Comparative secretome analysis of benign prostate hyperplasia and prostate cancer cell models.

by

Takudzwa Kevin Muvirimi

Thesis presented in fulfilment of the requirements for the degree of Master of Science in the Faculty of Science at Stellenbosch University

Supervisor: Dr. Karl-Heinz Storbeck
Co-supervisor: Dr. Maré Vlok

December 2017

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Kevin Muvirimi

Date: December 2017

Summary

Prostate cancer is the second most prevalent non-cutaneous malignancy affecting men in the world. The current method for screening and diagnosis of prostate cancer relies on the measurement of serum kallikrein 3 (KLK-3) in conjunction with the digital rectal exam. However, serum KLK-3 or prostate specific antigen (PSA) as it is better known, is also elevated by benign prostate hyperplasia (BPH). Consequently, a major challenge in the diagnosis of prostate cancer is the ability to discriminate between prostate cancer and BPH. Misdiagnosis of prostate cancer entails unnecessary treatment for patients which results in psychological stress and physical discomfort. Despite the ineffectiveness of PSA as a biomarker, it is still the primary diagnostic tool. This is largely due to lack of an alternative and it is therefore crucial to find a new biomarker which can distinguish malignant prostate cancer from BPH. In this study, the secretome and proteome of four prostate cell lines (BPH-1, LNCaP, PC3 and PNT2) were investigated to identify novel biomarkers for BPH using a MS based proteomic approach. The analysis of the secretome necessitated that the growth conditions of prostate cell lines were optimized to reduce cell death and maximize secreted proteins. Optimal seeding densities were selected following measurements of total protein, LDH and the LDH/protein ratio. Subsequent optimization of protein precipitation methods identified acetone precipitation as the most suitable method for isolating protein from the conditioned media of prostate cell lines. The cell proteome of the four prostate cell lines was comparatively analysed by LC-MS/MS. The total number of proteins identified per cell line were as follows: 2079 in BPH-1, 2081 in LNCaP, 1853 in PC3 and 2137 proteins in the PNT2 cell line. A literature search of the proteins unique to BPH-1 identified mesencephalic astrocyte-derived neurotrophic factor (MANF) as a potential biomarker for BPH. Ingenuity pathway analysis was subsequently used to analyse and identify aberrant pathways in BPH-1. The following pathways were significantly altered in BPH-1: antiproliferative role of transducer of ERB2 in T cell signalling, pyrimidine ribonucleotide *de novo* biosynthesis, cell signalling and vitamin and mineral metabolism. The optimised secretome samples could not be analysed due to technical difficulties beyond our control. As a proof of concept preliminary analysis of the cell secretome from PC3 and BPH-1 cell lines yielded forty-six and ninety-nine proteins which were unique to each cell line, respectively. The identification of MANF and the antiproliferative role of transducer of ERB2 in T cell signalling during the proteomic analysis both pointed towards an inflammatory response in BPH. Differences in inflammation could therefore be explored in future in order to develop biomarkers unique to BPH and prostate cancer.

Opsomming

Prostaatkanker is die tweede mees algemene kanker ter wêreld wat in mans voorkom. Die huidige metode vir die diagnosering van prostaatkanker is afhanklik van serum kallikrein-3(KLK-3)vlakke tesame met 'n digitale rektumondersoek. Serumvlakke van KLK-3, beter bekend as prostaat-spesifieke-antigeen (PSA), is egter ook verhoog tydens nie-kwaadaardige prostaathiperplasie (BPH). Gevolglik is 'n groot uitdaging in die diagnose van prostaatkanker die vermoë om te onderskei tussen prostaatkanker en BPH. Die misdiagnose van prostaatkanker lei tot onnodige behandelings vir pasiënte wat sielkundige stres en fisiese ongemak tot gevolg kan hê. Ten spyte van die feit dat PSA oneffektief is as 'n biomerker word dit steeds gebruik as 'n primêre of hoof diagnostiese hulpmiddel. Die rede hiervoor kan grootliks toegeskryf word aan 'n tekort aan alternatiewe hulpmiddels en dus om hierdie rede is dit noodsaaklik om 'n nuwe biomerker te identifiseer wat kan onderskei tussen prostaatkanker en BPH.

In hierdie studie is die sekretoom en proteoom van vier sellyne (BPH-1, LNCaP, PC3 en PNT2) ondersoek om nuwe biomerkers te identifiseer vir BPH deur gebruik te maak van 'n MS gebasseerde proteomiese benadering. Die analiese van die sekretoom het vereis dat die groeikondisies van die prostaatkanker-sellyne geoptimiseer word om seldood te vermy, maar terselfde tyd die maksimum afskeiding van proteiene tot gevolg het. Optimale groeidigthede was geselekteer gevolg deur die bepaling van die totale proteïne, LDH en die LDH/proteïen-verhouding. Die optimisering van proteïenpresipitasie-metodes het asetoonpresipitasie identifiseer as die mees geskikte metode vir die isolering van proteïene vanuit prostaatsellyne se gekondisioneerde groeimedium. Die selproteoom van die vier prostaat-sellyne is in 'n relatiewe manier geanaliseer deur gebruik te maak van LC-MS/MS. Die totale aantal proteïene wat geidentifiseer is per sellyn was as volg: 2079 in BPH-1, 2081 in LNCaP, 1853 in PC3 en 2137 proteïene in die PNT2 sellyn. 'n Literatuurstudie van proteïene uniek aan BPH-1 het die mesencefaliese-astrosiet-afgeleide neurotrofe-faktor (MANF) as 'n potensiële biomerker vir BPH identifiseer. "Ingenuity pathway analysis"-sagteware is gevolglik gebruik om padweë wat afwyk te analiseer en identifiseer. Die volgende padweë is aansienlik verander in BPH-1: die antiprolifererende rol van die "transducer" ERB2 in T-sel seinoordrag, pirimidienribonukleotied *de novo* biosintese sowel as selsein-, vitamien- en mineral-metabolisme. Die optimiseerde sekretoommonsters kon weens onvoorsiene tegniese probleme nie geanaliseer word nie. Ten spyte hiervan dien die voorlopige analieses van die selsekretoon van PC3 and BPH-1 sellyne as konseponderstuening en het onderskeidelik ses-en-veertig en nege-en-negentig proteïene opgelewer wat uniek is aan elke sellyn. Die identifikasie van MANF en die antiprolifererende rol van die "transducer" ERB2 in T-sel seinoordrag gedurende die proteomiese analieses is beide 'n aanduiding van inflammatoriese reaksies in BPH. In toekomstige studies kan die verskille in inflammasie dus gebruik word om biomerkers uniek aan BPH en prostaatkanker te ondersoek.

Dedicated to my parents Lilian and Richard Muvirimi and to my sisters Nyasha and Tinashe Muvirimi. I am grateful for your support and prayers, none of this would have been possible without them.

Acknowledgements

I hereby wish to express my sincere gratitude and appreciation to the following people and institutions:

Dr. Karl-Heinz Storbeck: for your patient guidance, encouragement and advice that you have provided throughout my time as a student in your lab. I feel extremely privileged to have had a supervisor who allowed me the space and freedom I needed to work and who responded to my questions and queries so promptly.

Dr. Maré Vlok: for your excellent co-supervision, advice and insights in protein biochemistry.

Ralie Louw: for training me in tissue culture, technical assistance and being an overall great lab manager.

Euan Johnstone: for indulging me in hours of scientific discourse and good banter.

Fellow Students: for creating an enjoyable work environment in the lab and your constant encouragement.

Stellenbosch University, **NRF** and **CANSA** for funding this project.

My Family: for providing me with a constant source of inspiration.

God Almighty, for being present with me every step of this amazing journey.

TABLE OF CONTENTS

CHAPTER 1

CHAPTER 2

CHAPTER 3

CHAPTER 4

LIST OF FIGURES

LIST OF TABLES

TABLES IN APPENDIX

LIST OF ABBREVIATIONS

AMACR	α-methyl CoA racemase
AR	Androgen receptor
ARE	Androgen receptor element
BOO	Bladder outlet obstruction
BPH	Benign prostatic hyperplasia
BPSA	Benign prostatic hyperplasia specific PSA
CA	Cancer antigen
CEA	Carcinoembryonic antigen
CM	Conditioned media
CRPC	Castration resistant prostate cancer
CYP17A1	Cytochrome P450 17- α -hydroxylase/17,20-lyase
DHEA	Dihydroepiandrosterone
DHT	Dihydrotestosterone
DRE	Digital rectal exam
FGF	Fibroblast growth factor
FT-ICR	Fourier transform ion cyclotron
ICAT	Isotope coded affinity tags
IL-6	Interleukin 6
IL-6R	Interleukin 6 receptor
IPA	Ingenuity pathway analysis
iTRAQ	Isobaric tag for relative and absolute quantification
KLK-2	Kallikrein 2
KLK-3	Kallikrein 3
LH	Luteinizing hormone
LHRH	Luteinizing hormone-releasing hormone
LUTS	Lower urinary tract symptoms
MANF	Mesencephalic astrocyte-derived neurotrophic factor
MCM	Minichromosome maintenance protein

MSR-1	Macrophage scavenger receptor 1
NMR	Nuclear magnetic resonance
PA	Plasminogen activator
PAP	Prostatic acid phosphatase
PCa	Prostate cancer
PCA3	Prostate cancer gene 3
PIA	Proliferative inflammatory atrophy
PIN	Prostatic intraepithelial neoplasia
PSA	Prostate specific antigen
PSCA	Prostate stem cell antigen
PTM	Post translational modification
RB-1	Retinoblastoma 1
RNASEL	Ribonuclease- L
ROS	Reactive oxidative species
TCA	Tricarboxylic acid
TMT	Tandem mass tag
TMPRSS2	Transmembrane serine protease 2
TOF	Time of flight
tPSA	Total PSA
TRUS	Transrectal ultrasound
TURP	Transurethral resection of the prostate
SILAC	Stable isotope labelling by amino acids in cell culture
SRD5A	5α-reductase
uPA	Urokinase-type plasminogen activator
uPAR	Urokinase-type plasminogen receptor
VGEF	Vascular endothelial growth factor

