECM-receptor interaction. ECM-receptor interaction pathways have been proven to be involved in various cancers.[163] This family directly interacts with the lncRNAs TUG1 and HOTAIR which have been shown to be significant lncRNAs in PCa. In a study by Andersen et al.[164], ECM was shown to be upregulated in PCa.[164] The KEGG pathway enrichment analysis also indicated that the TGF-β-signalling pathway was enriched by miRNA, hsa-miR-106a-5p, which was shown to interact with various lncRNAs including HOTAIR, LINC00657, MAGI1-IT1, and TUG1 in this study. The regulation of every signalling network, at multiple levels, is necessary for the normal physiological function of these pathways.[165] The TGF-β cytokine regulates cell differentiation, cell death and proliferation[166] by its interaction with kinases. TGF-β-signalling dysregulation is therefore associated with the pathogenesis of various diseases and the tumorigenesis of cancers, PCa included[166] and its blockage is the target of cancer therapies as this suppresses metastasis and tumour angiogenesis.[167]

The results discussed in this study are supported by various literature sources. The investigation of lncRNAs and miRNA interactions is a cornerstone to further understanding

the exact molecular mechanisms of gene expression regulation by lncRNAs. This is a developing field of study that has great potential to elucidate the processes of tumourigenesis and *in silico* bioinformatics tools are key to this elucidation. The next step was to further analyse the lncRNAs in PCa and the complex lncRNA-miRNA-mRNA networks.

5.1.4. ceRNA networks reveal that several cancer-related genes are targeted by TERC lncRNA through miRNA sponging

The generation of a basic ceRNA network based on the lncRNA TERC revealed some interesting miRNA-mRNA interactions (Figure 4.16). The ADAM10 mRNA was shown to be targeted by tumour suppressor miRNA hsa-miR-320a. ADAM10 protein is an important enzyme for proteins such as receptors and growth factors which are membrane-anchored proteins. Its expression is seen in many tissue types, and it plays roles in cell migration and adhesion. Its expression is elevated in PCa where it is involved in malignant transformation.[168] In a study by Cai et al.[168], ADAM 10 was shown to contribute to PCa metastasis by cleaving ephrin-A5 which is the primary ligand of EphA3, a receptor tyrosine kinase. The interaction of miRNA-320a and this mRNA suggests that miRNA-320a may exert its tumour suppressive attributes on ADAM10 to inhibit its action. This also suggests that if TERC lncRNA sponges hsa-miR-320a, ADAM10 may continue to promote tumour metastasis.

Transmembrane protein 106B (TMEM106B) is a type II transmembrane protein associated with Alzheimer's disease when it is repressed. It is under studied in the cancer landscape however, recently it has been shown to mediate lung cancer metastasis.[130] In this study, the authors state that the elevated expression of this gene predicts poor overall survival.[130] In the current study, TMEM106B was predicted to interact with hsa-miR-320a, which suggests that hsa-miR-320a may regulate the expression of TMEM106B mRNA. Additionally, hsa-miR-338-3p is a miRNA that also has tumour suppressive capabilities. This has been shown in gastric cancer[169] and PCa.[170] In PCa specifically, it was shown to be downregulated, leading to PCa proliferation and invasion. In normal physiology, it was shown to suppress tumorigenesis by targeting the *RAB23* gene, a member of the Fold discovery rate RAS oncogene family (*RAB23*).[170] In the current study, however, *RAB23* was not predicted to interact with hsa-miR-338-3p. Instead, this miRNA was shown to interact with *RAB14* which has been shown to promote lung cancer invasion and proliferation.[131] *RAB14*'s relevance in cancer proliferation is further supported by evidence that its overexpression activated the AKT and Wnt signalling pathways in ovarian and gastric cancers.[131]

In PCa, hsa-miR-4429 suppresses cell proliferation by inactivation of the Wnt/β-catenin pathway, targeting distal-less homeobox 1. This pathway is involved in several cellular processes including cell growth and stem cell maintenance, and induces the invasiveness of PCa cells.[161] Furthermore, this miRNA was shown to sensitise cervical cells to irradiation,[160] demonstrating its tumour suppressive role. Cyclin-dependent kinase 6 (CDK6) was shown to be a target for this miRNA in the current study. Literature supports that this miRNA is involved in suppressing colorectal cancer progression via the miR-500a-3p/CDK axis.[171] This does not mean that CDK6 exclusively interacts with hsa-miR-500a-3p, as there may be several other miRNAs interacting with the same gene target as shown in Figure 4.16.

The hypermethylated TERC interacting miRNAs act as tumour suppressors, implicating the involvement of TERC in PCa progression. Coupled with the understanding that lncRNAs can function as ceRNAs to sponge or absorb miRNAs, one can conclude that TERC may be oncogenic, sponging tumour-suppressive miRNAs, thus promoting PCa progression (Figure 5.1).

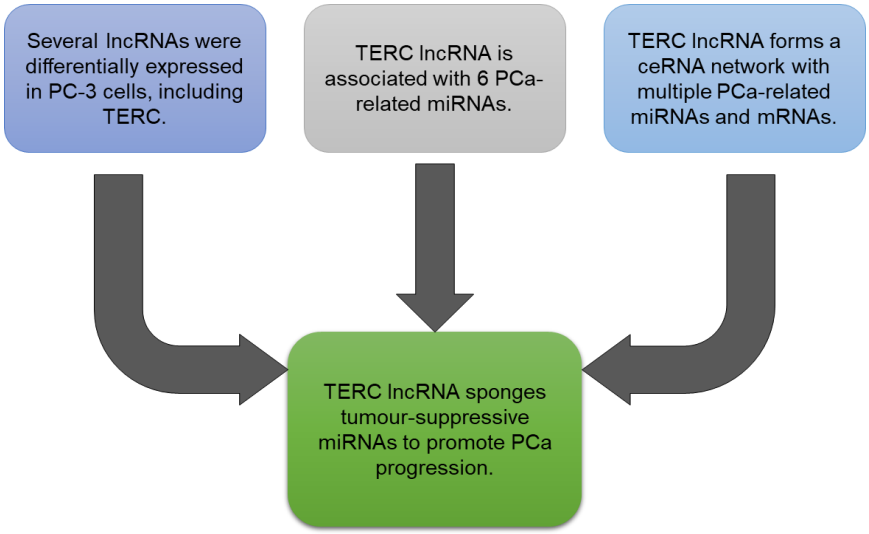


Figure 5.1: Diagram summarizing the main findings of this study. TERC lncRNA was differentially expressed in the PrimePCR Array. This was validated using RT-qPCR. TERC lncRNA was demonstrated to interact with six miRNAs and their associated mRNA targets for the progression of PCa.

5.2. Study limitations and perspectives

In the clinical setting, benign samples would have to be used for these differentially expressed genes to be used as prognostic markers, compared to low-grade and high-grade metastatic PCa samples. In future, one could explore the molecular profiling of lncRNAs in PCa tissues using samples from the SSA populations to see how this compares with the

literature on the lncRNAs expression in other regions. This would add much needed knowledge about the disease in African populations.

5.3. Conclusions

The burden of PCa continues to rise throughout the world; and therefore, the development of novel and more innovative prognostic and therapeutic approaches is necessary. The data and the information deduced from this project make worthy contributions to the existing knowledge on lncRNAs' role in PCa progression. Differential lncRNA gene expression was shown to be associated with PCa progression. The identified lncRNAs and associated enriched biological pathways provide better insights into mechanisms employed by the progression and poor prognosis of PCa. Such mechanisms include the epigenetic methylation patterns of these lncRNAs and their interactions with miRNAs. LncRNA-miRNA-mRNA interactions revealed an interplay between several miRNAs and gene targets, which shows the complexity of gene regulation by lncRNAs. Some of the aberrantly expressed lncRNAs revealed in this study such as the KRT81 lncRNA still lack data on its association with PCa. Thus, downstream experiments may include both wet-lab and *in silico* analysis to elucidate precise roles and associated mechanisms of these lncRNAs-miRNA-mRNA networks in PCa. Molecular profiling of cancer is emerging as a potent diagnostic, prognostic, and therapeutic target to improve overall patient survival.

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global cancer statistics 2020: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2021; 71(3):209-49. doi:10.3322/caac.21660
2. Culp MB, Soerjomataram I, Efstathiou JA, Bray F, Jemal A. Recent global patterns in prostate cancer incidence and mortality rates. Eur Urol. 2020; 77(1):38-52. doi:10.1016/j.eururo.2019.08.005
3. Rebbeck TR, Devesa SS, Chang B-L, Bunker CH, Cheng I, Cooney K, Eeles, R, Fernandez P, Giri VN, Gueye SM, Haiman CA, Henderson BE, Heyns CF. Global patterns of prostate cancer incidence, aggressiveness, and mortality in men of African descent. Prostate Cancer. 2013; 2013:560857-. doi:10.1155/2013/560857

4. Munretnam K, Alex L, Ramzi NH, Chahil JK, Kavitha IS, Hashim NA, Lye SH, Velapasmy S, Ler LW. Association of genetic and non-genetic risk factors with the development of prostate cancer in Malaysian men. Mol Biol Rep. 2014; 41(4):2501-8. doi:10.1007/s11033-014-3107-8
5. Ferlay J, Ervik M, Lam F, Colombet M, Mery L, Piñeros M, Znaor A, Soerjomataram IFB. [Internet]. Global cancer observatory: cancer today. [12 December 2022]. Available from: https://gco.iarc.fr/Today.
6. Aprile M, Katopodi V, Leucci E, Costa V. LncRNAs in cancer: from garbage to junk. Cancers. 2020; 12(11):3220. doi:10.3390/cancers12113220
7. Kashi K, Henderson L, Bonetti A, Carninci P. Discovery and functional analysis of lncRNAs: methodologies to investigate an uncharacterized transcriptome. Biochim Biophys Acta. 2015; 1859 doi:10.1016/j.bbagrm.2015.10.010
8. Kung JTY, Colognori D, Lee JT. Long noncoding RNAs: past, present, and future. Genetics. 2013; 193(3):651-69. doi:10.1534/genetics.112.146704
9. Ge M, Li A, Wang M. A bipartite network-based method for prediction of long non-coding RNA–protein interactions. Genomics Proteomics Bioinformatics. 2016; 14(1):62-71. doi:10.1016/j.gpb.2016.01.004
10. Li CH, Chen Y. Insight into the role of long noncoding RNA in cancer development and progression. In: Jeon KW, Galluzzi L, editors. Int rev cell mol bio: Academic Press; 2016. p. 33-65.
11. Jin Y, Liu S, Yuan Z, Yang Y, Tan S, Liu Z. Catfish genomic studies: progress and perspectives. In: MacKenzie S, Jentoft S, editors. Genomics in aquaculture. San Diego: Academic Press; 2016. p. 73-104.
12. Saldanha SN, Tollefsbol TO. Epigenetic approaches to cancer therapy. In: Tollefsbol TO, editor. Epigenetics in human disease. 2 ed: Academic Press; 2018. p. 219-47.
13. Petrov TLD, Riddle NC. The evolution of new technologies and methods in clinical epigenetics research. In: García-Giménez JL, editor. Epigenetic biomarkers and diagnostics. Boston: Academic Press; 2016. p. 67-89.
14. Gray SG. The potential of epigenetic compounds in treating diabetes. In: Tollefsbol TO, editor. Epigenetics in human disease. 2nd ed: Academic Press; 2018. p. 489-547.
15. Hu G, Zhao K. Identification of intergenic long noncoding RNA by deep sequencing. In: Huang S, Litt MD, Blakey CA, editors. Epigenetic gene expression and regulation. Oxford: Academic Press; 2015. p. 223-35.
16. Hosseini ES, Meryet-Figuiere M, Sabzalipoor H, Kashani HH, Nikzad H, Asemi Z. Dysregulated expression of long noncoding RNAs in gynecologic cancers. Mol Cancer. 2017; 16(1):107. doi:10.1186/s12943-017-0671-2
17. Walters K, Sarsenov R, Too WS, Hare RK, Paterson IC, Lambert DW, Brown S, Bradford JR. Comprehensive functional profiling of long non-coding RNAs through a novel pan-cancer integration approach and modular analysis of their protein-coding gene association networks. BMC Genomics. 2019; 20(1):454. doi:10.1186/s12864-019-5850-7
18. Wang Kevin C, Chang Howard Y. Molecular mechanisms of long noncoding RNAs. Mol Cell. 2011; 43(6):904-14. doi:10.1016/j.molcel.2011.08.018

19. Grillone K, Riillo C, Scionti F, Rocca R, Tradigo G, Guzzi PH, Alcaro S, Di Martino MT, Tagliaferri P, Tassone P. Non-coding RNAs in cancer: platforms and strategies for investigating the genomic "dark matter". J Exp Clin Cancer Res. 2020; 39(1):117. doi:10.1186/s13046-020-01622-x
20. Gibb EA, Vucic EA, Enfield KSS, Stewart GL, Lonergan KM, Kennett JY, Becker-Santos DD, MacAUlay CE, Lam S, Brown CJ, Lam WL. Human cancer long non-coding RNA transcriptomes. PLoS One. 2011; 6(10):e25915-e. doi:10.1371/journal.pone.0025915
21. Ma L, Bajic VB, Zhang Z. On the classification of long non-coding RNAs. RNA Biol. 2013; 10(6):925-33. doi:10.4161/rna.24604
22. Tsagakis I, Douka K, Birds I, Aspden J. Long non-coding RNAs in development and disease: conservation to mechanisms: long non-coding RNAs in development and disease. J Pathol. 2020; 250 doi:10.1002/path.5405
23. Moran VA, Perera RJ, Khalil AM. Emerging functional and mechanistic paradigms of mammalian long non-coding RNAs. Nucleic Acids Res. 2012; 40(14):6391-400. doi:10.1093/nar/gks296
24. Merriel SWD, Funston G, Hamilton W. Prostate cancer in primary care. Adv Ther. 2018; 35(9):1285-94. doi:10.1007/s12325-018-0766-1
25. Seraphin T, Joko Fru YW, Kamaté B, Chokunonga E, Wabinga H, Somdyala N, Manraj S, Ogunbiyi O, Dzamalala C, Finesse A, Korir A, N'da Guy, Lorenzoni C, Liu B, Kantelhardt E, Parkin D. Rising prostate cancer incidence in Sub-Saharan Africa: A trend analysis of data from the African cancer registry network. Cancer Epidemiol Biomarkers Prev. 2020; 30 doi:10.1158/1055-9965.EPI-20-1005

26. Acha-Sagredo A, Uko B, Pantazi P, Bediaga NG, Moschandrea C, Rainbow L, Marcus MW, Davies MPA, Field JK, Liloglou T. Long non-coding RNA dysregulation is a frequent event in non-small cell lung carcinoma pathogenesis. Br J Cancer. 2020; 122(7):1050-8. doi:10.1038/s41416-020-0742-9
27. Prensner J, Zhao S, Erho N, Schipper M, Iyer M, Dhanasekaran S, Magi-Galluzzi C, Mehra R, Sahu A, Siddiqui J, Davicioni E, Den R, Dicker A, Jeff R, Karnes R, Wei J, Klein E, Jenkins R, Chinnaiyan A, Feng F. RNA biomarkers associated with metastatic progression in prostate cancer: a multi-institutional high-throughput analysis of SChLAP1. Lancet Oncol. 2014; 15 doi:10.1016/S1470-2045(14)71113-1
28. Tai S, Sun Y, Squires JM, Zhang H, Oh WK, Liang CZ, Huang J. PC3 is a cell line characteristic of prostatic small cell carcinoma. The Prostate. 2011; 71(15):1668-79. doi:10.1002/pros.21383
29. Misawa A, Takayama K-I, Inoue S. Long non-coding RNAs and prostate cancer. Cancer Sci. 2017; 108(11):2107-14. doi:10.1111/cas.13352
30. Marima R, Hull R, Mbeje M, Molefi T, Mathabe K, Elbagory AM, Demitriou D, DLamini Z. Role of precision oncology in type II endometrial and prostate cancers in the African population: global cancer genomics disparities. Int J Mol Sci. 2022; 23(2):628. doi:10.3390/ijms23020628
31. Dunn MW, Kazer MW. Prostate cancer overview. Semin Oncol Nurs. 2011; 27(4):241-50. doi:10.1016/j.soncn.2011.07.002
32. Turner B, Drudge-Coates L. Prostate cancer: risk factors, diagnosis and management. Cancer Nurs Pract. 2010; 9:29-36. doi:10.7748/cnp2010.12.9.10.29.c8126
33. Sauer AK, Vela H, Vela G, Stark P, Barrera-Juarez E, Grabrucker AM. Zinc deficiency in men over 50 and its implications in prostate disorders. Front Oncol. 2020; 10:1293. doi:10.3389/fonc.2020.01293
34. Martin RM, Vatten L, Gunnell D, Romundstad P, Nilsen TI. Lower urinary tract symptoms and risk of prostate cancer: the hunt 2 cohort, Norway. Int J Cancer. 2008; 123(8):1924-8. doi:10.1002/ijc.23713
35. Kearns JT, Holt SK, Wright JL, Lin DW, Lange PH, Gore JL. PSA screening, prostate biopsy, and treatment of prostate cancer in the years surrounding the USPSTF recommendation against prostate cancer screening. Cancer. 2018; 124(13):2733-9. doi:10.1002/cncr.31337

36. Webber MM, Waghray A, Bello D. Prostate-specific antigen, a serine protease, facilitates human prostate cancer cell invasion. Clin Cancer Res. 1995; 1(10):1089.
37. Leslie SW, Soon-Sutton TL, R IA, Sajjad H, Siref LE. Prostate cancer. Statpearls. Treasure Island (FL): StatPearls Publishing Copyright © 2022, StatPearls Publishing LLC.; 2022.