Chapter 1

An introduction to prostatic diseases and the role of proteomics in biomarker discovery

1.1 Anatomy of the prostate

The prostate is an ovoid shaped gland located below the bladder, anterior to the rectum and surrounding the urethra (figure 1.1) (1). It is composed of tubuloalveolar glands encapsulated by a layer of connective tissue made up of collagen and a layer composed of abundant smooth muscle which makes up the prostatic stroma (1). The connective tissue and fibrous smooth muscle provide support for glandular epithelial and stromal cells. A double layer of epithelial cells lines the ducts that branch out from the urethra before terminating into acini (3,4). Luminar epithelial cells are cuboidal in shape and lie above flattened basal epithelial cells which is anchored to a basal membrane (2).
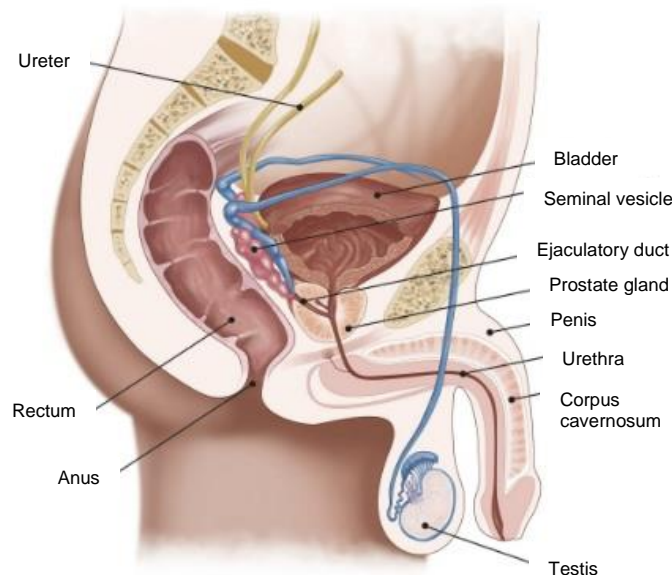
Figure 1.1: Location of the prostate relative to the genitourinary system (Modified from http://intimatehealthhelp.net).

The prostate is divided into four regions, namely the anterior fibromuscular stroma, central zone, transition zone and peripheral zone (figure 1.2). The peripheral zone contributes 70% of the glandular prostate and spans the lateral and dorsal section (4). Most prostate cancers originate in the peripheral zone and it is from this site that they typically metastasize to the lymph nodes or bones (1). The transition

zone contributes 5-10% of the glandular prostate and is located lateral and anterior to the urethra (5,6). It is separated from the peripheral zone by dense fibromuscular stroma. The transition zone is the site from which benign prostatic hyperplasia (BPH) develops and progresses (1). The central zone represents 25% of the glandular prostate and surrounds the ejaculatory ducts (5). The anterior fibromuscular stroma is the non-glandular part of the prostate comprising of dense fibromuscular tissue (5,6). It represents approximately one third of the prostate extending from the anterior of the urethra before fusing with the bladder muscle and neck of the external sphincter (2,5)

The major function of the prostate gland is the production of seminal fluid which nourishes, protects and aids in the mobility of sperm (1). During ejaculation the prostate dispenses a slightly acidic fluid which contributes approximately 30% of the total ejaculate volume (4). Prostatic fluid contains hormones, ions, citrate, lipids and growth factors which preserve and promote sperm viability. The remaining 70% of
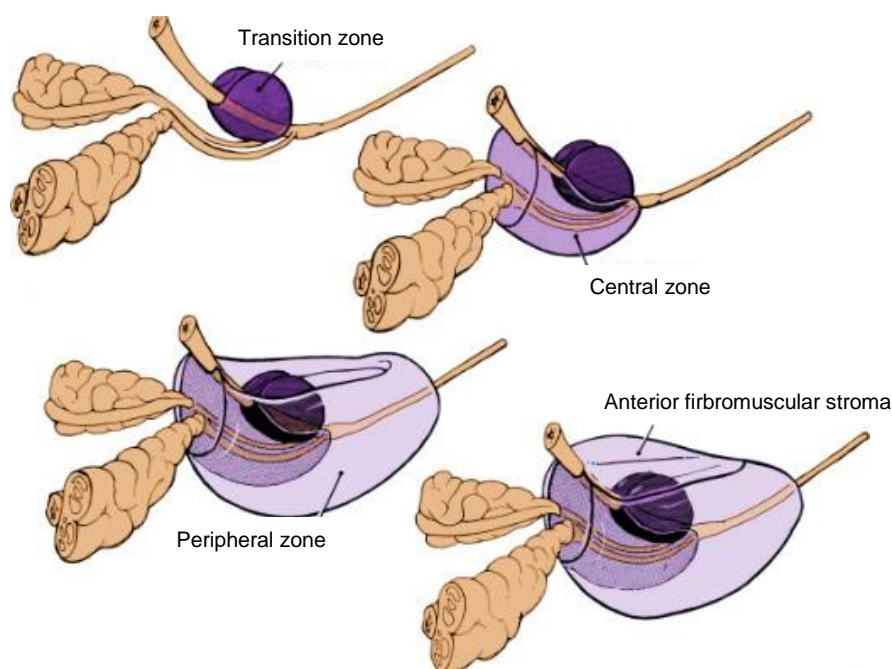
Figure 1.2: Zonal anatomy of the prostate (Modified from http://www.aboutcancer.com).

seminal fluid is derived from seminal vesicles, testicles and bulbourethral glands (1). The prostatic fraction of the ejaculate is also characterized by the presence of two enzymes, namely prostatic acid phosphatase (PAP) and kallikrein-3 (KLK3), which is more commonly known as prostate specific antigen (PSA). PAP is a tyrosine phosphatase implicated in the regulation of PCa cell proliferation (6). PSA is a serine protease which cleaves seminogelin in the seminal coagulum. This serves to liquefy the seminal coagulum and hence increase motility of sperm (2). Disturbed prostate function caused by malignant or benign tumours results in disruption of the basal cell layer and basement membrane causing leakage of PSA into general circulation. PSA can therefore be utilized as a serological marker for prostate cancer (PCa) screening or for monitoring the progress of treatment (5).

Androgens, which are steroid hormones, and growth factors are essential for the optimal functioning, growth and differentiation of the prostate (7). Testosterone is the principal circulatory androgen and is produced by the Leydig cells of the testes. In the prostate testosterone is converted to 5α-dihydrotestosterone (DHT) by the enzyme steroid 5α-reductase (SRD5A) (1). DHT is the most potent natural androgen and readily binds to an intracellular androgen receptor (AR) which is present in prostate epithelial and stromal cells (2,9). Activated AR's form homodimers which translocate into the nucleus and subsequently bind to a DNA sequence known as the androgen response element (ARE) (1). Binding of the DHT-AR complex to promoter or enhancer regions of AR regulated genes induces the expression of genes implicated in cell proliferation, the inhibition of apoptosis and differentiation (1).

Androgen action in the prostate effects reciprocal interaction between stromal and epithelial cells to maintain normal growth and function (9). Androgens stimulate the stromal cells to secrete peptide growth factors known as andromedins, which in turn diffuse into the epithelial cells where they induce the expression of genes controlling cell differentiation (7). Concomitantly, androgens also have an inhibitory effect on andromedin induced growth in epithelial cells thereby promoting epithelial differentiation (7). During carcinogenesis the AR-axis and growth factor network is dysregulated resulting in atypical androgen-growth factor interactions. The dysregulation of this network confers a selective advantage to malignant epithelial cells by enhancing cell proliferation and invasion ability with 99% of prostate tumours originating from the prostate epithelium (10).

1.2 Prostate cancer and associated risk factors

PCa is the second most prevalent cancer globally with 1.1 million people being diagnosed in 2012[1]. It is the sixth leading cause of cancer related deaths accounting for 8% of the total cancer mortalities worldwide[1]. PCa is the most prevalent cancer in South African men with 1 in 23 men being diagnosed with the disease[2]. Incidence rates vary substantially worldwide with the highest rates occurring in the western world. The high incidence rate of PCa in western countries coincides with the widespread use of biopsies and routine PSA screening tests which results in higher detection of the disease (12,13). The rate of diagnosis does not, however, correlate to the number of PCa related fatalities owing to the early detection of treatable lower staged cancer (12). Studies have shown that several risk factors increase the chances of developing PCa.

1.2.1 Age

There is a strong correlation between age and PCa incidence and fatality with over 90% of fatalities occurring in men over the age of 65 (12). While men under the age of 50 have only a 0.3% risk of developing PCa, the risk increases seven-fold in the next decade and rises sharply to 11.2% for men aged over 70. PCa progresses slowly and is preceded by asymptomatic prostate neoplasia which is often incidentally revealed during an autopsy. It is predicted that in most cases these localized and

---

[1]http // www.cdc.gov/cancer/international/statistics
[2]http // za.movember.com/mens-health/prostate-cancer

microscopic sized neoplastic lesions would only have become clinically relevant if men lived beyond 100 years of age (13).

1.2.2 Race and ethnicity

Population based studies have revealed trends between the incidence of PCa and different geographical locations. Approximately 75% of new PCa cases are reported in developed nations with North America having the highest incidence rate followed by Scandinavian countries. The lowest incidence rate occurs in Mediterranean countries and Japan (12). Incidence of PCa also differs according to ethnicity within different countries. In the USA for example the incidence of PCa is 1.6 times higher in African-Americans than Caucasian men (12). Migratory studies have also shown an elevated incidence rate for Japanese men who had immigrated to Western countries indicating a link between PCa and lifestyle factors. Disparities between PCa incidence and different racial groups may be attributed to differences in access to healthcare and genetic factors (2).