38. Marima R, Mbeje M, Hull R, Demetriou D, Mtshali N, Dlamini Z. Prostate cancer disparities and management in Southern Africa: insights into practices, norms and values. Cancer Manag Res. 2022; 14:3567-79. doi:10.2147/cmar.S382903
39. NICE guidance - prostate cancer: diagnosis and management. BJU Int. 2019; 124(1):9-26. doi:10.1111/bju.14809
40. Streicher J, Meyerson BL, Karivedu V, Sidana A. A review of optimal prostate biopsy: indications and techniques. Ther Adv Urol. 2019; 11:1756287219870074-. doi:10.1177/1756287219870074
41. Ebell MH, Thai TN, Royalty KJ. Cancer screening recommendations: an international comparison of high income countries. Public Health Rev. 2018; 39:7. doi:10.1186/s40985-018-0080-0
42. Zhou CK, Check DP, Lortet-Tieulent J, Laversanne M, Jemal A, Ferlay J, Bray F, Cook MB, Devesa SS. Prostate cancer incidence in 43 populations worldwide: an analysis of time trends overall and by age group. Int J Cancer. 2016; 138(6):1388-400. doi:10.1002/ijc.29894
43. Cassell A, Yunusa B, Jalloh M, Ndoye M, Mbodji MM, Diallo A, Kouka SC, Labou I, Niang L, Gueye SM. Management of advanced and metastatic prostate cancer: a need for a sub-Saharan guideline. J Oncol. 2019; 2019:1785428-. doi:10.1155/2019/1785428
44. Ngcamphalala C, Östensson E, Ginindza TG. The economic burden of prostate cancer in Eswatini. BMC Health Serv Res. 2022; 22(1):483. doi:10.1186/s12913-022-07817-6
45. Cheng L, Montironi R, Bostwick DG, Lopez-Beltran A, Berney DM. Staging of prostate cancer. Histopathology. 2012; 60(1):87-117. doi:10.1111/j.1365-2559.2011.04025.x
46. Sharma SV, Haber DA, Settleman J. Cell line-based platforms to evaluate the therapeutic efficacy of candidate anticancer agents. Nat Rev Cancer. 2010; 10(4):241-53. doi:10.1038/nrc2820
47. Wu X, Gong S, Roy-Burman P, Lee P, Culig Z. Current mouse and cell models in prostate cancer research. Endocr Relat Cancer. 2013; 20(4):R155-R70. doi:10.1530/ERC-12-0285
48. Castanares MA, Copeland BT, Chowdhury WH, Liu MM, Rodriguez R, Pomper MG, Lupold SE, Foss CA. Characterization of a novel metastatic prostate cancer cell line of LNCAP origin. Prostate. 2016; 76(2):215-25. doi:10.1002/pros.23115
49. Alimirah F, Chen J, Basrawala Z, Xin H, Choubey D. Du-145 and PC-3 human prostate cancer cell lines express androgen receptor: Implications for the androgen receptor functions and regulation. FEBS Lett. 2006; 580(9):2294-300. doi:10.1016/j.febslet.2006.03.041
50. Heinlein CA, Chang C. Androgen receptor in prostate cancer. Endocr Rev. 2004; 25(2):276-308. doi:10.1210/er.2002-0032
51. Ferraldeschi R, Welti J, Luo J, Attard G, de Bono JS. Targeting the androgen receptor pathway in castration-resistant prostate cancer: Progresses and prospects. Oncogene. 2015; 34(14):1745-57. doi:10.1038/onc.2014.115
52. Jaratlerdsiri W, Jiang J, Gong T, Patrick SM, Willet C, Chew T, Lyons RJ, Haynes A, Pasqualim G, Louw M, Kench JG, Campell R, Horvath LG, Chan EKF, Wedge DC, Sadsad R, Brum IS, Mutambirwa SBA, Stricker PD, Bornman MSR, Hayes VM. African-specific molecular taxonomy of prostate cancer. Nature. 2022; 609(7927):552-9. doi:10.1038/s41586-022-05154-6
53. Gong T, Jaratlerdsiri W, Jiang J, Willet C, Chew T, Patrick SM, Lyons RJ, Haynes A, Pasqualim G, Louw M, Kench JG, Campell R, Horvath LG, Chan EKF, Wedge DC, Sadsad R, Brum IS, Mutambirwa SBA, Stricker PD, Bornman MSR, Hayes VM. Genome-wide interrogation of structural variation reveals novel African-specific prostate cancer oncogenic drivers. Genome Med. 2022; 14(1):100. doi:10.1186/s13073-022-01096-w

54. Kiciński M, Vangronsveld J, Nawrot TS. An epidemiological reappraisal of the familial aggregation of prostate cancer: A meta-analysis. PLoS One. 2011; 6(10):e27130-e. doi:10.1371/journal.pone.0027130
55. Alvarez-Cubero MJ, Pascual-Geler M, Martinez-Gonzalez LJ, Expósito Ruiz M, Saiz M, Cozar JM, Lorente JA. Association between RNASEL, MSR1, and ELAC2 single nucleotide polymorphisms and gene expression in prostate cancer risk. Urol Oncol. 2016; 34(10):431.e1-.e8. doi:10.1016/j.urolonc.2016.05.018
56. Dayal S, Zhou J, Manivannan P, Siddiqui MA, Ahmad OF, Clark M, Awadia S, Garcia-Mata R, Shemshedini L, Malathi K. RNase L suppresses androgen receptor signalling, cell migration and matrix metalloproteinase activity in prostate cancer cells. Int J Mol Sci. 2017; 18(3) doi:10.3390/ijms18030529
57. Prensner JR, Chinnaiyan AM. The emergence of lncRNAs in cancer biology. Cancer Discov. 2011; 1(5):391-407. doi:10.1158/2159-8290.Cd-11-0209
58. Takayama K-i, Fujimura T, Suzuki Y, Inoue S. Identification of long non-coding RNAs in advanced prostate cancer associated with androgen receptor splicing factors. Commun Biol. 2020; 3(1):393. doi:10.1038/s42003-020-01120-y
59. de Koning APJ, Gu W, Castoe TA, Batzer MA, Pollock DD. Repetitive elements may comprise over two-thirds of the human genome. PLoS Genet. 2011; 7(12):e1002384-e. doi:10.1371/journal.pgen.1002384
60. Yi H-C, You Z-H, Huang D-S, Li X, Jiang T-H, Li L-P. A deep learning framework for robust and accurate prediction of ncRNA-protein interactions using evolutionary information. Mol Ther Nucleic Acids. 2018; 11:337-44. doi:10.1016/j.omtn.2018.03.001
61. Djebali S, Davis CA, Merkel A, Dobin A, Lassmann T, Mortazavi A, Tanzer A, Lagarde J, Lin W, Schlesinger F, Xue C, Marinov GK, Khatun J, Williams BA, Zaleski C. Landscape of transcription in human cells. Nature. 2012; 489(7414):101-8. doi:10.1038/nature11233
62. Esteller M. Non-coding RNAs in human disease. Nat Rev Genet. 2011; 12(12):861-74. doi:10.1038/nrg3074
63. Ling H, Vincent K, Pichler M, Fodde R, Berindan-Neagoe I, Slack FJ, Calin GA. Junk DNA and the long non-coding RNA twist in cancer genetics. Oncogene. 2015; 34(39):5003-11. doi:10.1038/onc.2014.456
64. Guttman M, Amit I, Garber M, French C, Lin MF, Feldser D, Huarte M, Zuk O, Carey BW, Cassady JP, Cabili MN, Jaenisch R, Mikkelsen TS, Jacks T. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. Nature. 2009; 458(7235):223-7. doi:10.1038/nature07672
65. Kumar M, DeVaux RS, Herschkowitz JI. Molecular and cellular changes in breast cancer and new roles of lncRNAs in breast cancer initiation and progression. In: Pruitt K, editor. Prog mol biol transl sci: Academic Press; 2016. p. 563-86.
66. Gu P, Chen X, Xie R, Xie W, Huang L, Dong W, Han J, Liu X, Shen J, Huang J, Lin T. A novel AR translational regulator lncRNA LBCS inhibits castration resistance of prostate cancer. Mol Cancer. 2019; 18(1):109. doi:10.1186/s12943-019-1037-8
67. Hu Q, Li Y, Li D, Yuan Y, Wang K, Yao L, Cheng Z, Han T. Amino acid metabolism regulated by lncRNAs: the propellant behind cancer metabolic reprogramming. Cell Commun Signal. 2023; 21(1):87. doi:10.1186/s12964-023-01116-1
68. Chen R, Liu Y, Zhuang H, Yang B, Hei K, Xiao M, Hou C, Gao H, Zhang X, Jia C, Li L, Li Y, Zhang N. Quantitative proteomics reveals that long non-coding RNA MALAT1 interacts with DBC1 to regulate p53 acetylation. Nucleic Acids Res. 2017; 45(17):9947-59. doi:10.1093/nar/gkx600
69. Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgner H, Guernec G, Martin D, Merkel A, Knowles D, Laragarde J, Veeravali L, Ruan X, Ruan Y. The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. Genome Res. 2012; 22:1775-89. doi:10.1101/gr.132159.111
70. Zhu J, Fu H, Wu Y, Zheng X. Function of lncRNAs and approaches to lncRNA-protein interactions. Sci China Life Sci. 2013; 56(10):876-85. doi:10.1007/s11427-013-4553-6
71. Liao Q, Liu C, Yuan X, Kang S, Miao R, Xiao H, Zhao G, Luo H, Bu D, Zhao H, Skogerbø G, Wu Z, Zhao Y. Large-scale prediction of long non-coding RNA functions in a

coding–non-coding gene co-expression network. Nucleic Acids Res. 2011; 39(9):3864-78. doi:10.1093/nar/gkq1348
72. Khalil AM, Rinn JL. RNA–protein interactions in human health and disease. Semin Cell Dev Biol. 2011; 22(4):359-65. doi:10.1016/j.semcdb.2011.02.016
73. Morillon A. Biological functions of long non-coding RNA. In: Morillon A, editor. Long non-coding rna: Elsevier; 2018. p. 55-67.
74. Thomson DW, Dinger ME. Endogenous microRNA sponges: evidence and controversy. Nat Rev Genet. 2016; 17(5):272-83. doi:10.1038/nrg.2016.20
75. Balas MM, Johnson AM. Exploring the mechanisms behind long noncoding RNAs and cancer. Noncoding RNA Res. 2018; 3(3):108-17. doi:10.1016/j.ncrna.2018.03.001
76. Fang Y, Fullwood MJ. Roles, functions, and mechanisms of long non-coding RNAs in cancer. Genomics, Proteomics & Bioinformatics. 2016; 14(1):42-54. doi:10.1016/j.gpb.2015.09.006
77. Jiang G, Su Z, Liang X, Huang Y, Lan Z, Jiang X. Long non-coding RNAs in prostate tumorigenesis and therapy (review). Mol Clin Oncol. 2020; 13(6):76. doi:10.3892/mco.2020.2146
78. Bussemakers M, Van Bokhoven A, Verhaegh GW, Smit FP, Karthaus H, Schalken J, Debruyne F, Ru N, Isaacs WB. DD3: a new prostate-specific gene, highly overexpressed in prostate cancer. Cancer Res. 2000; 59:5975-9.
79. Schalken JA, Hessels D, Verhaegh G. New targets for therapy in prostate cancer: sifferential display code 3 (DD3(PCA3)), a highly prostate cancer-specific gene. Urology. 2003; 62(5 Suppl 1):34-43. doi:10.1016/s0090-4295(03)00759-3
80. Wei JT, Feng Z, Partin AW, Brown E, Thompson I, Sokoll L, Chan DW, Lotan Y, Kibel AS, Busby JE, Bidair M, Lin DW, Taneja SS, Viterbo, R, Joon AY, Dahlgren J, Kagan J, Srivastava S, Sanda MG. Can urinary PCA3 supplement PSA in the early detection of prostate cancer? J Clin Oncol. 2014; 32(36):4066-72. doi:10.1200/JCO.2013.52.8505
81. Truong M, Yang B, Jarrard DF. Toward the detection of prostate cancer in urine: a critical analysis. J Urol. 2013; 189(2):422-9. doi:10.1016/j.juro.2012.04.143
82. Lemos AE, Ferreira LB, Batoreu NM, de Freitas PP, Bonamino MH, Gimba ER. PCA3 long noncoding RNA modulates the expression of key cancer-related genes in LNCaP prostate cancer cells. Tumour Biol. 2016; 37(8):11339-48. doi:10.1007/s13277-016-5012-3
83. Lemos AEG, Matos AdR, Ferreira LB, Gimba ERP. The long non-coding RNA PCA3: an update of its functions and clinical applications as a biomarker in prostate cancer. Oncotarget. 2019; 10(61)
84. Chen CD, Welsbie DS, Tran C, Baek SH, Chen R, Vessella R, Rosenfeld MG, Sawyers CL. Molecular determinants of resistance to antiandrogen therapy. Nat Med. 2004; 10(1):33-9. doi:10.1038/nm972
85. Chakravarty D, Sboner A, Nair SS, Giannopoulou E, Li R, Hennig S, Mosquera JM, Pauwels J, Park K, Kossai M, MacDOnald TY, Fontugne J, Erho N, Vergara IA. The oestrogen receptor alpha-regulated lncRNA NEAT1 is a critical modulator of prostate cancer. Nat Commun. 2014; 5(1):5383. doi:10.1038/ncomms6383
86. Zhang A, Zhao JC, Kim J, Fong KW, Yang YA, Chakravarti D, Mo YY, Yu J. LncRNA HOTAIR enhances the androgen-receptor-mediated transcriptional program and drives castration-resistant prostate cancer. Cell Rep. 2015; 13(1):209-21. doi:10.1016/j.celrep.2015.08.069
87. Crawford ED, Heidenreich A, Lawrentschuk N, Tombal B, Pompeo ACL, Mendoza-Valdes A, Miller K, Debruyne FMJ, Koltz L. Androgen-targeted therapy in men with prostate cancer: Evolving practice and future considerations. Prostate Cancer Prostatic Dis. 2019; 22(1):24-38. doi:10.1038/s41391-018-0079-0
88. Yuan X, Cai C, Chen S, Chen S, Yu Z, Balk SP. Androgen receptor functions in castration-resistant prostate cancer and mechanisms of resistance to new agents targeting the androgen axis. Oncogene. 2014; 33(22):2815-25. doi:10.1038/onc.2013.235
89. Takayama K-I, Horie-Inoue K, Katayama S, Suzuki T, Tsutsumi S, Ikeda K, Urano T, Fujimara T, Takagi K, Takahashi S, Hooma Y. Androgen-responsive long noncoding RNA

CTBP1-AS promotes prostate cancer. EMBO J. 2013; 32(12):1665-80. doi:10.1038/emboj.2013.99
90. Mehra R, Udager AM, Ahearn TU, Cao X, Feng FY, Loda M, Metimar JS, Kantoff P, Mucci LA, Chinnaiyan AM. Overexpression of the long non-coding RNA SChLAP1 independently predicts lethal prostate cancer. Eur Urol. 2016; 70(4):549-52. doi:10.1016/j.eururo.2015.12.003
91. Cao Y, Bryan TM, Reddel RR. Increased copy number of the TERT and TERC telomerase subunit genes in cancer cells. Cancer Sci. 2008; 99(6):1092-9. doi:10.1111/j.1349-7006.2008.00815.x
92. Collins K. Physiological assembly and activity of human telomerase complexes. Mech Ageing Dev. 2008; 129(1-2):91-8. doi:10.1016/j.mad.2007.10.008
93. Kim NW, Piatyszek MA, Prowse KR, Harley CB, West MD, Ho PL, Coviello GM, Wright, WE, Weinrich SL, Shay JW. Specific association of human telomerase activity with immortal cells and cancer. Science. 1994; 266(5193):2011-5. doi:10.1126/science.7605428
94. Hahn WC, Stewart SA, Brooks MW, York SG, Eaton E, Kurachi A, Beijersbergen RL, Knoll JH, Meyerson M, Weinberg RA. Inhibition of telomerase limits the growth of human cancer cells. Nat Med. 1999; 5(10):1164-70. doi:10.1038/13495
95. Bodnar AG, Ouellette M, Frolkis M, Holt SE, Chiu CP, Morin GB, Harley CB, Shay JW, Lichtsteiner S, Wright WE. Extension of life-span by introduction of telomerase into normal human cells. Science. 1998; 279(5349):349-52. doi:10.1126/science.279.5349.349
96. Wang J, Wu M, Chang L, Jin Z, Yang X, Li D, Wang J, Qu J, Hou Q, Huang X, Xu C. The lncRNA TERC promotes gastric cancer cell proliferation, migration, and invasion by sponging mir-423-5p to regulate sox12 expression. Ann Transl Med. 2022; 10(18):963.
97. Baena-Del Valle JA, Zheng Q, Esopi DM, Rubenstein M, Hubbard GK, Moncaliano MC, Hruszkewycz A, Vaghasia A, Yegnasubramanian S, Wheelan SJ, Meeker AK, Heaphy CM, Graham MK, De Marzo AM. MYC drives overexpression of telomerase RNA (hTR/TERC) in prostate cancer. J Pathol. 2018; 244(1):11-24. doi:10.1002/path.4980
98. Sakurai K, Reon BJ, Anaya J, Dutta A. The lncRNA DRAIC/PCAT29 locus constitutes a tumor-suppressive nexus. Mol Cancer Res. 2015; 13(5):828. doi:10.1158/1541-7786.MCR-15-0016-T
99. Zhang Y, Su X, Kong Z, Fu F, Zhang P, Wang D, Wu H, Wan X, Li Y. An androgen reduced transcript of lncRNA GAS5 promoted prostate cancer proliferation. PloS One. 2017; 12(8):e0182305-e. doi:10.1371/journal.pone.0182305
100. Pickard MR, Mourtada-Maarabouni M, Williams GT. Long non-coding RNA GAS5 regulates apoptosis in prostate cancer cell lines. Biochim Biophys. 2013; 1832(10):1613-23. doi:10.1016/j.bbadis.2013.05.005
101. Wu M, Huang Y, Chen T, Wang W, Yang S, Ye Z, Xi X. LncRNA MEG3 inhibits the progression of prostate cancer by modulating mir-9-5p/QKI-5 axis. J Cell Mol Med. 2019; 23(1):29-38. doi:10.1111/jcmm.13658
102. Prensner JR, Chen W, Han S, Iyer MK, Cao Q, Kothari V, Evans JR, Knudsen KE, Paulsen MT, Ljungman ML, Theodore S, Chinnaiyan AM, Feng FY. The long non-coding RNA PCAT-1 promotes prostate cancer cell proliferation through cMyc. Neoplasia. 2014; 16(11):900-8. doi:10.1016/j.neo.2014.09.001
103. Ylipää A, Kivinummi K, Kohvakka A, Annala M, Latonen L, Scaravilli M, Kartasalo K, Leppänen S, Karakurt S, Seppälä J, Yli-Harja O, Tammela TLJ. Transcriptome sequencing reveals PCAT5 as a novel ERG-regulated long noncoding RNA in prostate cancer. Cancer Res. 2015; 75(19):4026. doi:10.1158/0008-5472.CAN-15-0217
104. Yang L, Lin C, Jin C, Yang JC, Tanasa B, Li W, Merkurjev D, Ohgi KA, Meng D, Zhang J, Evans CP, Rosenfeld MG. LncRNA-dependent mechanisms of androgen-receptor-regulated gene activation programs. Nature. 2013; 500(7464):598-602. doi:10.1038/nature12451
105. Li J, Zhang Z, Xiong L, Guo C, Jiang T, Zeng L, Li G, Wang J. SNHG1 lncRNA negatively regulates mir-199a-3p to enhance CDK7 expression and promote cell proliferation in prostate cancer. Biochem Biophys Res Commun. 2017; 487(1):146-52. doi:10.1016/j.bbrc.2017.03.169

106. Misawa A, Takayama K-i, Urano T, Inoue S. Androgen-induced long noncoding RNA (lncrna) socs2-as1 promotes cell growth and inhibits apoptosis in prostate cancer cells J Biol Chem. 2016; 291(34):17861-80. doi:10.1074/jbc.M116.718536
107. Fonseca J, Moradi F, Maddalena LA, Ferreira-Tollstadius B, Selim S, Stuart JA. Resveratrol integrates metabolic and growth effects in PC3 prostate cancer cells-involvement of prolyl hydroxylase and hypoxia inducible factor-1. Oncol Lett. 2019; 17(1):697-705. doi:10.3892/ol.2018.9526
108. Tsoi LC, Iyer MK, Stuart PE, Swindell WR, Gudjonsson JE, Tejasvi T, Sarkar MK, Li B, Ding J, Woorhees JJ, Kang HM, Nair RP, Chinnaiyan AM, Abecasis GR, Elder JT. Analysis of long non-coding RNAs highlights tissue-specific expression patterns and epigenetic profiles in normal and psoriatic skin. Genome Biol. 2015; 16(1):24-. doi:10.1186/s13059-014-0570-4
109. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-delta delta c(t)) method. Methods. 2001; 25(4):402-8. doi:10.1006/meth.2001.1262
110. Bechtel W. Using computational models to discover and understand mechanisms. Stud Hist Philos Sci. 2016; 56:113-21. doi:10.1016/j.shpsa.2015.10.004
111. Chen J, Zhang J, Gao Y, Li Y, Feng C, Song C, Ning Z, Zhou X, Zhao J, Feng M, Zhang Y, Wei L, Pan Q, Jiang Y, Qian F, Han J, Yang Y, Wang Q, Li C. LncSEA: a platform for long non-coding RNA related sets and enrichment analysis. Nucleic Acids Res. 2021; 49(D1):D969-d80. doi:10.1093/nar/gkaa806

112. Li JH, Liu S, Zhou H, Qu LH, Yang JH. StarBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. Nucleic Acids Res. 2014; 42(Database issue):D92-7. doi:10.1093/nar/gkt1248
113. Bao Z, Yang Z, Huang Z, Zhou Y, Cui Q, Dong D. LncRNADisease 2.0: an updated database of long non-coding RNA-associated diseases. Nucleic Acids Res. 2019; 47(D1):D1034-d7. doi:10.1093/nar/gky905
114. Paraskevopoulou MD, Vlachos IS, Karagkouni D, Georgakilas G, Kanellos I, Vergoulis T, Zagganas K, Tsanakas P, Floros E, Dalamagas T, Hatzigeorgiou AG. DIANA-LncBase v2: indexing microRNA targets on non-coding transcripts. Nucleic Acids Res. 2016; 44(D1):D231-8. doi:10.1093/nar/gkv1270
115. Cai X, Cullen BR. The imprinted H19 noncoding RNA is a primary microRNA precursor. RNA. 2007; 13(3):313-6. doi:10.1261/rna.351707
116. Vlachos IS, Zagganas K, Paraskevopoulou MD, Georgakilas G, Karagkouni D, Vergoulis T, Dalamagas T, Hatzigeorgiou AG. DIANA-miRPath v3.0: deciphering microRNA function with experimental support. Nucleic Acids Res. 2015; 43(W1):W460-6. doi:10.1093/nar/gkv403
117. Karagkouni D, Paraskevopoulou MD, Chatzopoulos S, Vlachos IS, Tastsoglou S, Kanellos I, Papadimitriou D, Kavakiotis I, Maniou S, Skoufos G, Vergoulis T, Dalamagas T, Hatzigeorgiou AG. DIANA-TarBase v8: a decade-long collection of experimentally supported miRNA-gene interactions. Nucleic Acids Res. 2018; 46(D1):D239-d45. doi:10.1093/nar/gkx1141

118. Lieb V, Weigelt K, Scheinost L, Fischer K, Greither T, Marcou M, Theil G, Klocker H Holzhausen HJ, Lai X, Vera J, Ekici AB, Horninger W, Fornara P, Wullich B, Taubert H, Wach S. Serum levels of miR-320 family members are associated with clinical parameters and diagnosis in prostate cancer patients. Oncotarget. 2018; 9(12):10402-16. doi:10.18632/oncotarget.23781

119. McGeary SE, Lin KS, Shi CY, Pham TM, Bisaria N, Kelley GM, Bartel DP. The biochemical basis of microRNA targeting efficacy. Science. 2019; 366(6472) doi:10.1126/science.aav1741
120. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003; 13(11):2498-504. doi:10.1101/gr.1239303

121. Ponchel F, Toomes C, Bransfield K, Leong FT, Douglas SH, Field SL, Bell SM, Combaret V, Puisieux A, Mighell AJ, Robinson PA, Inglehearn CF, Isaacs JD, Markham AF. Real-time PCR based on SYBR-Green I fluorescence: an alternative to the TaqMan assay for a relative quantification of gene rearrangements, gene amplifications and micro gene deletions. BMC Biotechnol. 2003; 3(1):18. doi:10.1186/1472-6750-3-18
122. Freitas FCP, Depintor TS, Agostini LT, Luna-Lucena D, Nunes FMF, Bitondi MMG, Simões ZLP, Lourenço AP. Evaluation of reference genes for gene expression analysis by real-time quantitative PCR (qPCR) in three stingless bee species (Hymenoptera: Apidae: Meliponini). Sci Rep. 2019; 9(1):17692. doi:10.1038/s41598-019-53544-0

123. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG. Primer3-new capabilities and interfaces. Nucleic Acids Res. 2012; 40(15):e115-e. doi:10.1093/nar/gks596
124. Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden TL. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. BMC Bioinformatics. 2012; 13(1):134. doi:10.1186/1471-2105-13-134
125. Trypsteen W, Mohammadi P, Van Hecke C, Mestdagh P, Lefever S, Saeys Y, De Bleser P. Vandesompele J, Ciuffi A, Vandekerckhove L, De Spiegelaere W. Differential expression of lncRNAs during the HIV replication cycle: an underestimated layer in the HIV-host interplay. Sci Rep. 2016; 6(1):36111. doi:10.1038/srep36111
126. Eitan E, Tamar A, Yossi G, Peleg R, Braiman A, Priel E. Expression of functional alternative telomerase RNA component gene in mouse brain and in motor neurons cells protects from oxidative stress. Oncotarget. 2016; 7(48):78297-309. doi:10.18632/oncotarget.13049
127. Xin X, Li Q, Fang J, Zhao T. LncRNA HOTAIR: a potential prognostic factor and therapeutic target in human cancers. Front Oncol. 2021; 11 doi:10.3389/fonc.2021.679244
128. Nanashima N, Horie K, Yamada T, Shimizu T, Tsuchida S. Hair keratin KRT81 is expressed in normal and breast cancer cells and contributes to their invasiveness. Oncol Rep. 2017; 37(5):2964-70. doi:10.3892/or.2017.5564
129. Wang Y, Zhang M, Zhou F. Biological functions and clinical applications of exosomal long non-coding RNAs in cancer. J Cell Mol Med. 2020; 24(20):11656-66. doi:10.1111/jcmm.15873
130. Kundu ST, Grzeskowiak CL, Fradette JJ, Gibson LA, Rodriguez LB, Creighton CJ, Scott KL, Gibbons DL. TMEM106B drives lung cancer metastasis by inducing TFEB-dependent lysosome synthesis and secretion of cathepsins. Nat Commun. 2018; 9(1):2731. doi:10.1038/s41467-018-05013-x
131. Zhang J, Zhao X, Luan Z, Wang A. *RAB14* overexpression promotes proliferation and invasion through yap signaling in non-small cell lung cancers. Onco Targets Ther. 2020; 13:9269-80. doi:10.2147/ott.S255644
132. Zhang K, Liang Y, Zhang W, Zeng N, Tang S, Tian R. KRT81 knockdown inhibits malignant progression of melanoma through regulating interleukin-8. DNA Cell Biol. 2021; 40(10):1290-7. doi:10.1089/dna.2021.0317
133. Da M, Zhuang J, Zhou Y, Qi Q, Han S. Role of long noncoding RNA taurine-upregulated gene 1 in cancers. Mol Med. 2021; 27(1):51. doi:10.1186/s10020-021-00312-4
134. Ding Y-M, Chan EC, Liu L-C, Liu Z-W, Wang Q, Wang J-L, Cui XP, Jiang F, Guo XS. Long noncoding RNAs: important participants and potential therapeutic targets for myocardial ischaemia reperfusion injury. Clin Exp Pharmacol Physiol. 2020; 47(11):1783-90. doi:10.1111/1440-1681.13375
135. Li H, Wang X, Wen C, Huo Z, Wang W, Zhan Q, Cheng D, Chen H, Deng X, Peng C, Shen B. Long noncoding RNA NORAD, a novel competing endogenous rRNAna, enhances the hypoxia-induced epithelial-mesenchymal transition to promote metastasis in pancreatic cancer. Mol Cancer. 2017; 16(1):169. doi:10.1186/s12943-017-0738-0
136. Yu SY, Peng H, Zhu Q, Wu YX, Wu F, Han CR, Yan B, Li Q, Xiang HG. Silencing the long noncoding RNA NORAD inhibits gastric cancer cell proliferation and invasion by the