1.2.3 Familial associations

A family history of PCa is a strong indicator of susceptibility to the disease. There is a 2-3 fold likelihood of developing PCa for men with first degree relatives diagnosed with the disease (13). Segregation studies have been employed to distinguish environmental factors from genetic factors implicated in familial PCa (1). The segregation analysis indicated that Mendelian inheritance of the RNASEL gene attributed to cases of familial early onset PCa (1). Expression of the RNASEL gene is pervasive as it is implicated in 88% of PCa cases for carriers aged over 85 (1).

1.2.4 Dietary factors

Several dietary components including fat, dairy products and red meat have been linked to PCa risk (14). High fat intake, especially polyunsaturated fats, have been shown to induce the increased production of reactive oxidative species (ROS) by various mechanisms (15). For example, a study of fatty acid metabolism revealed α-methyl-CoA remarcase (AMACR) is over expressed in PCa tumours when compared to normal prostate tissue. AMACR is a peroxisomal enzyme involved in β-oxidation of branched fatty acids and which also produces hydrogen peroxide, a highly oxidative carcinogen, as a by-product. Dairy products are rich in dietary branched fatty acids and this may account for their link to PCa risk. Red meat is considered a high risk dietary component due to the production of carcinogenic heterocyclic amines generated during cooking. Heterocyclic amines are potent carcinogens also implicated in bladder and colorectal cancer (14).

Certain dietary factors play a protective role resulting in reduced risk and incidence of PCa. Vitamin E (α-tocopherol) possesses anti-oxidative properties and inhibits tumour cell growth by inducing apoptosis (15). The hormonal form of vitamin D (1-25-dihydroxyvitamin D) reduces the proliferative capacity and differentiation of PCa cells (15). Consumption of tomatoes which contain lycopene reduces the risk of PCa by 16%. Lycopene is a carotenoid with powerful anti-oxidative properties and has been reported to reduce serum PSA in men with localized PCa (14). The trace element selenium shows the most potential as a protective dietary element in PCa. Selenium inhibits tumorigenesis by boosting the

immune system, inducing apoptosis, scavenging reactive oxidative species and modulating testosterone levels (14).

1.3 Pathobiology of prostate cancer

Carcinogenesis in prostate tissue is the result of an accumulation of genetic alterations that confer unlimited proliferative ability and reduce apoptosis in affected cells. A multitude of molecular pathways have been implicated in PCa initiation and progression. Alterations in tumour suppressors, proto-oncogenes and inflammatory pathway genes negatively impacts on the regulation of cell proliferation, growth and apoptosis.

Phenotypic changes associated with PCa initiation are well characterized and indicate that 95% of prostate cancers are adenocarcinomas. Adenocarcinomas typically arise in epithelial cells of glandular tissue and acini (16).

Neoplastic transformation of the epithelial lining of secretory ducts is defined as prostatic intraepithelial neoplasia (PIN). PIN can be classified into low grade or high grade PIN based on cytological characteristics (17). High grade PIN lesions are considered a precursor to malignant PCa as they are co-localized in the peripheral zone with adenocarcinomas (18). Additionally, high grade PIN shares similar atypical genetic alterations and aberrant molecular pathways also observed in adenocarcinomas (19). Individuals with PIN are likely to develop PCa within ten years of detection (16).

Inflammation is a physiological response to tissue damage that facilitates and initiates the healing process. However, a chronic inflammatory state particularly in prostate tissue is thought to play a role in PCa development and progression (19). Immune system cells recruited during the inflammatory release highly reactive molecules such as nitic oxide and hydrogen peroxide (19). These reactive chemical species cause damage to the DNA of proliferating epithelial cells resulting in mutagenesis (19). Recent studies have identified several genes involved in the inflammatory pathway linked to PCa, namely ribonuclease L (RNASEL), macrophage scavenger receptor-1 (MSR1), interleukin-8 (IL-8) and vascular endothelial growth factor gene (VGEF) (20). The RNASEL gene encodes a ribonuclease that triggers apoptosis on the onset of a viral infection (2). This pro-apoptotic function of RNASEL qualifies it as a prostate tumour suppressor gene and its defective form is hence linked to PCa (20). The MSR-1 gene encodes a transmembrane receptor expressed by macrophages that binds several ligands such as gram-negative and gram-positive bacteria, lipoproteins and polynucleotides (21). MSR-1 is commonly found at sites of inflammation in the prostate. IL-8 is a cytokine that initiates and amplifies the inflammatory response. It has been suggested that the multitude of upstream signalling pathways activated by IL-8 promote malignant cell proliferation, metastasis and angiogenesis in endothelial cells (22). The VGEF gene encodes a growth factor that is considerably elevated in the plasma of PCa patients. VGEF possesses angiogenic properties and is known to promote endothelial cell proliferation, vasculogenesis and cell migration (20). Taken together, inflammation is a complex process that is effected by a multitude of genes, many of which may contribute to PCa.

Tumour suppressors are molecular regulators of cell proliferation that inhibit malignant cell growth. Inactivation, deletion or suppression of tumour suppressor genes therefore promotes tumour cell

proliferation. Several tumour suppressor genes have been identified which when defective contribute to PCa, these include retinoblastoma-1 (RB-1), CDKN2A and PTEN genes (23). The RB-1 gene inhibits tumorigenesis by arresting the progression of the cell cycle. Suppression of RB-1 has been linked to abnormal AR activity, a major contributing factor to castration resistant prostate cancer (CRPC) (24). Deletions in the CDKN2A gene are common in several types of cancer and this can be attributed to the prominent role it serves as a negative regulator of the cell cycle (23). Mutation of the PTEN gene, specifically mono allelic gene loss occur in 60% of localized PCa while homozygous deletions are common in CRPC and metastatic PCa (25). PTEN is therefore the most common defective tumour suppressor gene linked with PCa. Inactivation of the PTEN gene results in accumulation of PIP3, a major substrate for the Akt pathway resulting in evasion of apoptosis and activation of cell survival mechanisms (18). The PTEN regulated pathway has therefore become a target of therapeutic agents in advanced PCa (23). Another gene commonly silenced during PCa is GSTP1 which encodes the enzyme glutathione S-transferase. Hypermethylation of the GSTP promoter region nullifies gene expression (23). This epigenetic event has been observed in 90-95% off cancer lesions and 70% of high grade PIN lesions (19). GSTP is a carcinogen detoxifying enzyme that removes electrophiles and oxidants that cause DNA damage (19). Inactivation of this enzyme increases the susceptibility of DNA in prostate tissue to damage caused by ROS resulting in accumulation of mutations associated with cancer initiation and progression (23). Methylation of the GSTP1 promoter region is now commonly used in as a biomarker during the diagnosis of PCa (23).

Proto-oncogenes encode proteins that control cell proliferation and apoptosis, and as such alteration in these genes initiate and contribute to the amplification of adenocarcinomas (26). Point mutations, gene amplifications and chromosomal translocations are three mechanisms by which proto-oncogenes are converted to oncogenes. The first mechanism results in a gain of function which sees the proto-oncogenes being constitutively activated. The remaining mechanisms result in the overexpression of the encoded oncogenes. The consequence of these alterations is continuous signalling which induces uninhibited cell proliferation. Some of the oncogenes associated with PCa include ERB-2, BLC-2, and TMPRSS2-ERG (23). ERB2 encodes a transmembrane tyrosine kinase belonging to the EGRF family and is involved in cell growth and differentiation. The deletion of the ligand binding domain results in constitutive activation of the receptor in the absence of ligand (26). Studies in PCa cell lines have revealed that ERB2 is implicated in the activation of the AR during CRPC via an MAPK and c-Jun regulated pathway (27). Additionally, ERB2 plays a crucial role in bone metastasis which is associated with a poor prognosis for PCa patients (27). BCL-2 is an anti-apoptotic factor that promotes neoplastic cell proliferation by inhibiting the cellular stress response pathway (26). BLC-2 encodes a cytosolic protein that is localized to the mitochondria where it inhibits the activity of caspases that induce cell death. The overexpressed of BLC-2 in PCa therefore allows malignant cells to by-pass apoptosis. (28). Continual chromosomal translocation occurs during prostate cancer and often results in the TMPRSS2 gene fusing with the ERG gene (27,28). This genomic rearrangement concomitantly results in the activation of ERG which belongs to the ETS family of transcription factors (26). Activated ETS related transcription factors are involved in promoting proliferation, inhibiting apoptosis, expression of

oncogenes and cell migration (27). Therefore, gene fusion of TMPRSS2 and ERG in prostate tissue promotes malignant transformation.

The AR is a steroid hormone nuclear receptor that plays a pivotal role in PCa initiation and development (29). In the normal prostate gland the AR binds its native ligands DHT and testosterone which promote growth and normal function (33,34). Studies have revealed that 80-90% of PCa is initiated by androgen stimulated signals that promote growth and evasion of apoptosis (34,35). Disruption of the AR pathway is therefore the primary endocrine therapy used for PCa treatment (29). This is accomplished by reducing the production of testosterone by the testes and/or the inhibition of the AR (30). Androgen suppression is initially achieved by surgical or chemical castration (androgen deprivation therapy) (29). This is followed by anti-androgen therapy which makes use of non-steroidal AR antagonists such as flutamide or bicaluatamide to inhibit the binding of androgens to the AR (31). Both treatment regimens are often implemented simultaneously and usually result in remission of the cancer. However, after 1-3 years the cancer often returns and is then known as CRPC (18). Despite the depletion of serum testosterone the AR is thought to mediate tumour progression through several alternative mechanisms. First, amplification of the AR gene is thought to increase sensitivity of the AR to low androgen levels. Second, mutation of the AR gene may promote promiscuity of the AR such that it is activated by anti-androgens or non-steroids. Third, dysregulated cytokine or peptide growth pathways may cause irregular activation of the AR. Fourth, altered expression of AR coactivators may affect AR ligand specificity in such a manner that it is activated by adrenal steroids and non-steroids (22,35). Finally, AR expression and signalling resumes due to changes in the intratumoral expression patterns of key steroidogenic enzymes, which enables CRPC tissue to convert adrenal androgen precursors such as DHEA and androstenedione, which have low androgenic activity, to more potent androgens such as testosterone and DHT (32).