RhoA/ROCK1 pathway. Eur Rev Med Pharmacol Sci. 2019; 23(9):3760-70. doi:10.26355/eurrev_201905_17802
137. Shan Q, Qu F, Yang W, Chen N. Effect of LINC00657 on apoptosis of breast cancer cells by regulating mir-590-3p. Cancer Manag Res. 2020; 12:4561-71. doi:10.2147/cmar.S249576
138. Chu L, Yu L, Liu J, Song S, Yang H, Han F, Liu F, Hu Y. Long intergenic non-coding LINC00657 regulates tumorigenesis of glioblastoma by acting as a molecular sponge of miR-190a-3p. Aging (Albany NY). 2019; 11(5):1456-70. doi:10.18632/aging.101845
139. Huang W, Li H, Yu Q, Xiao W, Wang DO. LncRNA-mediated DNA methylation: an emerging mechanism in cancer and beyond. Journal of Experimental & Clinical Cancer Research. 2022; 41(1):100. doi:10.1186/s13046-022-02319-z
140. Ehrlich M. DNA hypermethylation in disease: Mechanisms and clinical relevance. Epigenetics. 2019; 14(12):1141-63. doi:10.1080/15592294.2019.1638701
141. Song Y, Wang R, Li LW, Liu X, Wang YF, Wang QX, Zhang Q. Long non-coding RNA HOTAIR mediates the switching of histone H3 lysine 27 acetylation to methylation to promote epithelial-to-mesenchymal transition in gastric cancer. Int J Oncol. 2019; 54(1):77-86. doi:10.3892/ijo.2018.4625
142. O'Brien J, Hayder H, Zayed Y, Peng C. Overview of microRNA biogenesis, mechanisms of actions, and circulation. Front Endocrinol. 2018; 9 doi:10.3389/fendo.2018.00402
143. Zhang XM, Wang J, Liu ZL, Liu H, Cheng YF, Wang T. LINC00657/miR-26a-5p/CKS2 ceRNA network promotes the growth of esophageal cancer cells via the MDM2/p53/Bcl2/Bax pathway. Biosci Rep. 2020; 40(6) doi:10.1042/bsr20200525
144. Sun B, Liu C, Li H, Zhang L, Luo G, Liang S, Lü M. Research progress on the interactions between long non-coding RNAs and microRNAs in human cancer (review). Oncol Lett. 2020; 19(1):595-605. doi:10.3892/ol.2019.11182
145. Xiong ZJ, Zhang Q, Wang DX, Hu L. Overexpression of TUG1 promotes neuronal death after cerebral infarction by regulating microRNA-9. Eur Rev Med Pharmacol Sci. 2018; 22(21):7393-400. doi:10.26355/eurrev_201811_16278
146. Dong Z, Zhang A, Liu S, Lu F, Guo Y, Zhang G. Aberrant methylation-mediated silencing of lncRNA MEG3 functions as a cerna in esophageal cancer. Mol Cancer Res. 2017; 15(7):800-10. doi:10.1158/1541-7786.Mcr-16-0385
147. Ma M-z, Li C-x, Zhang Y, Weng M-z, Zhang M-d, Qin Y-y, Gong W, Quan Z-w. Long non-coding RNA HOTAIR, a c-Myc activated driver of malignancy, negatively regulates miRNA-130a in gallbladder cancer. Molecular Cancer. 2014; 13(1):156. doi:10.1186/1476-4598-13-156
148. Wang S, Wang Y, Wang S, Tong H, Tang Z, Wang J. Long non-coding RNA FIRRE acts as a miR-520a-3p sponge to promote gallbladder cancer progression via mediating YOD1 expression. Frontiers in Genetics. 2021; 12 doi:10.3389/fgene.2021.674653
149. Li J, Zhang Z, Xiong L, Guo C, Jiang T, Zeng L, Li G, Wang J. SNHG1 lncRNA negatively regulates miR-199a-3p to enhance CDK7 expression and promote cell proliferation in prostate cancer. Biochem Biophys Res Commun. 2017; 487(1):146-52. doi:https://doi.org/10.1016/j.bbrc.2017.03.169
150. Cao W, Sun Y, Liu L, Yu J, Ji J, Wang Y, Yang J. HOTAIR mediates cisplatin resistance in nasopharyngeal carcinoma by regulating miR-106a-5p/SOX4 axis. Bioengineered. 2022; 13(3):6567-78. doi:10.1080/21655979.2022.2038429
151. Lu J, Mu X, Yin Q, Hu K. MiR-106a contributes to prostate carcinoma progression through PTEN. Oncol Lett. 2019; 17(1):1327-32. doi:10.3892/ol.2018.9697
152. Poliseno L, Salmena L, Riccardi L, Fornari A, Song MS, Hobbs RM, Sportoletti P, Varmeh S, Egia A, Fedele G, Rameh L, Loda M, Pandolfi PP. Identification of the miR-106b~25 microRNA cluster as a proto-oncogenic pten-targeting intron that cooperates with its host gene MCM7 in transformation. Sci Signal. 2010; 3(117):ra29. doi:10.1126/scisignal.2000594
153. Drobna M, Szarzyńska B, Jaksik R, Sędek Ł, Kuchmiy A, Taghon T, Van Vlierberghe P, Szczepanski T, Witt M, Dawidowska M. Hsa-miR-20b-5p and hsa-miR-363-3p affect

expression of *PTEN* and *BIM* tumor suppressor genes and modulate survival of t-all cells in vitro. Cells. 2020; 9(5) doi:10.3390/cells9051137
154. Sagar SK. MiR-106b as an emerging therapeutic target in cancer. Genes Dis. 2022; 9(4):889-99. doi:10.1016/j.gendis.2021.02.002
155. Gorur A, Balci Fidanci S, Dogruer UN, Ayaz L, Akbayir S, Yildirim YH, Dirlik M, Serin MS, Tamer L. Determination of plasma microRNA for early detection of gastric cancer. Mol Biol Rep. 2013; 40(3):2091-6. doi:10.1007/s11033-012-2267-7
156. Liang Y, Li S, Tang L. MicroRNA 320, an anti-oncogene target miRNA for cancer therapy. Biomedicines. 2021; 9(6) doi:10.3390/biomedicines9060591
157. Hong H, Zhu H, Zhao S, Wang K, Zhang N, Tian Y, Li Y, Wang Y, Lv X, Wei T, Liu Y, Fan S, Liu Y, Li Y, Cai A, Jin S, Qin Q, Li H. The novel circCLK3/miR-320a/FoxM1 axis promotes cervical cancer progression. Cell Death Dis. 2019; 10(12):950. doi:10.1038/s41419-019-2183-z

158. Lv QL, Du H, Liu YL, Huang YT, Wang GH, Zhang X, Chen SH, Zhou HH. Low expression of microRNA-320b correlates with tumorigenesis and unfavorable prognosis in glioma. Oncol Rep. 2017; 38(2):959-66. doi:10.3892/or.2017.5762
159. Wang WY, Lu WC. Reduced expression of hsa-miR-338-3p contributes to the development of glioma cells by targeting mitochondrial 3-oxoacyl-acp synthase (oxsm) in glioblastoma (gbm). Onco Targets Ther. 2020; 13:9513-23. doi:10.2147/ott.S262873
160. Sun H, Fan G, Deng C, Wu L. MiR-4429 sensitized cervical cancer cells to irradiation by targeting RAD51. J Cell Physiol. 2020; 235(1):185-93. doi:10.1002/jcp.28957
161. Wang J, Xie S, Liu J, Li T, Wang W, Xie Z. MicroRNA-4429 suppresses proliferation of prostate cancer cells by targeting distal-less homeobox 1 and inactivating the Wnt/β-catenin pathway. BMC Urol. 2021; 21(1):40. doi:10.1186/s12894-021-00810-x
162. Stoen MJ, Andersen S, Rakaee M, Pedersen MI, Ingebriktsen LM, Bremnes RM, Donnem T, Lombardi APG, Kilvaer TK, Busund LT, Richardsen E. High expression of miR-17-5p in tumor epithelium is a predictor for poor prognosis for prostate cancer patients. Sci Rep. 2021; 11(1):13864. doi:10.1038/s41598-021-93208-6
163. Bao Y, Wang L, Shi L, Yun F, Liu X, Chen Y, Chen C, Ren Y, Jia Y. Transcriptome profiling revealed multiple genes and ECM-receptor interaction pathways that may be associated with breast cancer. Cell Mol Biol Lett. 2019; 24(1):38. doi:10.1186/s11658-019-0162-0
164. Andersen MK, Rise K, Giskeødegård GF, Richardsen E, Bertilsson H, Størkersen Ø, Bathen TF, Rye M, Tessem MB. Integrative metabolic and transcriptomic profiling of prostate cancer tissue containing reactive stroma. Sci Rep. 2018; 8(1):14269. doi:10.1038/s41598-018-32549-1
165. Heldin CH, Lu B, Evans R, Gutkind JS. Signals and receptors. Cold Spring Harb Perspect Biol. 2016; 8(4):a005900. doi:10.1101/cshperspect.a005900
166. Kaminska B, Wesolowska A, Danilkiewicz M. TGF beta signalling and its role in tumour pathogenesis. Acta Biochim Pol. 2005; 52(2):329-37.
167. Akhurst RJ, Hata A. Targeting the TGFβ signalling pathway in disease. Nat Rev Drug Discov. 2012; 11(10):790-811. doi:10.1038/nrd3810
168. Cai C, Zhang M, Liu L, Zhang H, Guo Y, Lan T, Xu Y, Ma P, Li S. ADAM10-cleaved ephrin-A5 contributes to prostate cancer metastasis. Cell Death Dis. 2022; 13(5):453. doi:10.1038/s41419-022-04893-8
169. Sun F, Yu M, Yu J, Liu Z, Zhou X, Liu Y, Ge X, Gao H, Li M, Jiang X, Liu S, Chen X, Guan W. MiR-338-3p functions as a tumor suppressor in gastric cancer by targeting PTP1B. Cell Death Dis. 2018; 9(5):522. doi:10.1038/s41419-018-0611-0
170. Wang Y, Qin H. MiR-338-3p targets RAB23 and suppresses tumorigenicity of prostate cancer cells. Am J Cancer Res. 2018; 8(12):2564-74.
171. Liu Y, Tang W, Ren L, Liu T, Yang M, Wei Y, Chen Y, Ji M, Chen G, Chang W, Xu J. Activation of miR-500a-3p/CDK6 axis suppresses aerobic glycolysis and colorectal cancer progression. J Transl Med. 2022; 20(1):106. doi:10.1186/s12967-022-03308-8

Appendices

Appendix A: Ethics approval

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Faculty of Health Sciences

Institution: The Research Ethics Committee, Faculty Health Sciences, University of Pretoria complies with ICH-GCP guidelines and has US Federal wide Assurance.

- FWA 00002567, Approved dd 18 March 2022 and Expires 18 March 2027.
- IORG #: IORG0001762 OMB No. 0990-0278 Approved for use through August 31, 2023.

Faculty of Health Sciences **Research Ethics Committee**

16 February 2023

Approval Certificate
Annual Renewal

Dear Miss MP Mbeje,

Ethics Reference No.: 561/2021 – Line 2
Title: Molecular profiling of long non-coding RNAs in prostate cancer cell lines

The **Annual Renewal** as supported by documents received between 2023-01-18 and 2023-02-15 for your research, was approved by the Faculty of Health Sciences Research Ethics Committee on 2023-02-15 as resolved by its quorate meeting.

Please note the following about your ethics approval:

- Renewal of ethics approval is valid for 1 year, subsequent annual renewal will become due on 2024-02-16.
- Please remember to use your protocol number (561/2021) on any documents or correspondence with the Research Ethics Committee regarding your research.
- Please note that the Research Ethics Committee may ask further questions, seek additional information, require further modification, monitor the conduct of your research, or suspend or withdraw ethics approval.

Ethics approval is subject to the following:

- The ethics approval is conditional on the research being conducted as stipulated by the details of all documents submitted to the Committee. In the event that a further need arises to change who the investigators are, the methods or any other aspect, such changes must be submitted as an Amendment for approval by the Committee.

We wish you the best with your research.

Yours sincerely

On behalf of the FHS REC, Dr R Sommers
MBChB, MMed (Int), MPharmMed, PhD

The Faculty of Health Sciences Research Ethics Committee complies with the SA National Act 61 of 2003 as it pertains to health research and the United States Code of Federal Regulations Title 45 and 46. This committee abides by the ethical norms and principles for research, established by the Declaration of Helsinki, the South African Medical Research Council Guidelines as well as the Guidelines for Ethical Research: Principles Structures and Processes, Second Edition 2015 (Department of Health)

Research Ethics Committee
Room 4-60, Level 4, Tswelopele Building
University of Pretoria, Private Bag x323
Gezina 0031, South Africa
Tel +27 (0)12 356 3084
Email: deepeka.behari@up.ac.za
www.up.ac.za

Fakulteit Gesondheidswetenskappe
Lefapha la Disaense tša Maphelo

Appendix B

Table B1. Details of the lncRNAs used in this study.

LncRNA	Full name	GenBank ID	Ensembl ID
AC003075.4	N/A	N/A	ENSG00000237773
AC005618.6	N/A	N/A	ENSG00000272070
AC006994.1	N/A	N/A	ENSG00000231063
AC009404.2	N/A	N/A	ENSG00000236255
AFG3L1P	AFG3-like AAA ATPase 1, pseudogene	NR_003228, NR_003226, NR_003227	ENSG00000223959
B2M	Beta-2-microglobulin	NC_000015.9, NG_012920.1, NT_010194.17	ENSG00000166710
BCDIN3D-AS1	BCDIN3D antisense RNA 1	NR_027499, NR_027501, NR_027500	ENSG00000258057
CBR3-AS1	CBR3 antisense RNA 1	NR_038893.1 NR_038892.1 NR_038894.1 (transcript variants 2,1, and 3)	ENSG00000236830
CFLAR-AS1	CFLAR antisense RNA 1	NR_040030.1	ENSG00000226312
COX6CP10 (COX6CP1)	Cytochrome c oxidase subunit 6C pseudogene 1	N/A	ENSG00000238008
CTA-363E19.2	N/A	N/A	ENSG00000272072
CTD-2010I16.1	N/A	N/A	ENSG00000254595
CTD-2339M3.1	N/A	N/A	ENSG00000242602
CTD-2562J17.7	N/A	N/A	ENSG00000254429
CTD-3157E16.1	N/A	N/A	ENSG00000265519
DUTP1	Deoxyuridine triphosphatase pseudogene 1	N/A	ENSG00000229048
FAM83-AS1	FAM83H antisense RNA 1 (head-to-head)	NR_033849	ENSG00000203499

FGD5-AS1	FGD5 antisense RNA 1	NR_046251, NR_046255, NR_046254, NR_046253, NR_046252	ENSG00000225733
FOXP4-AS1	FOXP4 antisense RNA 1	NR_126415, NR_126416, NR_126417	ENSG00000234753
GNG12-AS1	GNG12 antisense RNA 1 (GNG12, DIRAS3 and WLS antisense RNA 1)	NR_040077.1	ENSG00000232284
GTF3AP2	General transcription factor IIIA pseudogene 2	N/A	ENSG00000258966
HMBS	Hydroxymethylbilane synthase	NC_000011.9, NG_008093.1, NT_033899.8	ENSG00000256269
HMGN3-AS1	HMGN3 antisense RNA 1	NR_040671.1	ENSG00000270362
HOTAIR	HOX transcript antisense RNA	NR_047517, NR_047518, NR_003716	ENSG00000228630
KB-1562D12.2	N/A	N/A	ENSG00000253991
KCNQ1OT1	KCNQ1 opposite strand/antisense transcript 1	NR_002728.3	ENSG00000269821
KRT81	N/A	N/A	ENSG00000205426
LINC00339	Long intergenic non-protein coding RNA 339	NR_110692.1 NR_023918.2 NR_109762.1 NR_023919.2 NR_109760.1 NR_109761.1 NR_109759.1	ENSG00000218510
LINC00493/SMIM26	Small integral membrane protein 26	NM_001348957.2 NM_001348958.2 (transcript var. 1 and 2	ENSG00000232388
LINC00657	Long intergenic non-protein coding RNA 657	NR_027451	ENSG00000260032
LINC01024/	Long intergenic non-protein coding RNA	NR_102741.1	ENSG00000245146

MALINC1	1024/ mitosis associated long intergenic non-coding RNA 1	NR_102740.1 NR_102739.1 NR_038842	
LINC01137	Long intergenic non-protein coding RNA 1137	N/A	ENSG00000233621
lnc-FAM83G-3	N/A	N/A	lnc-FAM83G-3
LURAP1L-AS1	LURAP1L antisense RNA 1	NR_125775	ENSG00000235448
MAGI1-IT1	MAGI1 intronic transcript 1	NR_145422.1	ENSG00000272610
NHS-AS1	NHS antisense RNA 1	NR_046632	ENSG00000230020
NRSN2-AS1	NRSN2 antisense RNA 1	N/A	ENSG00000225377
PCA3	Prostate cancer-associated 3 (non-protein coding)	NR_132312.1 NR_015342.2 NR_132313.1 (transcript variants 1,2 and 3)	ENSG00000225937
PCAT1	Prostate cancer-associated transcript 1 (non-protein coding)	NR_045262	ENSG00000253438
PCAT29	Prostate cancer-associated transcript 29 (non-protein coding)	NR_126437.1 NR_126438.1	ENSG00000259641
PCGEM1	Prostate-specific transcript (non-protein coding)	NR_002769.1	ENSG00000227418
PRNCR1	Prostate cancer-associated non-coding RNA 1	NR_109833.1	ENSG00000282961
PSMD6-AS2	PSMD6 antisense RNA 2	NR_038286	ENSG00000239653
PTPRJAS1	PTPRJ antisense RNA 1	N/A	ENSG00000254879
RN7SL15P	N/A	N/A	ENSG00000264573
RN7SL277P	RNA, 7SL, cytoplasmic 277,	N/A	ENSG00000240490

	pseudogene		
RNF139-AS1	RNF139 antisense RNA 1 (head-to-head)	NR_108047, NR_108048	ENSG00000245149
RP11-119B16.2	N/A	N/A	ENSG00000229539
RP11-121A14.2	N/A	N/A	ENSG00000235204
RP11-146E13.4	N/A	N/A	ENSG00000277128
RP11-156E6.1	N/A	N/A	ENSG00000259623
RP11-174G6.5	N/A	N/A	ENSG00000261324
RP11-196G18.22	N/A	N/A	ENSG00000261716
RP11-226L15.5	N/A	N/A	ENSG00000260766
RP11-244F12.3	N/A	N/A	ENSG00000259498
RP11-338I21.1	N/A	N/A	ENSG00000271933
RP11-363E7.4	N/A	N/A	ENSG00000260912
RP11-379H18.1	N/A	N/A	ENSG00000271122
RP11-382A20.4	N/A	N/A	ENSG00000259986
RP11-390P2.4	N/A	N/A	ENSG00000225177
RP11-46C24.7	N/A	N/A	ENSG00000259877
RP11-473I1.5	N/A	N/A	ENSG00000260349
RP11-48B3.3	N/A	N/A	ENSG00000254162
RP11-545E17.3	N/A	NR_046240	ENSG00000223478
RP11-556O9.2	N/A	N/A	ENSG00000264513
RP11-619J20.1	Pre-mRNA processing factor 31	NM_015629.4	ENSG00000249678
RP11-67L2.2	N/A	NR_024618	ENSG00000273033
RP11-759F5.1	N/A	N/A	ENSG00000236327
RP11-884K10.6	N/A	N/A	ENSG00000271916
RP11-98I9.4	N/A	NR_125843	ENSG00000228506
RP1-30M3.6	N/A	N/A	ENSG00000272402
RP3-443C4.2	N/A	N/A	ENSG00000233823
RP3-523E19.2	N/A	N/A	ENSG00000271218

RP4- 635E18.8	N/A	N/A	ENSG00000226984
RP4-591B8.2	N/A	N/A	ENSG00000271895
RP4-724E16.2	N/A	NR_110051	ENSG00000197670
SCHLAP1	SWI/SNF complex antagonist associated with prostate cancer 1 (non-protein coding)	NR_104322.1, NR_104325.1, NR_104324.1, NR_104323.1, NR_104319.1, NR_104320.1, NR_104321.1	ENSG00000281131
SNHG19	Small nucleolar RNA host gene 19	NR_132114	ENSG00000260260
SNHG8	Small nucleolar RNA host gene 8	NR_003584.3, NR_034010.1, NR_034011.1	ENSG00000269893
SOCS2-AS1	SOCS2 antisense RNA 1	NR_038263.1	ENSG00000246985
TBP	TATA box binding protein	NC_000006.11, NT_025741.15, NG_008165.1	ENSG00000112592
TERC	Telomerase RNA component	NR_001566.1	ENSG00000270141
TINCR	Terminal differentiation-induced non-coding RNA	NM_001396408.1	ENSG00000223573
TP53TG1	TP53 target 1 (non-protein coding)	NR_015381	ENSG00000182165
TUG1	Taurine up-regulated 1 (non-protein coding)	NR_110492, NR_002323, NR_110493	ENSG00000253352
UBA6-AS1	UBA6 antisense RNA 1 (head-to-head)	NR_015439	ENSG00000248049
WAC-AS1	WAC antisense RNA 1 (head-to-head)	N/A	ENSG00000254635
XIST	X inactive specific transcript (non-protein coding)	NR_001564.2	ENSG00000229807
ZNF767P	Zinc finger family member 767, pseudogene	NR_027788, NR_027789	ENSG00000133624

Appendix C: Supplementary tables and figures

C1: cDNA verification

Table C1.1: Quantification cycle (Cq) values from real-time qPCR.

Sample ID	Cq Value
LNCaP sample 1	16.41
LNCaP sample 2	17.47
LNCaP sample 3	17.1
LNCaP sample 4	18.47
PC-3 sample 1	15.25
PC-3 sample 2	16.97
PC-3 sample 3	17.09
PC-3 sample 4	18.08

Table C1.2: The melting temperatures of the various cDNA samples.

Sample ID	Melt Temperature (°C)
LNCaP sample 1	83.71
LNCaP sample 2	83.70
LNCaP sample 3	83.68
LNCaP sample 4	83.75
PC-3 sample 1	83.55
PC-3 sample 2	83.59
PC-3 sample 3	83.58
PC-3 sample 4	83.71

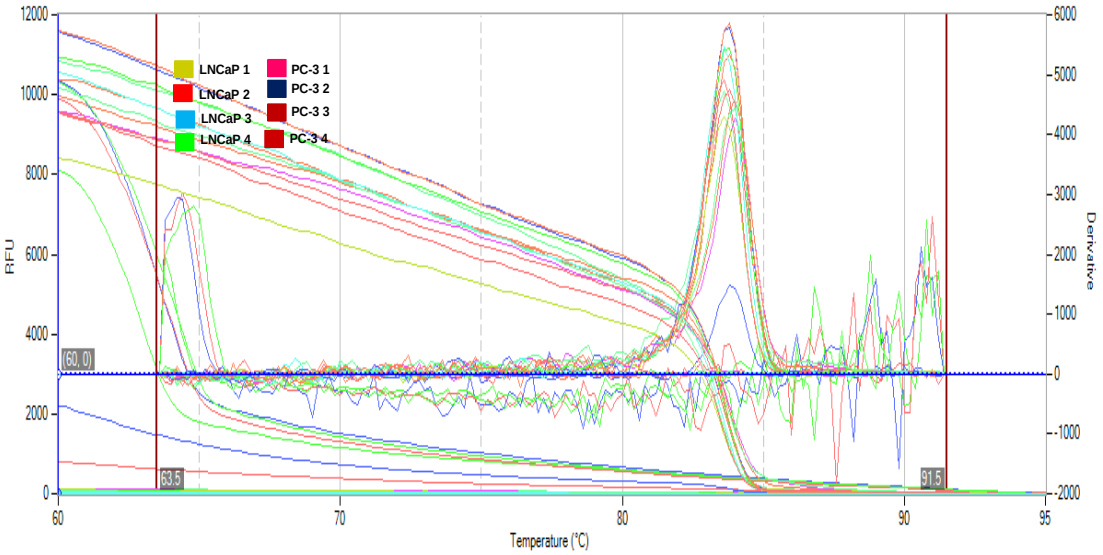
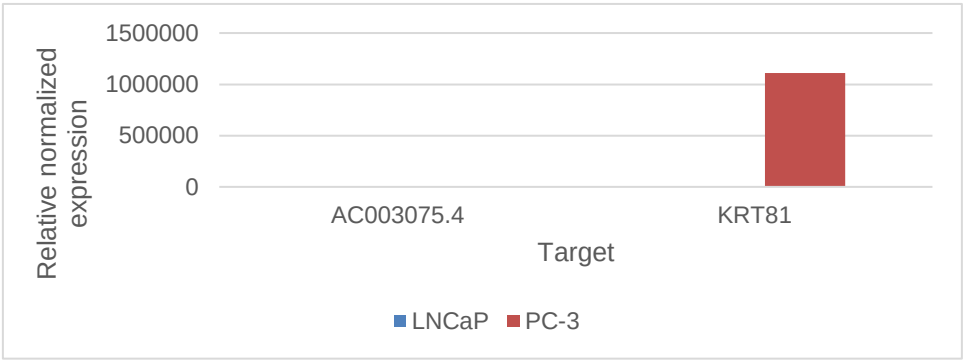


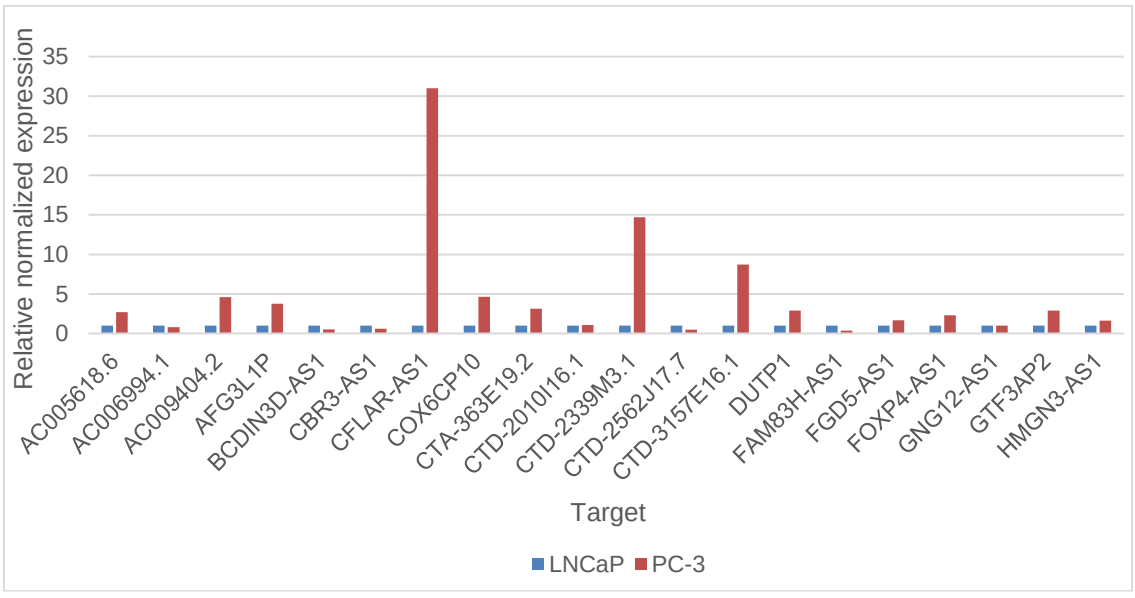
Figure C1.1: Melt curve generated using human GAPDH primers and the cDNA samples from PC-3 and LNCaP cell lines. The melt curve shows that the samples have similar melting temperatures of about 83°C.

C2: CFX Maestro normalized lncRNA expression bar graphs.

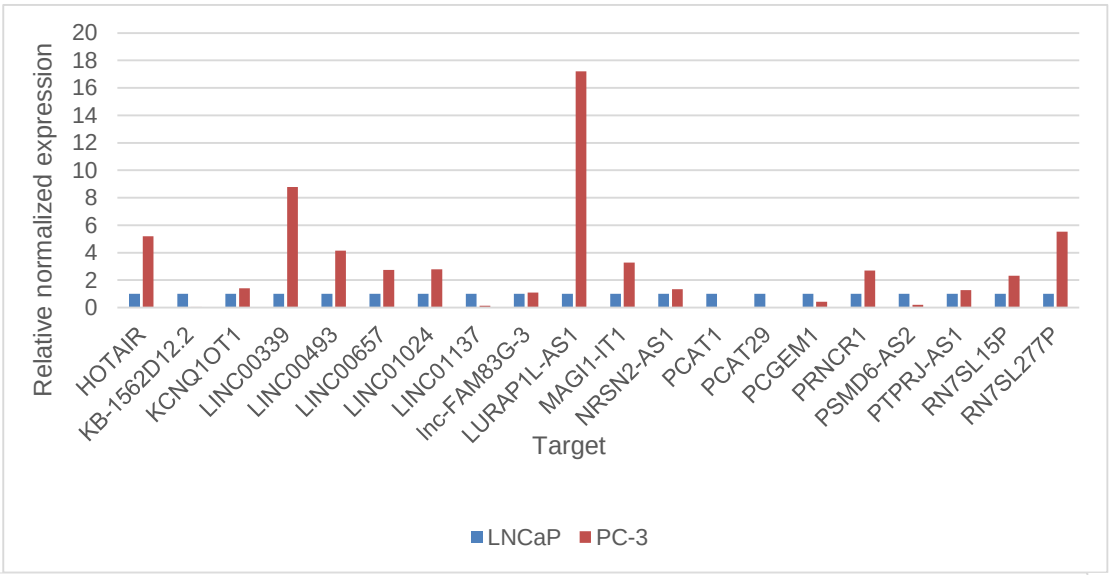
A.



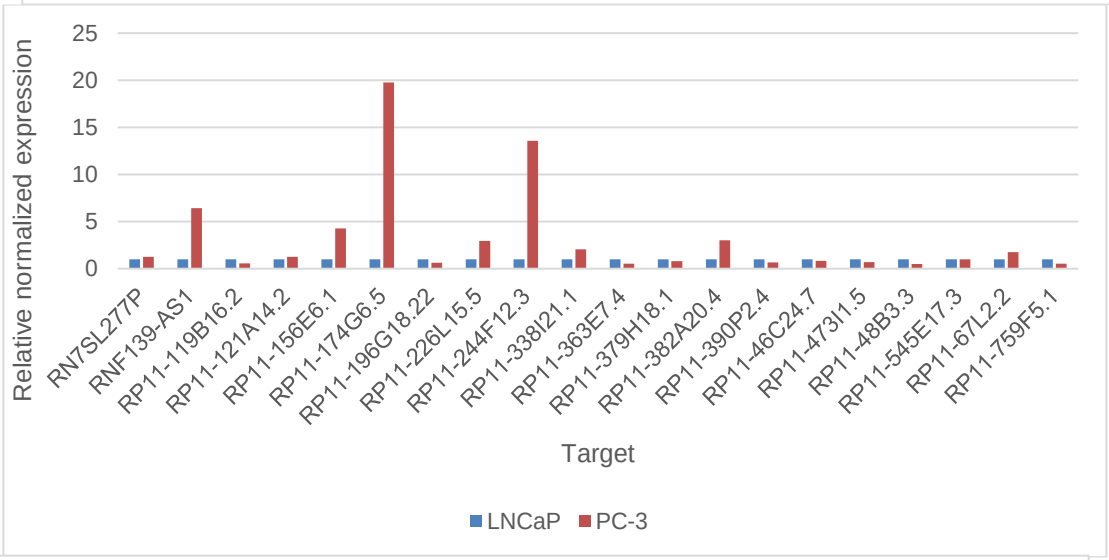
B.



C.



D.



E.