The study of molecular pathways involved in carcinogenesis and tumour progression aids in the identification of targets for anti-cancer drug development. Targeted therapy shows immense potential as a remedy to PCa as it directly inhibits tumour cell growth and survival. Therefore, elucidation of the precise molecular pathobiology of PCa may offer insights on the mechanisms of carcinogenesis and present targets for therapeutic intervention.

1.4 Screening and diagnosis of prostate cancer

Early stage PCa is not associated with any discernible symptoms. These only manifest during the later stages of the disease. It is therefore imperative to diagnose PCa in the early stages when therapeutic intervention increases the chances of a favourable clinical outcome. Screening for PCa is normally carried out in men over the age of 49 years and in men exhibiting symptoms of lower urinary tract symptoms (LUTS) (33). The screening procedure utilizes a PSA blood serum test, palpitation of the prostate to detect physical abnormalities via digital rectal exam (DRE) and a transrectal ultrasound (TRUS) guided biopsy for histopathological staging of prostate tissue (34).

The PSA test measures the serum levels of PSA and is often used in conjunction with the DRE. The use of the PSA test in tandem with the DRE improves the detection rate by 81% (33). A PSA reading

above 4 ng/ml indicates the potential presence of PCa and warrants a subsequent biopsy (34). However, PSA is a poor diagnostic tool as it fails to distinguish malignant tumours from benign tumours thus resulting in overtreatment (35). The lack of specificity of the PSA test results in 30% of patients with PCa having normal PSA readings while 20% of patients without PCa have elevated PSA readings (33). While the PSA test is a biochemical evaluation, the DRE is a physical examination of the prostate gland. During the DRE a gloved finger is inserted into the rectum and the prostate is physically examined for firmness, nodularity, asymmetry and induration (36). Following the preliminary screening, an abnormal PSA test and DRE warrant subsequent histopathological analysis of prostate tissue. A TRUS guided needle is used to extract cells from the transition and peripheral zone of the prostate (2). These prostate biopsy samples are assed for the presence of neoplastic cells and the degree of differentiation on the basis of the Gleason grading system (figure 1.3) (37).

Diagnosis of PCa is based on the preliminary imaging of the prostate by TRUS or MRI, determination of prostate volume and histological analysis of prostate tissue (2). Malignant tumours are hyperechoic, asymmetrical and dilated when compared to healthy prostate tissue (38). These anomalies can be visualized by the use of TRUS or MRI and on the basis of the results a biopsy of abnormal tissue is carried out (2). An ultrasound guided needle biopsy is performed transrectally via a spring loaded needle gun (39). The biopsy samples that are obtained are evaluated for PCa stage using histological analysis with reference to the Gleason grading system. The Gleason grading system is based on the glandular differentiation and growth patterns of malignant cells in prostate stroma as visualized under a low to medium power microscope (38,40). The degree of differentiation is stratified into 5 grades ranging from 1 to 5 which increase from well differentiated to poorly differentiated tissue architecture (3,40). Due to the heterogeneity of PCa the prostate usually contains two or more differentiation patterns. The two most prevalent patterns, primary and secondary differentiation patterns are summed to obtain the Gleason score which ranges from 2 to 10 (3,40). The Gleason score therefore provides a prognosis which is then used to implement appropriate treatment strategies.
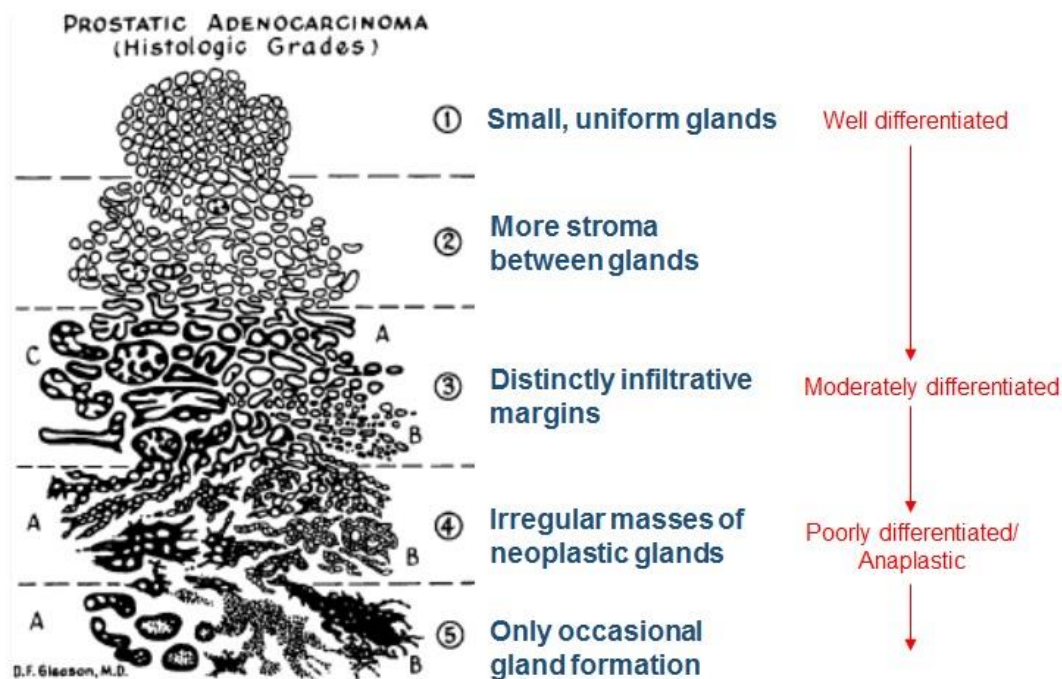
Figure 1.3: Gleason grading histological score showing progressive loss of prostatic glandular differentiation. Modified with permission from Humphrey P., 2004 (37).

1.5 Prostate cancer treatment

Several treatment strategies are used to treat PCa and are dependent on patient age, tumour stage and physical health. Implementation of the correct treatment therapy is crucial, but is challenging due to the highly variable prognosis of PCa.

For localized PCa, watchful waiting, radiation therapy and prostatectomies are the typical routes of treatment (16). A radical prostatectomy involves the surgical removal of the prostate gland (12). Excision of the prostate may be performed trans-urethrally through the tip of the penis or laparoscopically through incisions that are made in the abdomen. The laparoscopic treatment method is relatively less invasive to trans-urethral prostate resection especially when it is performed with robotic guides. Radical prostatectomy poses numerous harmful side effects with a 50% risk of erectile dysfunction, urinary incontinence, infection and excessive bleeding at the site of surgical incisions (12).

Following removal of the primary tumour through prostatectomy radiotherapy is routinely implemented as a follow up treatment and may be administered post-operatively as "adjuvant therapy" or as salvage therapy if PSA levels become elevated again. Studies have shown that early radiation therapy results in more favourable prognostic outcomes, hence adjuvant therapy is superior to salvage therapy (40). Radiotherapy may be further categorized on the basis of radiation delivery method into external beam radiation or internal radiation (brachytherapy) [3]. External beam radiation makes use of x-rays or gamma rays propagated from a linear accelerator to kill tumour cells. This type of radiotherapy is usually administered for 6-10 weeks and may be enhanced by ancillary use of computed tomography imaging for precise targeting of tumours (2). In brachytherapy, small radioactive pellets (^{125}I or ^{103}Pd) are

implanted directly into tumours allowing for high ionization dosage delivery with minimal cytotoxic effects on proximal cells (2). The pellets deliver low level x-ray radiation which incidentally can only travel short distances thereby limiting the cytotoxic ionization energy to the targeted tumour. Both external beam radiation and brachytherapy carry several undesirable side effects such as urinary incontinence, irritation of the rectum, erectile dysfunction and fatigue[3]. Side effects of radiotherapy used as a primary treatment are relatively less severe when compared to radical prostatectomy[3].To avoid the undesirable side effects of both prostatectomies and radiotherapy watchful waiting can be considered as an alternative. Using this strategy, clinical symptoms are closely monitored and therapeutic agents are administered correspondingly (12).

Androgen deprivation therapy is the gold standard for treatment of aggressive metastatic PCa. This therapy drastically reduces the levels of circulating testosterone, by inhibiting testosterone production in the testes by surgical (orchiectomy) or chemical castration, thereby halting the growth and proliferation of tumour cells (12). Luteinizing hormone-releasing hormone (LHRH) agonists are synthetic peptides that inhibit the production of luteinizing hormone (LH) in the pituitary gland. The reduction in LH removes the stimulus for the testes to produce testosterone. LHRH agonists are administered subcutaneously and usually comprise either leuprolide acetate or goserelin acetate (2). Seventy percent of patients initially respond well to this treatment up until the disease progresses to CRPC. Two drugs, abiraterone and enzulatamide have recently been approved for the treatment of metastatic CRPC as both significantly increase life expectancy. Abiraterone inhibits androgen biosynthesis by inhibiting cytochrome P450 17α-hydroxylase/17,20-lyase (CYP17A1). The inhibition of CYP17A1 results in a considerable decrease in the levels of circulating androgens and androgen precursors. Enzulatamide disrupts the AR signalling axis at three junctions: specifically inhibition of nuclear translocation of the AR into the nucleus; blocking the AR binding to the DNA; and inhibiting recruitment of co-activators. Patients with CRPC have a poor prognosis and at this stage treatment is only palliative (12). Chemotherapy, using drugs such as docetaxel and paclitaxel are also used for the treatment for CRPC. These drugs do not preferentially target tumour cells and are therefore only used as a last resort to treat symptoms of terminal CRPC (2).

1.6 Benign Prostatic Hyperplasia

BPH is a urological disease characterized by an increase in cellular proliferation of the stromal and glandular components of the prostate gland (41). This condition is non-malignant, but causes an increase in size of the prostate. BPH is the most common benign tumour in men affecting one third of men between the ages of 50-60 and approximately 90% of men aged 85 and above (42).

1.6.1 Epidemiology of BPH

Several risk factors predispose men to developing BPH, namely age, race, familial history and diet (43). Prevalence of BPH increases with age with approximately 40% of men aged 50 and 90% of men aged

---

[3] http://www.cancer.org/cancer/prostatecancer/detailedguide/prostate-cancer-treating-hormone-therapy

over 80 exhibiting microscopic BPH (43). Microscopic BPH is usually asymptomatic as only 25-50% of cases present clinically relevant (macroscopic) manifestations (44). A study in the USA showed that the incidence of BPH was greatest in African Americans and the lowest in Asian (Chinese and Japanese) immigrants much like in PCa (44). Another case study in Italy showed a link between dietary habits and the risk of BPH (45). Dietary fat derived from dairy products was found to increase risk of BPH while lycopene found in tomatoes seemed to play a protective role, presumably due to its antioxidant properties (45,46,47). Familial history, specifically inherited genetic polymorphisms of the SRD5A2 gene is linked to an increased risk of BPH (43). Single nucleotide polymorphisms of the SRD5A2 gene cause dysregulation of the SRD5A2 enzyme which contributes to the pathological state of the androgen dependent prostate gland observed in BPH (47). More epidemiological studies need to be conducted to gain deeper insights into the hormonal changes and the influence of the environment in BPH risk.

1.6.2 Pathobiology of BPH

The molecular aetiology of BPH is a multifactorial abrogation of pathways related to tissue remodelling of the ageing prostate, hormonal changes and chronic inflammation in the prostate (48). The pathogenesis of this disease is not clearly defined and is therefore attributed to these several factors. Further evidence of the multifaceted pathobiology of BPH is shown by the cytological and morphological variation seen in BPH tissue specimens (43).

Tissue remodelling of cells in the transition zone of the prostate gland is a typical consequence of ageing in males. Basal cells undergo the greatest age related modifications of all cell types in the prostate. Gradual alterations in intracellular metabolism results in the enlargement and cellular hypertrophy of these cells (48). Alteration of prostate epithelial cells sees this cell type adopt a "senescent phenotype" resulting in atypical responses to growth factors and other signalling molecules (43). Furthermore, the aberrant interaction of epithelial cells and complex signalling pathways disrupts epithelial-stromal interactions causing an increase in prostate volume (48).

The dysregulation of peptide growth factor secretion that occurs in BPH stimulates epithelial and stromal cell growth (figure 1.4) (43). Several fibroblast growth factors (FGF) are also known to be overexpressed in BPH where they play a growth stimulatory role. FGFs are produced in response to the localized hypoxic conditions created by increased cell proliferation in the prostate (49). Stromal cells produce FGF-2, -7 and 10 which stimulate epithelial cell growth (43). Epithelial cells in turn produce ILα-1 and IL-8 which stimulates the production of FGF-2 and -7 in stromal cells (44,50). FGF-2 is also know to promote the growth of stromal components. Taken together, two positive paracrine loop promote cellular growth in BPH.

Cell growth in the healthy prostate is tightly controlled by cellular mechanisms that balance apoptosis and cell proliferation. Dysregulation of pathways involved in these two processes is thought to be a causative factor in the development of BPH. The proto-oncogene BCL-2 and the growth factor TGFβ are known to play pivotal roles in the apoptotic pathway in prostate tissue (43). BCL-2 is an inhibitor of apoptosis that is overexpressed in the prostatic epithelium of BPH patients (43). This anti-apoptotic molecule is also overexpressed in PCa and other types of cancer further highlighting its prominent role

in abnormal cell proliferation. TGFβ is a negative growth factor that plays a dual role as an inhibitor of epithelial cell growth and a promoter of stromal cell differentiation (43). Similar to BCL-2, TGFβ is overexpressed in BPH and this may account for the increased smooth muscle tone that is seen in BPH patients (43).

1.6.3 Pathogenesis and clinical presentation

BPH typically originates in the transition zone adjacent to the prostatic urethra (43). The onset of hyperplasia is characterized by the formation of nodules which initially compress prostate tissue (50). Progressive nodular proliferation results in compression of the prostatic urethra which in turn causes bladder outlet obstruction (BOO) (51). Obstruction of the bladder causes urinary voiding problems, bladder hypertrophy and overactive bladder syndrome (42). The clinical manifestation of BOO presents varied symptoms among patients collectively termed lower urinary tract symptoms LUTS (52). Approximately 25% of men suffer from LUTS with symptoms ranging from mild to severe (53) LUTS are characterized by chronic retention, hesitancy, nocturia, dysuria, incomplete voiding and overflow incontinence (42). In addition to BOO, symptoms of LUTS may also be attributed to detrusor over activity and increased prostatic smooth muscle tone (41).

1.6.4 Diagnosis and treatment of BPH

The diagnosis of BPH relies on a physical examination by DRE and subsequent PSA testing if necessary. The DRE determines the prostate size, surface texture and symmetry all of which are evaluated to assess the risk of BPH (42). A PSA test and TRUS guided biopsy are performed to rule out PCa following a DRE which indicates an abnormal prostate. The use of PSA remains controversial as it is elevated in a number of pathological states of the prostate including BPH. A urologist may sanction further tests to confirm enlargement of the prostate namely, uroflowmetry, residual volume measurements and urodynamics (50). Uroflowmetry measures the speed of urination relative to time and is used to determine the presence of an obstruction. The residual volume is measured subsequent to urine voiding via an abdominal ultrasound. It is used to determine the best treatment course with a volume above 300 ml indicating acute urinary retention and necessitating surgery. Urodynamics are employed when an obstruction is not clearly detected and gives predictive information for the best treatment course. Urondynamics measure both pressure and flow rate, and is used to distinguish between bladder obstruction and detrusor instability (50).

Treatment for BPH is used to relieve the symptoms of LUTS, improve the overall quality of life of the patient and prevent complications that may arise from LUTS. Depending on the severity of LUTS treatment may be pharmacological or surgery based (54). Pharmacological treatment is usually prescribed to men presenting mild to moderate symptoms and includes SRD5A inhibitors and α-blockers (54). α-blockers are the first line of treatment for mild LUTS and effect their mode of action by binding the α1 adrenoreceptors in the prostate (54). This relaxes smooth muscle in the prostate and bladder neck thereby improving urine flow (42). The SRD5A inhibitors prevent the conversion of testosterone to DHT by SRD5A (42). This reduces prostate size by 20-30%, improves urine flow and prevents secondary complications of LUTS such as acute urinary retention (54). Surgical treatment is

reserved for patients with severe LUTS or complications thereof such as haematuria, acute urinary retention, renal failure and bladder stones (53). The most commonly used surgical procedure is the transurethral resection of the prostate (TURP). This method sees hyperplastic tissue being excised from the bladder neck and has a success rate of 80% (41). To conclude, BPH is a complex disease and progress is being made to improve diagnosis and provide non-invasive treatment options that improve the quality of life of patients.

1.7 Biomarkers

Biomarkers can be defined as alterations of cellular biochemical processes resulting in quantifiable changes in body tissue, cells or fluids which are unique to pathological or physiological processes (56,57). The myriad of cell types in the body possess unique molecular signatures which are altered during disease states (57). Detailed knowledge of this molecular signature under normal physiological conditions may be key to unlocking the precise mechanism of disease aetiology. Consequently, a biomarker may be used to measure specific parameters that are indicative of an abnormal state (57). Aberrations of the genome caused by somatic or germline mutations may result in an altered phenotype which is reflected in the atypical expression of genes, metabolic pathways and proteomic signatures (58). Therefore, the primary utility of biomarkers is to measure and assess the normal and pathologic state as well as determine the response to therapeutic intervention (57,60) . One such example is the urine test for glucose used to identify patients suffering from diabetes. The dysregulation of metabolic pathways involved in sugar metabolism results in elevated blood-glucose subsequent excretion in the urine where it is detected (60).

A biomarker may also be defined as a specific physical, chemical or biological feature that characterizes processes occurring in living organisms. Biomarkers render insights into the biochemical pathways connected to pathogenesis which allows disease progression or regression to be tracked (56). This definition deviates from the steadfast interpretation of biomarkers as biological entities and incorporates high throughput techniques such as MS and NMR that aid in defining pathogenesis and monitoring its development at the cellular level (55). For this purpose, biomarkers are prone to modulation and therefore increase our knowledge of drug pharmacodynamics (57).

1.7.1 Cancer Biomarkers

Cancer is a multi-faceted disease defined by irreversible alterations of the genomic framework and gene expression. Modifications in three subsets of genes, namely oncogenes, tumour suppressor genes and DNA repair genes consequently results in the cellular machinery by-passing the usual biological regulatory system (61). This distinct genotype bestows a survival advantage to the expressed phenotype due to the dysregulation of specific cell cycle checkpoints (57). The genetic diversity of the disease is reflected in the clinical setting where the pathogenicity and aggressiveness of the disease varies across individuals. Consequently, these variations present a formidable problem in the treatment of the disease especially when it is in its advanced stages (62). A key factor in the successful treatment of cancer lies in early detection of the disease when tumour progression can be halted (59). The progression of a normal cell into a malignant tumour cell is a complex process defined at each stage by aberrant cellular biochemistry. These aberrant molecular signatures may indicate hyperproliferation, gene over expression, abnormal gene expression, inflammation, hyperplasia or atypical enzymatic activity (61). Therefore, by deciphering the complex molecular signature of malignant transformation, biomarkers can be utilized as early indicators of cancer or risk thereof (59).

A cancer biomarker refers to a quantifiable biological molecule localized in body tissue or fluids that serves as an indicator of a malignant tumour. These biomarkers are produced either by the tumour or

host tissue in response to the tumour (57). Cancer biomarkers play a pivotal role in the early detection, screening, prognosis, disease staging and prediction of therapeutic intervention (60,62). They may be measured either quantitatively or qualitatively, and span many different forms such as nucleic acids, proteins, metabolites, or cellular events associated with proliferation or angiogenesis (3,63). These biomarkers may be present in the blood, cerebrospinal fluid, urine, saliva, or tissue samples (56,62).