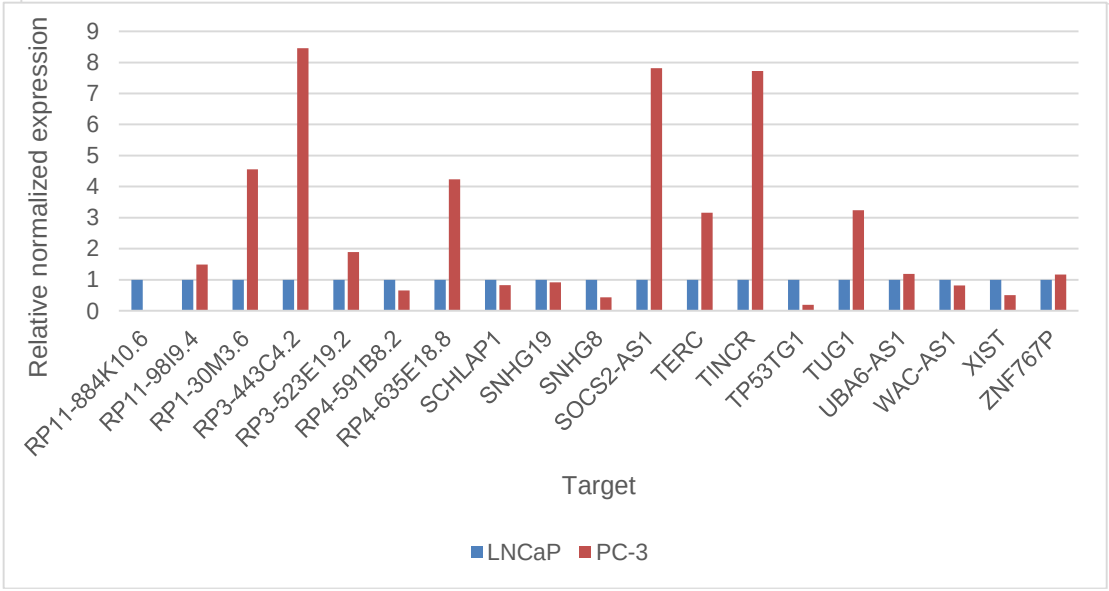


Figure C2.1: Bar graphs showing the relative normalized expression of each lncRNA for the two cell lines, LNCaP (control) and PC-3 (experiment). The relative normalized expression of LNCaP is 1.00 as this is the control.

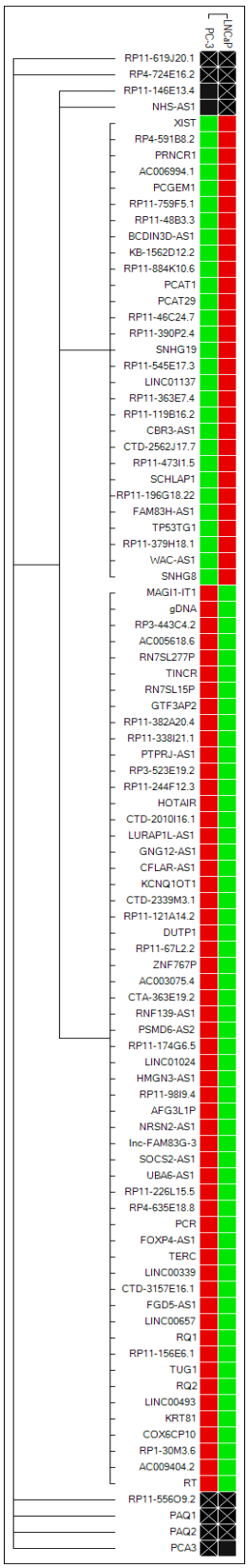


Figure C2.2: Clustergram with the various differentially expressed lncRNAs. The overexpressed genes are represented in red. The downregulated lncRNAs are shown in green. The lncRNAs with unchanged expressions are shown in black. The lncRNAs with similar expression patterns are grouped (clustered) together.

C3: LncSEA analysis of upregulated and downregulated lncRNAs

Table C3.1: LncSEA enrichment analysis of upregulated lncRNAs

Enrichment set	Class	Subclass	Count	LncRNA	P-value	FDR
proliferation	Cancer Hallmark	_	5	HOTAIR; LINC00657; SOCS2-AS1; TINCR; TUG1	1.39E-07	9.73E-07
apoptosis	Cancer Hallmark	_	4	HOTAIR; SOCS2-AS1; TINCR; TUG1	3.55E-07	1.24E-06
metastasis	Cancer Hallmark	_	3	HOTAIR; TINCR; TUG1	2.60E-06	6.07E-06
invasion	Cancer Hallmark	_	3	HOTAIR; TINCR; TUG1	5.18E-05	9.06E-05
migration	Cancer Hallmark	_	3	HOTAIR; TINCR; TUG1	6.78E-05	9.49E-05
prognosis	Cancer Hallmark	_	3	HOTAIR;LINC00657;TUG1	0.000139	0.000162
EMT	Cancer Hallmark	_	2	HOTAIR;TUG1	0.000313	0.000313
Oesophageal squamous cancer	Disease	LncRNADisease2.0	4	CFLAR-AS1;HOTAIR;TINCR;TUG1	1.09E-08	4.20E-06
Oesophageal squamous cancer	Disease	MNDR2.0	3	HOTAIR;TINCR;TUG1	4.70E-08	1.56E-05
Glioma	Disease	Lnc2Cancer2.0	4	HOTAIR;LINC00339;TINCR;TUG1	1.95E-07	2.05E-05
Oesophageal squamous cancer	Disease	Lnc2Cancer2.0	4	CFLAR-AS1;HOTAIR;TINCR;TUG1	2.72E-07	2.05E-05
Prostate cancer	Disease	LncRNADisease2.0	4	HOTAIR;SOCS2-AS1;TERC;TUG1	2.98E-07	5.74E-05
Colorectal carcinoma	Disease	MNDR2.0	4	FOXP4-AS1;HOTAIR;TINCR;TUG1	7.43E-07	0.000123
Squamous cell carcinoma	Disease	MNDR2.0	3	HOTAIR;TINCR;TUG1	1.51E-06	0.000167
Squamous cell carcinoma	Disease	LncRNADisease2.0	3	HOTAIR;TINCR;TUG1	1.64E-06	0.00021
Non-small cell lung cancer	Disease	Lnc2Cancer2.0	4	HOTAIR;LINC00339;TINCR;TUG1	2.12E-06	0.000107
Bladder urothelial cancer	Disease	MNDR2.0	2	HOTAIR;TUG1	2.21E-06	0.000172
Bladder cancer	Disease	MNDR2.0	3	HOTAIR;TINCR;TUG1	2.60E-06	0.000172
Lung small cell carcinoma	Disease	LncRNADisease2.0	2	HOTAIR;TUG1	3.68E-06	0.000326
Colorectal carcinoma	Disease	LncRNADisease2.0	4	FOXP4-AS1;HOTAIR;TINCR;TUG1	4.24E-06	0.000326
Prostate cancer	Disease	EVLncRNAs	3	HOTAIR;TERC;TUG1	4.38E-06	0.00115
Pancreatic ductal adenocarcinoma	Disease	Lnc2Cancer2.0	3	HOTAIR;LINC00339;TUG1	4.64E-06	0.000175
B cell lymphoma	Disease	LncRNADisease2.0	2	HOTAIR;TUG1	5.52E-06	0.000354

B cell lymphoma	Disease	MNDR2.0	2	HOTAIR;TUG1	5.52E-06	0.000305
Multiple myeloma	Disease	MNDR2.0	2	HOTAIR;TUG1	7.73E-06	0.000366
Small cell lung cancer	Disease	Lnc2Cancer2.0	2	HOTAIR;TUG1	7.73E-06	0.000233
Bladder cancer	Disease	LncRNADisease2.0	3	HOTAIR;TINCR;TUG1	1.09E-05	6.00E-04
Colorectal carcinoma	Disease	EVLncRNAs	3	HOTAIR;TINCR;TUG1	1.14E-05	0.00115
Bladder urothelial cancer	Disease	LncRNADisease2.0	2	HOTAIR;TUG1	1.32E-05	0.000635
Multiple myeloma	Disease	EVLncRNAs	2	HOTAIR;TUG1	1.32E-05	0.00115
Prostatic Neoplasms	Disease	MNDR2.0	3	HOTAIR;SOCS2-AS1;TUG1	1.46E-05	0.000604
Pre-eclampsia	Disease	LncRNADisease2.0	2	HOTAIR;TUG1	1.65E-05	0.000706
Urothelial cancer	Disease	Lnc2Cancer2.0	2	HOTAIR;TUG1	1.65E-05	0.000347
Malignant melanoma	Disease	Lnc2Cancer2.0	2	HOTAIR;TUG1	1.65E-05	0.000347
Bladder cancer	Disease	Lnc2Cancer2.0	3	HOTAIR;TINCR;TUG1	1.84E-05	0.000347
Chronic obstructive pulmonary disease	Disease	LncRNADisease2.0	2	TINCR;TUG1	2.02E-05	0.000778
Gastric cancer	Disease	EVLncRNAs	3	HOTAIR;TINCR;TUG1	2.85E-05	0.00187
Prostate cancer	Disease	Lnc2Cancer2.0	3	HOTAIR;SOCS2-AS1;TUG1	3.04E-05	0.00051
Acute myeloid leukaemia	Disease	LncRNADisease2.0	2	HOTAIR;TUG1	5.61E-05	0.00196
Glioma	Disease	EVLncRNAs	2	HOTAIR;TUG1	6.96E-05	0.00365
Gastric cancer	Disease	MNDR2.0	3	HOTAIR;TINCR;TUG1	7.44E-05	0.00255
Pancreatic ductal adenocarcinoma	Disease	MNDR2.0	2	HOTAIR;LINC00339	7.69E-05	0.00255
Gallbladder cancer	Disease	Lnc2Cancer2.0	2	HOTAIR;TUG1	7.69E-05	0.00116
Oral squamous cell carcinoma	Disease	LncRNADisease2.0	2	HOTAIR;TUG1	8.45E-05	0.00259
Oral squamous cell carcinoma	Disease	Lnc2Cancer2.0	2	HOTAIR;TUG1	8.45E-05	0.00116
Lung squamous cell carcinoma	Disease	LncRNADisease2.0	2	TERC;TUG1	9.25E-05	0.00259
Epithelial ovarian cancer	Disease	Lnc2Cancer2.0	2	HOTAIR;TUG1	9.25E-05	0.00116
Non-small cell lung cancer	Disease	LncRNADisease2.0	3	HOTAIR;TINCR;TUG1	9.64E-05	0.00259
Multiple myeloma	Disease	LncRNADisease2.0	2	HOTAIR;TUG1	0.000101	0.00259

Gallbladder cancer	Disease	LncRNADisease2.0	2	HOTAIR;TUG1	0.00011	0.00265
Multiple myeloma	Disease	Lnc2Cancer2.0	2	HOTAIR;TUG1	0.00011	0.00128
Bladder cancer	Disease	EVLncRNAs	2	HOTAIR;TUG1	0.000119	0.0052
Thyroid cancer	Disease	Lnc2Cancer2.0	2	HOTAIR;TUG1	0.000148	0.0016
Endometrial cancer	Disease	Lnc2Cancer2.0	2	HOTAIR;TUG1	0.000181	0.00182
Breast cancer	Disease	Lnc2Cancer2.0	3	HOTAIR;TINCR;TUG1	0.000252	0.00238
Colorectal carcinoma	Disease	Lnc2Cancer2.0	3	HOTAIR;TINCR;TUG1	0.000287	0.00255
Exosome	Exosome	_	2	HOTAIR;TUG1	0.00017	0.00017
hypermethylation	Methylation Pattern	_	2	HOTAIR;TERC	0.000138	0.00051
differential methylation	Methylation Pattern	_	2	HOTAIR;TERC	0.000204	0.00051
hsa-miR-4698	MicroRNA	lncbase	4	LINC00657;RP11-156E6.1;RP11-174G6.5;TUG1	1.77E-08	1.26E-05
hsa-let-7b-5p	MicroRNA	lncbase	6	AC009404.2;AFG3L1P;HOTAIR;LINC00657;RP11-156E6.1;TUG1	2.64E-08	1.26E-05
hsa-miR-485-5p	MicroRNA	lncbase	5	LINC00657;RP11-156E6.1;RP11-226L15.5;TINCR;TUG1	5.57E-08	1.57E-05
hsa-miR-93-5p	MicroRNA	starbase	4	HOTAIR;LINC00657;MAGI1-IT1;RP11-226L15.5	5.76E-08	2.28E-06
hsa-miR-106b-5p	MicroRNA	starbase	4	HOTAIR;LINC00657;MAGI1-IT1;RP11-226L15.5	5.76E-08	2.28E-06
hsa-miR-17-5p	MicroRNA	starbase	4	HOTAIR;LINC00657;MAGI1-IT1;RP11-226L15.5	5.76E-08	2.28E-06
hsa-miR-20a-5p	MicroRNA	starbase	4	HOTAIR;LINC00657;MAGI1-IT1;RP11-226L15.5	5.76E-08	2.28E-06
hsa-miR-106a-5p	MicroRNA	starbase	4	HOTAIR;LINC00657;MAGI1-IT1;RP11-226L15.5	5.76E-08	2.28E-06
hsa-miR-20b-5p	MicroRNA	starbase	4	HOTAIR;LINC00657;MAGI1-IT1;RP11-226L15.5	5.76E-08	2.28E-06
hsa-miR-519d-3p	MicroRNA	starbase	4	HOTAIR;LINC00657;MAGI1-IT1;RP11-226L15.5	5.76E-08	2.28E-06
hsa-miR-92b-3p	MicroRNA	lncbase	4	AFG3L1P;LINC00657;LINC01024;TINCR	6.58E-08	1.57E-05
hsa-miR-548n	MicroRNA	lncbase	4	LINC00657;RP11-156E6.1;RP11-174G6.5;TUG1	1.35E-07	2.58E-05
hsa-miR-143-3p	MicroRNA	lncbase	5	HOTAIR;LINC00657;MAGI1-IT1;RNF139-AS1;TUG1	1.74E-07	2.77E-05
hsa-miR-196b-5p	MicroRNA	starbase	3	RP11-156E6.1;RP11-226L15.5;TUG1	2.06E-07	6.34E-06

hsa-miR-196a-5p	MicroRNA	starbase	3	RP11-156E6.1;RP11-226L15.5;TUG1	2.06E-07	6.34E-06
hsa-let-7d-5p	MicroRNA	lncbase	5	AC009404.2;AFG3L1P;LINC00657;RP11-156E6.1;TUG1	2.21E-07	2.93E-05
hsa-miR-320a	MicroRNA	lncbase	5	LINC00657;RP11-156E6.1;RP11-174G6.5;TERC;TUG1	2.59E-07	2.93E-05
hsa-miR-101-3p	MicroRNA	lncbase	5	AFG3L1P;HOTAIR;LINC00657;RP11-156E6.1;TUG1	2.76E-07	2.93E-05
hsa-miR-125b-2-3p	MicroRNA	lncbase	4	LINC00493;LINC00657;RP11-174G6.5;TUG1	4.37E-07	4.12E-05
hsa-miR-196a-5p	MicroRNA	lncbase	4	AC009404.2;AFG3L1P;RP11-156E6.1;TUG1	4.93E-07	4.12E-05
hsa-miR-92a-3p	MicroRNA	lncbase	4	AFG3L1P;LINC00657;LINC01024;TINCR	5.75E-07	4.12E-05
hsa-miR-106a-5p	MicroRNA	lncbase	4	LINC00657;RP11-156E6.1;RP11-226L15.5;TUG1	7.43E-07	4.12E-05
hsa-miR-186-5p	MicroRNA	lncbase	4	AC005618.6;LINC00657;RP11-156E6.1;TUG1	7.70E-07	4.12E-05
hsa-miR-382-5p	MicroRNA	starbase	3	LINC00657;RP11-174G6.5;TUG1	7.75E-07	2.15E-05
hsa-miR-320b	MicroRNA	lncbase	4	RP11-156E6.1;RP11-174G6.5;TERC;TUG1	8.25E-07	4.12E-05
hsa-miR-320d	MicroRNA	lncbase	4	RP11-156E6.1;RP11-174G6.5;TERC;TUG1	8.25E-07	4.12E-05
hsa-miR-548d-5p	MicroRNA	lncbase	3	LINC00657;RP11-226L15.5;TUG1	8.61E-07	4.12E-05
hsa-let-7a-3p	MicroRNA	lncbase	3	LINC00657;RP11-174G6.5;TUG1	8.61E-07	4.12E-05
hsa-miR-25-3p	MicroRNA	lncbase	4	LINC00657;LINC01024;TINCR;TUG1	1.01E-06	4.12E-05
hsa-let-7f-5p	MicroRNA	lncbase	5	AC009404.2;AFG3L1P;LINC00657;RP11-156E6.1;TUG1	1.03E-06	4.12E-05
hsa-let-7i-5p	MicroRNA	lncbase	5	AC009404.2;AFG3L1P;LINC00657;RP11-156E6.1;TUG1	1.03E-06	4.12E-05
hsa-miR-548c-5p	MicroRNA	lncbase	3	LINC00657;RP11-226L15.5;TUG1	1.05E-06	4.12E-05
hsa-let-7f-2-3p	MicroRNA	lncbase	3	LINC00657;RP11-174G6.5;TUG1	1.05E-06	4.12E-05
hsa-let-7g-5p	MicroRNA	lncbase	5	AC009404.2;AFG3L1P;LINC00657;RP11-156E6.1;TUG1	1.07E-06	4.12E-05
hsa-let-7a-5p	MicroRNA	lncbase	5	AC009404.2;AFG3L1P;LINC00657;RP11-156E6.1;TUG1	1.08E-06	4.12E-05
hsa-miR-26a-5p	MicroRNA	lncbase	5	AC005618.6;LINC00657;RP11-156E6.1;RP4-	1.20E-06	4.40E-05