The first recognized cancer biomarker was reported in 1965 (Table 1.1) by Dr Joseph Gold based on his investigation of tumour specific antigens associated with colon cancer. He identified a protein in the blood of colon cancer patients normally found in fetal tissue and accordingly named it carcinoembryonic antigen (CEA) (64,65). Gold's research laid the foundation for advancement in this field with the 1970's being notable for a variety of serum tests developed for numerous types of cancers. The 1980's saw even more remarkable progress with the introduction of biomarkers for colorectal, pancreatic cancer (CA19-9), ovarian cancer (CA-125) and breast cancer (CA15-3) (64). At present the most well-known biomarker, commonly utilized for prostate cancer screening tests is the prostate specific antigen (PSA). PSA was first characterized in 1979 and later approved for monitoring PCa in 1986 by the United States Food and Drug Administration (FDA) (66,67). The PSA blood test for monitoring PCa progression provided a major boost for treatment of PCa patients. The greatest utility of PSA as a biomarker however was discovered after a study by Catalona *et al.* reported elevated PSA serum levels in asymptomatic men (67,68). Following this study PSA was approved as a PCa screening tool in 1994 by the FDA and has since significantly reduced PCa induced mortality (67). The drawback of PSA as with most of the fore mentioned cancer biomarkers is its lack of specificity which causes misdiagnosis. In brief, these biomarkers serve as indicators of cancer since they are present at the conventional basal level in normal individuals and significantly elevated in cancer sufferers. Limited progress has been made in the development and implementation of new cancer biomarkers despite recent advances in technology (64).

15

Table 1.1 Cancer biomarkers that are in current use. DRE, digital rectal examination; ER, estrogen receptor; NACB, National Academy of Clinical Biochemistry; NSGCT, nonseminomatous germ cell; PgR, progesterone receptor. Reproduced from Kulasingam *et al.*, 2007 (69).

Tumor Marker	Cancer Type	Year of Discovery	Application based on ASCO and /or NACB recommendations
α-fetoprotein	Germ-cell hepatoma	1963	Diagnosis Differential diagnosis of NSGCT Staging
Calcitonin	Medullary thyroid carcinoma	1970s	Diagnosis Monitoring therapy
CA125	Ovarian	1981	Prognosis Detecting recurrence Monitoring therapy
CA 15-3	Breast	1984-5	Monitoring therapy
CA 19-9	Pancreatic	1979	Monitoring therapy
Carcinoembryonic antigen	Colon	1965	Monitoring therapy Prognosis Detecting recurrence Screening for metastases
ER and PgR	Breast	1970s	Select patients for endocrine therapy
HER2	Breast	1985-6	Select patients for trastuzumab therapy
Lactate dehydrogenase	Germ cell	1954	Diagnosis Prognosis Detecting recurrence Monitoring therapy
Prostate specific antigen	Prostate	1979	Screening (with DRE) Diagnosis (with DRE)

1.7.2 Classification of cancer biomarkers

Different criteria has been employed to classify cancer biomarkers but no clear-cut standardized system has yet been devised. The absence of a conventional classification system can be attributed to the plethora of new information coming from the biomedical sciences field coupled to advances in technology (59). A simple yet effective method of biomarker classification places emphasis on monitoring the progression of the disease (70). This classification system delineates the continuum of events that characterize cancer beginning from primary exposure to the advanced metastatic disease (figure 1.4). Additionally, this approach may aid clinical processes such as early diagnosis, prognosis and prediction of drug efficacy (55). In the clinical setting however, tumour biomarkers will frequently overlap in such a manner that they cannot be exclusively designated to a single classification strata (59). Therefore, tumour marker classification in this context relies on close observation of the disease from the earliest phase up to advanced metastatic cancer (55). Cancer biomarkers classified according to the characterization of disease progression can be classified into screening, diagnostic, prognostic, predictive, pharmacodynamic and surrogate end-point biomarkers.
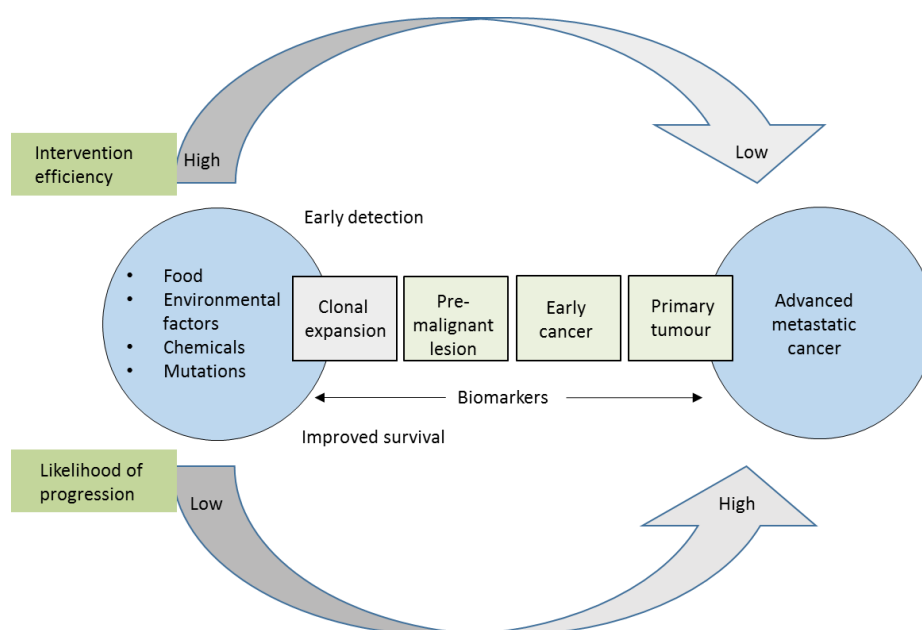
Figure 1.4: Cancer initiation and progression. Modified with permission from Srinivas *et al.*, 2001 (61).

Screening (Early Detection) Biomarkers

Screening biomarkers are used to identify individuals at risk of developing cancer in an asymptomatic population. They essentially serve as an early warning system for individuals at high risk of developing cancer. Caution must be exercised before the sanctioned use of these biomarkers as the benefit of early detection must outweigh the risk of erroneous diagnosis (71).

Diagnostic biomarkers

The ideal diagnostic biomarker possesses a high level of sensitivity and specificity (71). According to Maruvada *et al.* sensitivity is defined as "the true-positive test results expressed as a percentage of all tested individuals who have that disease" while specificity is defined as "true-negative results expressed as a percentage of all tested individuals who do not have that disease" (72).

Prognostic biomarkers

Prognostic biomarkers are utilized when clinical manifestation of cancer have been established (63). These biomarkers aid in predicting the likely course of neoplastic progression, recurrence and therefore assist oncologists in selecting appropriate treatment methods (60,63). They may also be selected based on histopathological grading of tumours which enables differentiation of malignant tumours from benign tumours (59).

Predictive/stratification biomarkers

Predictive biomarkers are utilized for evaluating the effectiveness of a particular treatment for an individual (60,63). They are typically derived from the analysis of the gene expression profile of the specific cancer type (63). Knowledge of the molecular profile of the specific tumour makes it possible to predict the outcome of the chemotherapeutic drugs that will be administered (59).

Pharmacodynamic biomarkers

Pharmacodynamic biomarkers can predict the effect of a drug on a biological system. These molecular markers can be used to evaluate the relationship between treatment modality, target response and closely related molecular pathways that may be affected. Single-molecule real-time sequencing, ion torrent sequencing and nanopore sequencing (discussed in section 1.8.1) are tools used to elucidate pharmacodynamic biomarkers. Pharmacodynamic biomarkers can be used in oncology to identify effective drug combinations and schedule for different treatment regimens to be implemented (73).

Surrogate end-point biomarkers

Surrogate end-point biomarkers are useful for the observation of patient response to a particular type of treatment (71). This type of biomarker provides a holistic outlook of disease progression which can be used to predict the end-result of the chosen treatment. Indeed surrogate end-point biomarkers serve as interim clinical endpoints and their predictions determine the potential benefit or harm the proposed treatment (74).

1.7.3 Current and potential biomarkers for prostate cancer

The process of malignant transformation in prostate cells is characterized by aberrant genomic alterations that confer unlimited proliferative capacity, angiogenesis, evasion of apoptosis and metastatic capability as discussed in section 1.3 (75). Pathways involved in these oncogenic events together with proliferation signalling molecules have been identified as key targets for biomarker mining. The implementation of a novel biomarker at the clinical level has thus far been confounded due to the heterogeneity of PCa caused by the multitude of genetic alterations (51). As a result there is a growing consensus to identify a panel of biomarkers that can be used to detect early stage PCa and inhibit its progression to late stage CRPC (75). At present potential molecular markers are primarily being

investigated in serum and tissue. Molecular signatures unique to tumour cells are also being investigated (51). The utility of these emerging PCa biomarkers include screening, staging, assessing metastatic potential and prognosis (75).

1.7.3.1 Serum biomarkers

Serum biomarkers are molecules produced by tumour cells or proximal cells in the tumour microenvironment (76). These molecules are released into circulation and can be measured using diagnostic assays. Serum biomarkers are gaining widespread utility in the clinical setting as they can be obtained economically and in a non-invasive manner. Currently the serum biomarker PSA is used for screening and staging PCa (75). It has however come under criticism as it is elevated in serum during benign prostate pathologies such as BPH and acute prostatitis (77). Therefore new biomarkers and PSA derivatives have been proposed as alternatives to increase specificity and distinguish between aggressive cancer and indolent cancer (77).