				635E18.8;TUG1		
hsa-miR-548b-5p	MicroRNA	lncbase	3	RP11-156E6.1;RP11-226L15.5;TUG1	2.25E-06	7.95E-05
hsa-miR-548l	MicroRNA	lncbase	3	LINC00657;RP11-226L15.5;TUG1	3.41E-06	0.000116
hsa-miR-302e	MicroRNA	starbase	3	LINC00657;MAGI1-IT1;RP11-156E6.1	3.64E-06	5.78E-05
hsa-miR-548k	MicroRNA	lncbase	3	LINC00657;RP11-226L15.5;TUG1	3.64E-06	0.00012
hsa-miR-520b	MicroRNA	starbase	3	LINC00657;MAGI1-IT1;RP11-156E6.1	4.12E-06	5.78E-05
hsa-miR-520c-3p	MicroRNA	starbase	3	LINC00657;MAGI1-IT1;RP11-156E6.1	4.12E-06	5.78E-05
hsa-miR-520e	MicroRNA	starbase	3	LINC00657;MAGI1-IT1;RP11-156E6.1	4.12E-06	5.78E-05
hsa-miR-520a-3p	MicroRNA	starbase	3	LINC00657;MAGI1-IT1;RP11-156E6.1	4.12E-06	5.78E-05
hsa-miR-373-3p	MicroRNA	starbase	3	LINC00657;MAGI1-IT1;RP11-156E6.1	4.12E-06	5.78E-05
hsa-miR-340-5p	MicroRNA	lncbase	3	LINC00657;RP11-174G6.5;TUG1	4.12E-06	0.000131
hsa-miR-372-3p	MicroRNA	starbase	3	LINC00657;MAGI1-IT1;RP11-156E6.1	4.38E-06	5.78E-05
hsa-miR-302b-3p	MicroRNA	starbase	3	LINC00657;MAGI1-IT1;RP11-156E6.1	4.38E-06	5.78E-05
hsa-miR-302c-3p	MicroRNA	starbase	3	LINC00657;MAGI1-IT1;RP11-156E6.1	4.38E-06	5.78E-05
hsa-miR-520d-3p	MicroRNA	starbase	3	LINC00657;MAGI1-IT1;RP11-156E6.1	4.38E-06	5.78E-05
hsa-miR-302d-3p	MicroRNA	starbase	3	LINC00657;MAGI1-IT1;RP11-156E6.1	4.38E-06	5.78E-05
hsa-miR-302a-3p	MicroRNA	starbase	3	LINC00657;MAGI1-IT1;RP11-156E6.1	4.64E-06	5.84E-05
hsa-miR-151a-3p	MicroRNA	lncbase	3	AFG3L1P;LINC00657;TUG1	5.50E-06	0.000169
hsa-miR-3613-3p	MicroRNA	lncbase	3	LINC00657;RP11-156E6.1;TUG1	5.81E-06	0.000173
hsa-miR-452-5p	MicroRNA	lncbase	3	LINC00657;RP11-156E6.1;TUG1	6.46E-06	0.000187
hsa-miR-219a-5p	MicroRNA	starbase	2	LINC00339;TUG1	7.73E-06	9.31E-05
hsa-miR-363-5p	MicroRNA	lncbase	2	LINC00657;TUG1	7.73E-06	0.000216
hsa-miR-222-3p	MicroRNA	lncbase	3	LINC00657;RP11-156E6.1;TUG1	7.91E-06	0.000216
hsa-miR-145-5p	MicroRNA	lncbase	4	AC005618.6;LINC00657;RP11-156E6.1;TUG1	9.55E-06	0.000253
hsa-miR-4782-3p	MicroRNA	starbase	2	LINC00339;TUG1	1.03E-05	0.000119
hsa-let-7e-3p	MicroRNA	lncbase	2	LINC00657;RP11-174G6.5	1.03E-05	0.000264
hsa-miR-16-1-3p	MicroRNA	lncbase	3	AFG3L1P;LINC00657;TUG1	1.05E-05	0.000264
hsa-let-7e-5p	MicroRNA	lncbase	4	AC009404.2;AFG3L1P;LINC00657;TUG1	1.10E-05	0.000269

hsa-miR-588	MicroRNA	lncbase	3	LINC00657;RP11-156E6.1;TUG1	1.19E-05	0.000282
hsa-miR-361-5p	MicroRNA	lncbase	4	LINC00657;LINC01024;RP11-156E6.1;TUG1	1.22E-05	0.000282
hsa-miR-4705	MicroRNA	lncbase	3	LINC00657;RP11-156E6.1;TUG1	1.24E-05	0.000282
hsa-miR-30b-5p	MicroRNA	lncbase	4	AC005618.6;LINC00657;RP4-635E18.8;TUG1	1.29E-05	0.000286
hsa-miR-617	MicroRNA	lncbase	2	LINC00657;TUG1	1.32E-05	0.000286
hsa-miR-30c-5p	MicroRNA	lncbase	4	AC005618.6;LINC00657;RP4-635E18.8;TUG1	1.36E-05	0.000288
hsa-miR-222-5p	MicroRNA	lncbase	3	LINC00657;RP11-174G6.5;TINCR	1.40E-05	0.00029
hsa-miR-501-3p	MicroRNA	lncbase	3	AFG3L1P;LINC00657;TUG1	1.52E-05	0.000309
hsa-miR-320c	MicroRNA	lncbase	3	RP11-156E6.1;TERC;TUG1	1.64E-05	0.000326
hsa-miR-486-5p	MicroRNA	lncbase	3	LINC00657;RP11-174G6.5;TUG1	1.71E-05	0.000333
hsa-miR-4429	MicroRNA	lncbase	3	RP11-174G6.5;TERC;TUG1	1.97E-05	0.000376
hsa-miR-1277-3p	MicroRNA	lncbase	2	AFG3L1P;LINC00657	2.42E-05	0.000453
hsa-miR-34a-3p	MicroRNA	lncbase	3	LINC00657;RP11-156E6.1;TUG1	2.68E-05	0.000482
hsa-miR-1305	MicroRNA	lncbase	3	LINC00657;RP11-156E6.1;TUG1	2.68E-05	0.000482
hsa-miR-30a-5p	MicroRNA	lncbase	3	LINC00657;RP11-156E6.1;TUG1	3.04E-05	0.000537
hsa-miR-21-5p	MicroRNA	lncbase	3	LINC00657;RP11-156E6.1;TUG1	3.13E-05	0.000543
hsa-miR-338-3p	MicroRNA	lncbase	4	LINC00339;LINC00657;TERC;TUG1	3.61E-05	0.000592
hsa-miR-501-5p	MicroRNA	lncbase	3	AFG3L1P;LINC00657;TUG1	3.64E-05	0.000592
hsa-miR-589-3p	MicroRNA	lncbase	3	AFG3L1P;LINC00657;TUG1	3.74E-05	0.000592
hsa-miR-766-5p	MicroRNA	lncbase	3	AC005618.6;LINC00657;TUG1	3.74E-05	0.000592
hsa-miR-1247-5p	MicroRNA	lncbase	2	LINC00657;TINCR	3.85E-05	0.000592
hsa-miR-548ay-5p	MicroRNA	lncbase	2	RP11-226L15.5;TUG1	3.85E-05	0.000592
hsa-miR-130a-3p	MicroRNA	lncbase	3	AFG3L1P;HOTAIR;TUG1	3.85E-05	0.000592
hsa-miR-30d-5p	MicroRNA	lncbase	3	LINC00657;RP11-156E6.1;TUG1	4.07E-05	0.000616
hsa-miR-330-3p	MicroRNA	lncbase	3	LINC00657;RP11-156E6.1;TUG1	4.30E-05	0.000641
hsa-miR-30e-5p	MicroRNA	lncbase	3	LINC00657;RP11-156E6.1;TUG1	4.67E-05	0.000685
hsa-miR-548j-5p	MicroRNA	lncbase	2	RP11-226L15.5;TUG1	4.99E-05	0.000721
hsa-miR-26b-3p	MicroRNA	lncbase	3	AFG3L1P;LINC00657;TUG1	5.18E-05	0.000727

hsa-miR-183-5p	MicroRNA	lncbase	3	LINC00657;RP11-226L15.5;TUG1	5.18E-05	0.000727
hsa-miR-221-3p	MicroRNA	starbase	2	HOTAIR;TUG1	6.26E-05	0.000694
hsa-miR-222-3p	MicroRNA	starbase	2	HOTAIR;TUG1	6.96E-05	0.000714
hsa-miR-410-3p	MicroRNA	starbase	2	LINC00657;TUG1v	6.96E-05	0.000714
hsa-miR-145-5p	MicroRNA	starbase	2	RP11-156E6.1;TUG1	0.000101	0.000999
hsa-miR-425-5p	MicroRNA	starbase	2	RP11-226L15.5;TUG1	0.00011	0.00105
hsa-miR-335-5p	MicroRNA	starbase	2	TINCR;TUG1	0.000128	0.00114
hsa-miR-31-5p	MicroRNA	starbase	2	TINCR;TUG1	0.000128	0.00114
hsa-miR-9-5p	MicroRNA	starbase	2	MAGI1-IT1;TUG1	0.000148	0.00124
hsa-miR-142-3p	MicroRNA	starbase	2	RP11-226L15.5;TUG1	0.000148	0.00124
hsa-miR-218-5p	MicroRNA	starbase	2	LINC00339;RP11-156E6.1	0.000159	0.0013

Table C3.2: LncSEA enrichment analysis of downregulated lncRNAs

Enrichment set	Class	Subclass	Count	LncRNA	P-value	FDR
proliferation	Cancer Hallmark	_	4	PCAT1;PCAT29;PCGEM1;PRNCR1	5.94E-08	4.16E-07
migration	Cancer Hallmark	_	3	PCAT1;PCAT29;PCGEM1	2.17E-06	7.60E-06
apoptosis	Cancer Hallmark	_	2	PCAT1;PCGEM1	0.000167	0.00039
invasion	Cancer Hallmark	_	2	PCAT1;PCGEM1	0.000254	0.000444
prognosis	Cancer Hallmark	_	2	PCAT1;PCAT29	0.000495	0.000693
Prostate cancer	Disease	EVLncRNAs	4	PCAT1;PCAT29;PCGEM1;PRNCR1	2.53E-10	6.63E-08
Prostatic Neoplasms	Disease	MNDR2.0	4	PCAT1;PCAT29;PCGEM1;PRNCR1	1.29E-09	4.27E-07
Prostate cancer	Disease	LncRNADisease2.0	4	PCAT1;PCAT29;PCGEM1;PRNCR1	2.57E-09	9.89E-07
Prostate cancer	Disease	Lnc2Cancer2.0	4	PCAT1;PCAT29;PCGEM1;PRNCR1	3.49E-09	5.27E-07
Prostate	Disease	LncRNADisease2.0	3	PCAT1;PCGEM1;PRNCR1	6.39E-09	1.23E-06
Prostate	Disease	MNDR2.0	3	PCAT1;PCGEM1;PRNCR1	6.39E-09	1.06E-06
Gastric cancer	Disease	Lnc2Cancer2.0	4	PCAT1;PRNCR1;SNHG8;TP53TG1	1.40E-07	1.06E-05
Colorectal carcinoma	Disease	LncRNADisease2.0	3	FAM83H-AS1;PCAT1;PRNCR1	5.69E-06	0.00073

Colorectal carcinoma	Disease	Lnc2Cancer2.0	3	FAM83H-AS1;PCAT1;PRNCR1	9.40E-06	0.000473
Gastric cancer	Disease	LncRNADisease2.0	3	PCAT1;PRNCR1;SNHG8	1.17E-05	0.00113
Non-small cell lung cancer	Disease	EVLncRNAs	2	PCAT1;TP53TG1	7.02E-05	0.00809
Cancer	Disease	LncRNADisease2.0	2	PCAT1;TP53TG1	7.98E-05	0.00577
Bladder cancer	Disease	LncRNADisease2.0	2	PCAT1;PCAT29	8.99E-05	0.00577
Colorectal carcinoma	Disease	EVLncRNAs	2	PCAT1;PRNCR1	9.26E-05	0.00809
Ovarian cancer	Disease	Lnc2Cancer2.0	2	PCGEM1;TP53TG1	0.000109	0.00374
Malignant glioma	Disease	MNDR2.0	2	PCGEM1;PRNCR1	0.000121	0.0134
Malignant glioma	Disease	LncRNADisease2.0	2	PCGEM1;PRNCR1	0.000124	0.00626
Glioma	Disease	Lnc2Cancer2.0	2	FAM83H-AS1;TP53TG1	0.000124	0.00374
Lung cancer	Disease	LncRNADisease2.0	2	FAM83H-AS1;PCAT1	0.00013	0.00626
Phytosterols	Drug	Lnc2Cancer	1	PCGEM1	0.000205	0.00384
g-oryzanol	Drug	Lnc2Cancer	1	PCGEM1	0.000205	0.00384
doxorubicin (DOX)	Drug	Lnc2Cancer	1	PCGEM1	0.000205	0.00384
3,3-Diindolylmethane (DIM)	Drug	Lnc2Cancer	1	PCGEM1	0.000205	0.00384
Temozolomide, TMZ	Drug	Lnc2Cancer	1	PRNCR1	0.000205	0.00384
Chemotherapy	Drug	Lnc2Cancer	1	SNHG8	0.000205	0.00384
1(SNHG8)	Drug	Lnc2Cancer	1	SNHG8	0.000205	0.00384
Irinotecan	Drug	LncMap	5	FAM83H-AS1;PCAT1;PCGEM1;RP11-884K10.6;SNHG8	0.000503	0.00984
Topotecan	Drug	LncMap	5	FAM83H-AS1;PCAT1;PCGEM1;RP11-884K10.6;SNHG8	0.00116	0.00984
Lapatinib	Drug	LncMap	4	FAM83H-AS1;PCAT1;SNHG8;TP53TG1	0.00123	0.00984
Exosome	Exosome	_	2	PCGEM1;PRNCR1	1.79E-05	1.79E-05
hsa-miR-1915-3p	MicroRNA	lncbase	2	FAM83H-AS1;SNHG8	1.25E-05	0.0119
Exosome	Subcellular Location	lncRNA	3	CTD-2562J17.7;SNHG8;TP53TG1	0.00756	0.0153

C4: KEGG and GO analyses supplementary figures and tables generated using DIANA-miRPath v3.0

Table C4.1: KEGG enrichment analysis from the top 30 miRNAs determined in LncSEA according to p-value.