1.7.3.1.1 Prostate Specific Antigen

The measurement of total prostate specific antigen (tPSA) in blood serum is the standard clinical practice for screening, monitoring and predicating outcome of therapy in PCa patients (78). The concentration of tPSA in serum is determined by the total number of cells producing PSA, expression of the PSA gene and the extent of leakage into circulation from disrupted cell architecture (78). Screening tests for PCa that employ tPSA serum measurement have a positive predicative value as low as 25% for patients with readings between 4-10 ng/ml (79). The relationship between tPSA and PCa remains unclear as evidenced by its poor predictive value for disease screening. Total PSA offers greater value as a prognostic tool and is often employed to predict outcome of radical prostatectomy such as tumour volume and disease stage (80). Total PSA can also be used to monitor patients on hormone therapy for late stage PCa. After a 7 month treatment period serum tPSA falls to a nadir level that can be interpreted to give an accurate prognosis (78). A tPSA reading above 4 ng/ml is associated with a poor clinical outcome whereas a reading below 0.2 ng/ml indicates a favourable prognosis[4].

The lack of specificity of PSA as a biomarker has led to the investigation of its derivatives as alternatives, these include free-to-total PSA ratio, PSA density, PSA velocity and PSA isoforms (75). Free-to-total PSA ratio can be used to enhance specificity for patients with a tPSA reading between 4-10 ng/ml (75). A relatively lower free-to-total PSA ratio is associated with PCa; a ratio between 0-10% carries 55% risk of PCa while a ratio greater than 25% carries only a 8% risk of PCa (78). PSA density is determined by dividing serum tPSA by prostate volume and is thought to aid in differentiating early stage PCa from BPH (74,77). The amount of PSA released into serum in PCa is relatively higher when compared to PSA serum levels associated with BPH (75). Therefore, a high PSA density is associated with PCa while a lower PSA density is indicative of BPH. PSA velocity is a measure of serum PSA over a given time period and is reflective of the physiological state of the prostate (78). A PSA velocity greater than

---

[4] https://www.auanet.org/education/guidelines/prostate-specific-antigen.cfm

0.75 ng/ml/year increases the likelihood of diagnosing PCa while it is in its early stages (78). Recent investigations have identified PSA isoforms, namely BPH specific PSA (BPSA) and pro-PSA as promising biomarkers with an enhanced ability to differentiate PCa from BPH (75).

1.7.3.1.2 Human Kallikrein-Related Peptidase 2

Human Kallikrein-Related Peptidase 2 (KLK-2) is a serine protease secreted by the prostate gland that functions to convert the precursor form of PSA to active PSA (81). KLK-2 shares an 80% sequence homology with PSA and is also present in serum at 2% the concentration of PSA (81). Several studies have reported tissue levels of KLK-2 being elevated in PCa and it may thus serve as a biomarker. Serum studies have reported the diagnostic utility of KLK-2 when used in combination with free PSA and tPSA for disease staging, extracapsular extension and prognosis for patients who have undergone radical prostatectomy (63,64). Measurement of complexed KLK-2 (KLK-2-P16) in the serum can improve specificity as it has been shown to be associated with prostate tumour tissue (66).

1.7.3.1.3 Kallikrein-4 (KLK-4)

KLK-4 is an androgen-regulated serine protease highly expressed in the prostate linked to PCa progression (75). Antibodies of KLK-4 are elevated in the serum of PCa patients while tumour tissue levels of KLK-4 are relatively higher than in normal prostate tissue (75). The overexpression of KLK-4 in the prostate during PCa is thought to mediate the transition of cell morphology from epithelial to mesenchymal (EMT) conformation (58,66). This transition is caused by KLK-4 mediated reduction in E-cadherin expression and increased vimentin expression which is a prominent EMT event (75). Additionally, PC3 cells that overexpressed KLK-4 showed a considerable increase in cell proliferation, motility and colony formation (82). In a study by Gao *et al.* it was discovered that KLK-4 facilitates metastasis by mediating interactions between osteoblasts and tumour cells (82). As the various roles of KLK-4 in PCa progression continue to be elucidated, this protease shows immense potential as a biomarker for the diagnosis of PCa.

1.7.3.1.4 E-cadherin

E-cadherin is a glycoprotein localized at cell junctions that maintain cell-cell adhesion and a connection to the encapsulating extracellular matrix (ECM) (74,82). It plays an indirect role in maintaining cell quiescence by inhibiting anoikis, cell death caused by detachment of cells from other cells or the ECM (75). Serum E-cadherin levels of advanced stage metastatic PCa patients are typically elevated since it is detached from cell junctions and free to enter circulation. Another molecular event associated with downregulation of E-cadherin is the switching of cadherin type expression to N-cadherin. N-cadherin is associated with metastasis, increased invasion capacity and increased mortality risk (75). E-cadherin therefore shows potential as a diagnostic PCa tool and a target for therapy in metastatic PCa.

1.7.3.1.5 Minichromosome maintenance protein (MCM2-7)

Minichromosome maintenance (MCM) proteins, 2-7 are involved in the initiation of DNA replication and maintenance of genomic stability in replicating cells (84). Expression of MCM proteins is linked to the cell cycle with upregulation in all phases of the cell cycle except for the G0 phase when they are down regulated (84). The crucial role of these proteins in DNA proliferation inevitably links them to

carcinogenesis as is seen by the overexpression of MCM-5 in PCa tissue and elevation in the urine sediments of PCa patients (58,68). As a consequence of its overexpression in PCa tissue MCM-5 serves as a prognostic tool for patients receiving radiotherapy, hormonal therapy and radical prostatectomy (75). MCM-5 levels are considerably reduced in the prostate tissue of BPH patients making it a promising diagnostic tool for differentiating between the benign and aggressive disease state. Another protein in the MCM family useful for diagnostic purposes is MCM-7 which can differentiate BPH from low PIN and aggressive cancer (75). Further research must be conducted to evaluate the clinical utility of these proteins as biomarkers.

1.7.3.1.6 Early Prostate Cancer Antigen (EPCA-2)

The nuclear matrix consists of a structural framework that maintains the spatial arrangement of components responsible for nucleotide synthesis, steroid transport and nuclear organization (85). Alteration of the nuclear matrix proteins is linked with carcinogenesis, most notably early prostate cancer antigen (EPCA-2) protein (79). EPAC-2 is a nuclear matrix protein that is exclusively expressed in prostate tissue and elevated in the serum of PCa patients (75). This nuclear matrix protein was first identified in a study that profiled the proteome of rat prostates. Following this study, a serum based assay with 94% specificity and 92% sensitivity was developed (74,79). More recent studies have focused on developing anti-bodies against epitopes of EPCA-2 that are simple to detect in serum and increase specificity (75). Two epitopes, EPCA-2.19 and EPCA-2.22 have shown promise and could potentially be used in tandem to diagnose and stage PCa (75).

1.7.3.1.7 Interleukin-6 (IL-6)/Interleukin-6 Receptor (IL-6R)

Interleukin-6 (IL-6) is a cytokine which stimulates the growth of prostate tumours (86). This cytokine's receptor exists in two forms, the membrane bound mIL-6R and the soluble sIL-6R (75). Studies have reported the elevation of IL-6 and sIL-6R in the serum of PCa patients (86). Elevated serum levels of IL-6 and sIL-6R are linked to tumour aggressiveness, increased tumour volume, metastasis and poor survival outcome (76,89). An *in vitro* investigation that studied the effect of IL-6 and sIL-6 on prostate cells by Santer *et al.* reported increased cell motility, migration and reduced adhesion (75). IL-6 and sIL-6R have shown potential as biomarkers for PCa and have already been incorporated in a panel of biomarkers as prognostic tools for radical prostatectomy (75).

1.7.3.1.8 Prostate Cancer Gene 3 (PCA3)

The prostate cancer antigen gene (PCA3) was first reported by Bussemakers *et al.* as a non-coding RNA that is overexpressed in 95% of PCa cases (88). PCA3 is upregulated approximately 66-fold in PCa tissue relative to normal prostate tissue and can thus serve as a viable biomarker (78). Its reliability as a biomarker is evidenced by a commercial testing kit, APTIMA® PCA3 for measuring urine levels of PCA3 (80). Also, a comparative study of PCA3 and PSA as first line screening tools for PCa showed that PCA3 was better suited to identifying early stage PCa than PSA (75). For these reasons PCA3 is likely to become an important diagnostic tool for PCa detection.

1.7.3.2 Tissue biomarkers

Prostate biopsies are a rich source of biomarkers and have recently started finding clinical utility due to advancements in immunohistochemistry, tumour microarrays and imaging technologies (89). Immunohistochemical staining antibodies bind biomarkers in solid tumour samples uniformly while tumour microarrays facilitate high throughput analysis of multiple samples (89). The combination of high throughput analysis and uniform staining enables the analysis of multiple biomarkers on tumour samples. Automated image analysis of the stained specimens enables accurate and precise quantitation of biomarkers in solid tumours (89). Expression of urokinase-type plasminogen activator, prostate stem cell antigen (PSCA) and α-methyl-co-racemase (AMACR) correlate with PCa and are being investigated as potential tissue biomarkers (75).

1.7.3.2.1 Urokinase-type plasminogen activator (uPA)

Urokinase-type plasminogen activator (uPA) is a serine protease that degrades protein components of the extracellular matrix and basement membrane (74,89). This protease is a component of the uPA complex which comprises its receptor (uPAR), plasminogen and plasminogen activation inhibitors (PA-1, PA-2) (90). Urokinase-type plasminogen activator exerts its effects by binding to its receptor which activates a signalling cascade that initiates membrane proteolysis via broad spectrum proteases (75). The activation of these proteases in tumour cells is associated with metastasis and angiogenesis. Immunohistochemical tissue analysis of PCa cells, lymph node metastases, and tumour associated stroma cells revealed overexpression of uPA, uPAR and its inhibitor PA-1 (91). In addition to its inhibitory role, PA-1 was also linked to tumour related processes such as migration, cell adhesion, proliferation and chemotaxis (91). A recent study by Skovgaard *et al.* described the imaging of malignant prostate tissue by positron emission tomography (PET) to generate a quantitative expression profile of the uPAR system (91). Urokinase-type plasminogen activator can therefore be used as a tissue biomarker for metastatic PCa and an inhibitory target for PCa treatment.