KEGG pathway	**#miRNAs**
Signalling pathways regulating pluripotency of stem cells	15
TGF-beta signalling pathway	11
ECM-receptor interaction	10
Proteoglycans in cancer	9
Axon guidance	9
FoxO signalling pathway	9
Mucin-type O-Glycan biosynthesis	8
Hippo signalling pathway	8
Glioma	8
Circadian rhythm	8
Oestrogen signalling pathway	7
Transcriptional misregulation in cancer	6
ErbB signalling pathway	6
Prostate cancer	6
Endocytosis	5
Wnt signalling pathway	5
Hepatitis B	5
Amphetamine addiction	5
Gap junction	5
Protein processing in the endoplasmic reticulum	5
mTOR signalling pathway	5
Viral carcinogenesis	5
Ras signalling pathway	5
Glutamatergic synapse	4
Thyroid hormone signalling pathway	4
MAPK signalling pathway	4

AMPK signalling pathway	4
Neurotrophin signalling pathway	4
Adherens junction	4
Valine, leucine, and isoleucine biosynthesis	4
Chronic myeloid leukaemia	4
Ubiquitin mediated proteolysis	4
Long-term depression	4
Non-small cell lung cancer	4
Glycosaminoglycan biosynthesis - chondroitin sulfate / dermatan sulfate	4
Renal cell carcinoma	3
PI3K-Akt signalling pathway	3
Colorectal cancer	3
Melanoma	3
Regulation of actin cytoskeleton	3
Prolactin signalling pathway	3
Lysine degradation	3
Amoebiasis	3
Choline metabolism in cancer	3
Dopaminergic synapse	3
Glycosphingolipid biosynthesis - ganglio series	3
Endometrial cancer	3
Oxytocin signalling pathway	3
Cell adhesion molecules (CAMs)	3
Prion diseases	2
Pathways in cancer	2
Focal adhesion	2
Adrenergic signalling in cardiomyocytes	2
Endocrine and other factor-regulated calcium reabsorption	2
cAMP signalling pathway	2
cGMP-PKG signalling pathway	2
Adipocytokine signalling pathway	2
Rap1 signalling pathway	2

Platelet activation	2
Cholinergic synapse	2
Retrograde endocannabinoid signalling	2
Steroid hormone biosynthesis	2
Circadian entrainment	2
Protein digestion and absorption	2
Drug metabolism - cytochrome P450	2
Tyrosine metabolism	2
Long-term potentiation	1
Pancreatic cancer	1
Oocyte meiosis	1
Arrhythmogenic right ventricular cardiomyopathy (ARVC)	1
Melanogenesis	1
p53 signalling pathway	1
Morphine addiction	1
Cocaine addiction	1
Cytokine-cytokine receptor interaction	1
Glycosaminoglycan biosynthesis - heparan sulphate/heparin	1
2-Oxocarboxylic acid metabolism	1
Bladder cancer	1
Bacterial invasion of epithelial cells	1
Tight junction	1
Calcium signalling pathway	1
N-Glycan biosynthesis	1
Pantothenate and CoA biosynthesis	1
Aldosterone-regulated sodium reabsorption	1
Salivary secretion	1
Mineral absorption	1
GABAergic synapse	1
Hedgehog signalling pathway	1
Biosynthesis of unsaturated fatty acids	1
Insulin secretion	1

Fatty acid degradation	1
Other types of O-glycan biosynthesis	1
Inositol phosphate metabolism	1
Glycosphingolipid biosynthesis – lacto- and neolacto- series	1
Vitamin B6 metabolism	1
Inflammatory mediator regulation of TRP channels	1
Gastric acid secretion	1
Chemical carcinogenesis	1

Table C4.2: GO analysis from the top 30 miRNAs determined in LncSEA according to p-value.

Go category	**#miRNAs**
Biosynthetic process	30
Cellular nitrogen compound metabolic process	30
Cellular protein modification process	28
Fc-epsilon receptor signalling pathway	26
Biological process	25
Neurotrophin trk receptor signalling pathway	25
Gene expression	18
Epidermal growth factor receptor signalling pathway	16
Cellular component assembly	16
Transcription, DNA-templated	15
Catabolic process	15
Response to stress	14
Blood coagulation	14
Fibroblast growth factor receptor signalling pathway	14
Phosphatidylinositol-mediated signalling	14
Transcription initiation from RNA polymerase II promoter	13
Cell death	13
Viral process	12

Symbiosis, encompassing mutualism through parasitism	12
Macromolecular complex assembly	12
Mitotic cell cycle	11
Post-translational protein modification	11
Membrane organization	9
Nucleobase-containing compound catabolic process	9
Small molecule metabolic process	9
Cellular lipid metabolic process	7
Protein complex assembly	7
Regulation of transcription from RNA polymerase II promoter in response to hypoxia	7
Synaptic transmission	6
Cellular protein metabolic process	6
Cell junction organization	6
Positive regulation of transcription, DNA-templated	6
Cell junction assembly	6
Cellular component disassembly involved in the execution phase of apoptosis	6
Apoptotic signalling pathway	6
Nervous system development	5
Fc-gamma receptor signalling pathway involved in phagocytosis	5
Positive regulation of protein insertion into mitochondrial membrane involved in the apoptotic signalling pathway	5
mRNA metabolic process	5
mRNA processing	5
Nuclear-transcribed mRNA poly(a) tail shortening	5
Glycosaminoglycan metabolic process	5
Cell-cell signalling	4
Axon guidance	4
Toll-like receptor 10 signalling pathway	4

Toll-like receptor tlr1:tlr2 signalling pathway	4
Toll-like receptor tlr6:tlr2 signalling pathway	4
Transforming growth factor beta receptor signalling pathway	4
Trif-dependent toll-like receptor signalling pathway	4
Regulation of small GTPase-mediated signal transduction	4
Toll-like receptor 9 signalling pathway	4
Toll-like receptor 5 signalling pathway	4
Nucleocytoplasmic transport	4
Chromatin organization	4
Cytoskeleton organization	4
Phosphatidylinositol biosynthetic process	4
Vasodilation by norepinephrine-epinephrine involved in the regulation of systemic arterial blood pressure	4
Negative regulation of transcription, DNA-templated	3
Cell-cell junction organization	3
RNA metabolic process	3
Negative regulation of transcription from RNA polymerase II promoter	3
Post-Golgi vesicle-mediated transport	3
In utero embryonic development	3
Extracellular matrix disassembly	3
Energy reserve metabolic process	3
Signal transduction	3
Cellular response to amino acid stimulus	3
Mitotic nuclear envelope disassembly	3
G2/m transition of mitotic cell cycle	2
Myd88-independent toll-like receptor signalling pathway	2
DNA metabolic process	2
Toll-like receptor 4 signalling pathway	2
Protein n-linked glycosylation via asparagine	2
Toll-like receptor signalling pathway	2

Innate immune response	2
Extracellular matrix organization	2
Intrinsic apoptotic signalling pathway	2
Nls-bearing protein import into the nucleus	2
Toll-like receptor 2 signalling pathway	2
Toll-like receptor 3 signalling pathway	2
Positive regulation of transcription from RNA polymerase II promoter	2
Cellular component movement	2
Glutamate receptor signalling pathway	2
Neuron cell-cell adhesion	2
Positive regulation of macroautophagy	2
Sulphur compound metabolic process	2
Viral entry into the host cell	2
O-glycan processing	2
Viral penetration into the host nucleus	2
Dolichol-linked oligosaccharide biosynthetic process	2
Carnitine shuttle	2
Negative regulation of neurotransmitter secretion	2
Nuclear-transcribed mRNA catabolic process, deadenylation-dependent decay	1
Transcription from RNA polymerase II promoter	1
Insulin receptor signalling pathway	1
Post-embryonic development	1
Negative regulation of translation involved in gene silencing by miRNA	1
Adherens junction organization	1
Immune system process	1
RNA splicing	1
Negative regulation of transforming growth factor beta receptor signalling pathway	1

Mammary gland development	1
Muscle cell differentiation	1
Termination of RNA polymerase II transcription	1
Activation of Jun kinase activity	1
Collagen catabolic process	1
Stress-activated mapk cascade	1
Anatomical structure morphogenesis	1
Odontogenesis of dentin-containing tooth	1
Plasma membrane organization	1
Positive regulation of the apoptotic process	1
Cytoskeleton-dependent intracellular transport	1
Chondroitin sulphate metabolic process	1
Protein targeting	1
Bmp signalling pathway	1
Cellular response to hypoxia	1
Regulation of insulin secretion	1
mRNA 3'-end processing	1
Circadian regulation of gene expression	1
Heart development	1
Regulation of protein kinase a signalling	1
Lens induction in camera-type eye	1
Circadian rhythm	1
Tight junction assembly	1
Negative regulation of macrophage-derived foam cell differentiation	1
Pre-miRNA processing	1
Regulation of the release of sequestered calcium ion into the cytosol by sarcoplasmic reticulum	1
Regulation of defence response to virus by virus	1
Regulation of synaptic transmission, glutamatergic	1
Heat generation	1

Cellular glucuronidation	1
Regulation of germinal centre formation	1
Neuronal stem cell maintenance	1
Regulation of cell communication by electrical coupling involved in cardiac conduction	1
Golgi to plasma membrane protein transport	1
Targeting of mRNA for destruction involved in RNA interference	1
Cargo loading into a vesicle	1
Regulation of translation in response to stress	1
Flavonoid glucuronidation	1
Xenobiotic glucuronidation	1
Cellular response to jasmonic acid stimulus	1

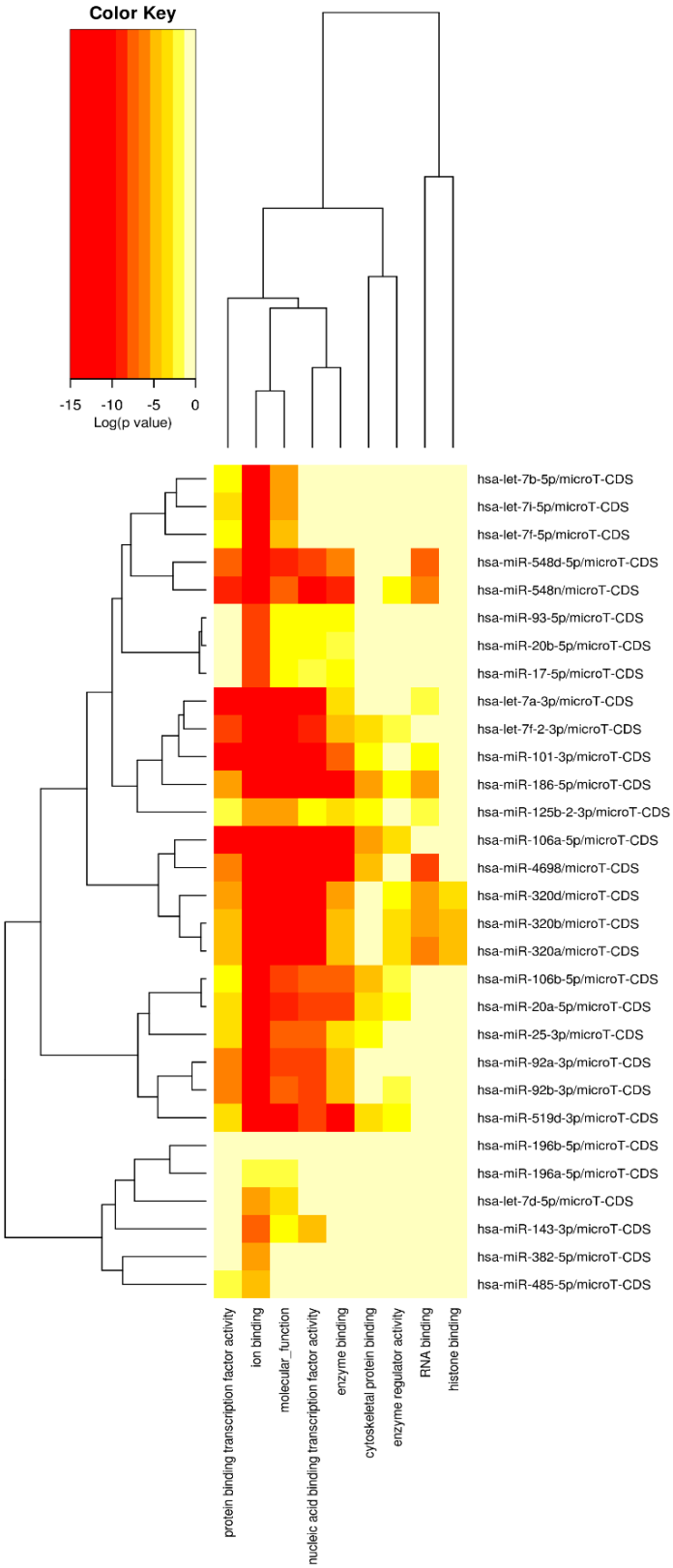


Figure C4.1: Heatmap of the miRNA versus the molecular function GO subcategory (generated by DIANA-miRPath v3.0). p-value<0.05; microT<0.8; Fischer's Test analysis method with FDR correction.

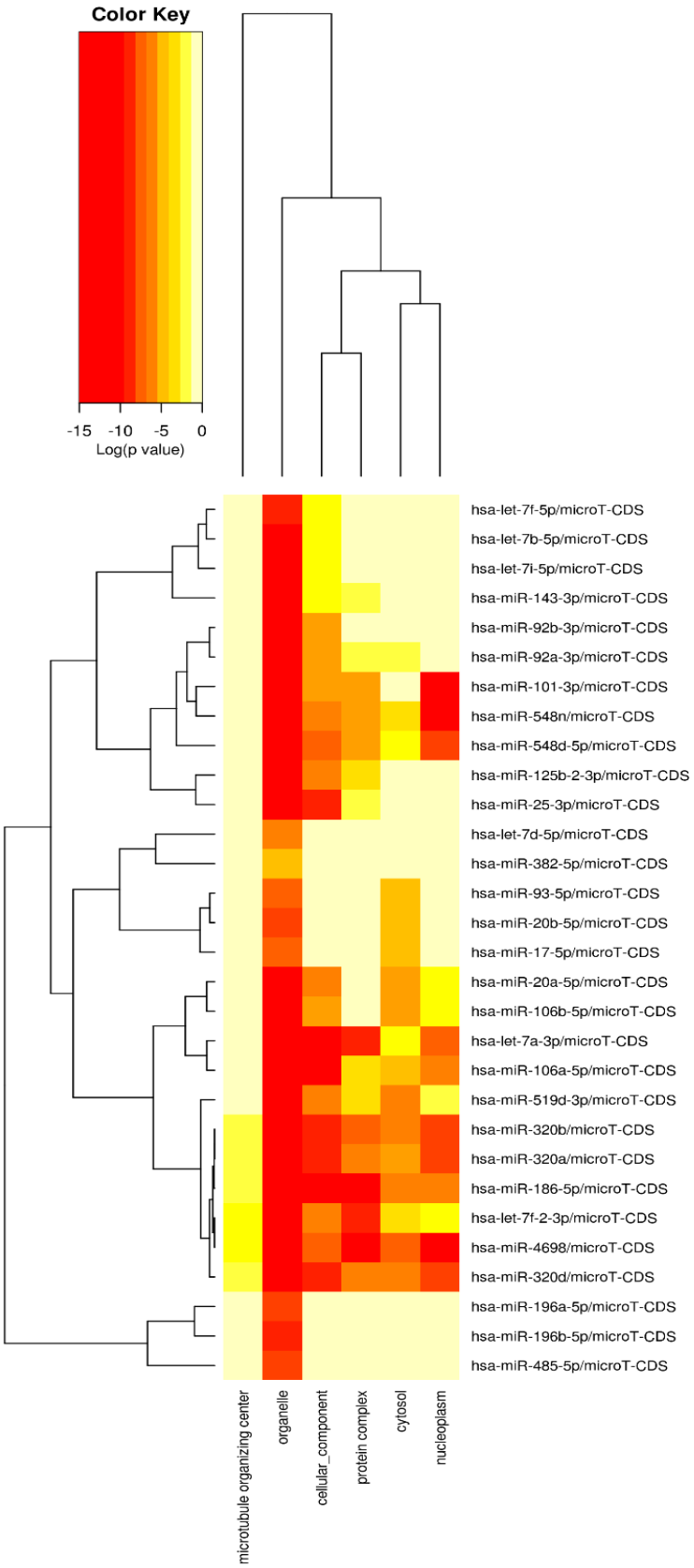


Figure C4.2: Heatmap of the miRNA versus the cellular component GO subcategory generated by DIANA-miRPath v3.0. p-value<0.05; microT<0.8; Fischer's Test analysis method with FDR correction.