1.7.3.2.2 Prostate Stem Cell Antigen (PSCA)

Prostate stem cell antigen (PSCA) is a membrane glycoprotein differentially expressed in prostate tissue (80). The function of PSCA in normal physiological processes or carcinogenesis is unknown but its expression is elevated in the prostate of PCa patients. PSCA is expressed in 100% of metastasis samples and 94% of primary tumour samples (75). The RNA of PSCA has also been detected in the serum of PCa patients. The differential expression of PSCA in the prostate makes it an ideal biomarker for disease screening and a staging tool to determine the aggressiveness of PCa (80).

1.7.3.2.3 α-methyl-CoA-racemase (AMACR)

α-methyl-CoA-racemase is an enzyme involved in the peroxisomal β-oxidation of branched chain fatty acids (79). This enzyme is highly expressed in malignant tumour tissue. Immunohistochemical staining of tumour tissue for AMACR is able to differentiate PCa from BPH with 97% diagnostic sensitivity and 92% specificity. AMACR also possesses prognostic value with expression levels being used to determine PCa recurrence and survival outcome. This enzyme is has already gained clinical utility as a diagnostic tool in combination with other biomarkers (80).

1.7.3.3 Molecular signature biomarkers

The molecular signature of a cancer cell is a characteristic protein or nucleic acid expression profile that distinguishes the tumour cell from normal proximal stromal or epithelial cells. High throughput analysis of molecular signatures offers insights into classifying cancer heterogeneity, staging, and aggressiveness. Delineation of molecular pathways involved in disease progression can provide targets for therapy and prediction the outcome of therapeutic intervention. With improvements in technology disease patterns can be profiled in individuals to increase diagnostic accuracy and administer personalized treatment regimens for patients. The desired outcome of selective therapy is accurate disease staging and increased patient survival (61,74).

1.8 Strategies for discovering novel biomarkers

The advent of high-throughput technologies aka "omic" technologies has enabled the extensive analysis of the molecular biology of the human body. Biomarker discovery is one such field of molecular biology that makes use of omics. The high throughput techniques used in these biomarker studies determine the data quality and experimental limits of the investigation (92). The powerful analytical capabilities of omics instrumentation has led to the widespread use of this technology in cancer biomarker research (93). Traditionally the approach used for cancer biomarker research relied principally on genomics and ancillary computational tools (92). However, due to rapid advances in high-throughput technology and a need for a more holistic analysis of the physiological microenvironment other techniques are now implemented to augment genomics data (92). Strategies commonly used to identify cancer biomarkers include genomics, proteomics, metabolomics and glycomics. These approaches all share the same general statistical principles beginning with preliminary data analysis of samples to reduce experimental noise followed by probabilistic calculations to evaluate the link between measured molecular dynamics to the disease state (93).Taking it altogether, interpretation of the data generated by omic technologies in the drive to find cancer biomarkers can be used to delineate mechanisms of tumour pathology and ultimately find effective therapeutic remedies.

1.8.1 Genomics

Completion of the Human Genome Project in 2001 ushered in new innovations in molecular biology that provided insights into cancer pathobiology (94). The clinical and biological features of the cancer phenotype are directly linked to alterations in the genome. Therefore, characterization of the tumour genome can be used to reveal and compile the disease causing mutational profile (95). Advances in next generation sequencing have made it possible to comparatively sequence the genotype of cancer cells and normal cells (96). The two high throughput methods commonly used to decipher genomic patterns in biological systems are DNA microarrays and whole genome sequencing (94).

Enhanced throughput utilized in DNA microarrays facilitates the genotyping and sequencing of patient tumours with a high degree of sensitivity and specificity (94). DNA microarrays comprise of probe sequences affixed to a solid surface that hybridize with fluorophore labelled sample DNA or RNA. The probe-DNA/RNA hybrid fluorescence signal is detected by a scanning confocal microscope and analysed by software (95,96). A two colour fluorescent tagging technique is often employed to compare

different tissue types, normal vs disease state or cell cultures subjected to different treatments (96). Whole genome sequencing provides a comprehensive, cost effective and robust method for sequencing the cancer genome (98). Using this technique, profiling of the human genome sequence has identified inherited DNA variants that increase susceptibility to several cancer types (99). This is the basis of genome wide association studies which evaluate the prevalence of these genetic variants across a given population (99). Whole genome sequencing employs a number of sequencing tools that provide high resolution and increased sensitivity such that even the smallest genetic variants can be detected[5]. Advances in high throughput next generation sequencing technology such as single-molecule real-time sequencing, ion torrent sequencing, nanopore sequencing and pyrosequencing facilitate detection of single nucleotide variants, deletions, copy number changes, and chromosomal translocations (102,103). A multitude of genetic alterations have been identified in the PCa genome, one of the most prominent being the differential expression of the AMACR gene. In summary the continuing advancements in next generation sequencing will accelerate and simplify the characterization of the cancer genome and eventually a central metric can be established to provide gene expression based diagnosis, prognosis and offer therapeutic targets (99).

1.8.2 Metabolomics

Metabolomics is an analytical approach that comprehensively studies the entire complement of metabolites in a biological system (101,102). In this approach low molecular weight metabolites are extracted from easily accessible biofluids and characterized to generate a metabolic responsive profile (103). Identification and quantification of metabolites relies on high through-put analytical technologies, namely nuclear magnetic resonance (NMR) spectroscopy and mass spectrometry (MS) (103). NMR provides rapid and non-destructive analysis of primary extracts while generating highly reproducible results. Gas chromatography (GC) and liquid chromatography (LC) coupled to MS are alternatives to NMR and offer higher resolution and separation reproducibility (103). NMR and MS technologies are often used in tandem to maximize the depth of metabolome analysis. Molecular targets of metabolomics constitute the metabolome and include lipids, thiols, amino acids, peptides, vitamins, nucleic acids and carbohydrates (102). These molecules take part in anabolic and catabolic reactions that sustain growth, maintenance and function (104). Therefore, analysis of metabolome can offer insights into molecular pathways that are disrupted in the disease state.

The metabolic profile of cancer cells shows enhanced glucose uptake, down regulation of oxidative phosphorylation and preferential conversion of pyruvate to lactate to replenish energy lost to dysregulated respiratory pathways (104). Several pathways are altered in tumour cells relative to normal cells, namely glycolysis, pentose phosphate pathway, tricarboxylic acid (TCA) cycle, nucleotide and protein biosynthesis (104). Different approaches may be used to study the metabolites of these altered pathways (figure 1.5) but the most commonly used analytical platforms for cancer research are metabolic fingerprinting which identifies and quantifies metabolites, and metabolic profiling which is

---

[5] http://www.illumina.com

typically used in tissue comparison analysis (104,105). In a study of PCa biomarkers by Soliman *et al.*, urine profiling by CE-ESI-MS/MS presented sarcosine as a potential biomarker (106).

1.8.3 Glycomics

Glycomics is a study that aims to characterize the structure and function of the full complement of glycans (glycome) produced by a cell or organism (107). Glycosylation is a posttranslational modification (PTM) occurring in the endoplasmic reticulum (ER) and Golgi apparatus which entails sugar moieties being attached to proteins (108). Approximately 50% of all proteins undergo glycosylation which contributes greatly to diversity of the proteome (109). Due to the abundance of glycoproteins in serum they are the principal focus for investigations of disease biomarker identification. Additionally, glycoproteins are ideal molecules for this analytical platform due to their ideal small size which makes them easy to quantitate (107). Extracellular matrices and cell surfaces are the predominant sites for glycosylation and this PTM therefore plays a pivotal role in cellular interactions (109). There is a considerable difference in the expression of glycosylation between tumour cells and normal cells due to the temporal and cell specific glycosylation that reflects the dynamic response of the phenotype to stimuli (108). MS based profiling of the glycome for disease biomarker identification revealed abnormal expression of glycosyltransferases and sugar nucleotide donors in tumour cells (109). A study by Saldova *et al.* that characterized the glycosylation of serum glycoproteins showed elevated α-2-3 linked sialic acid in PCa relative to BPH (109).
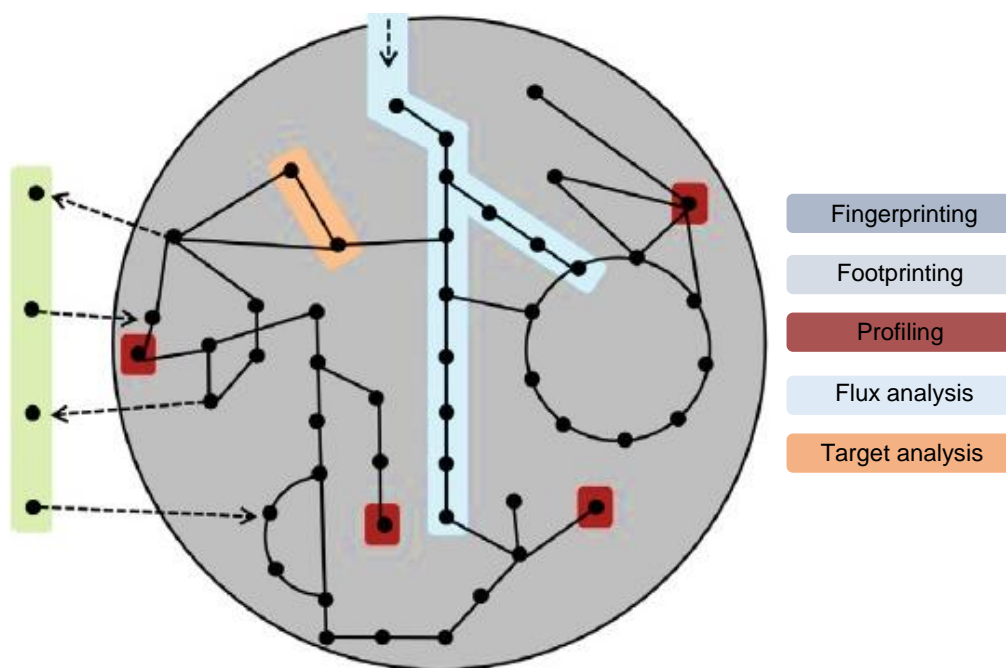
Figure 1.5: Different metabolomics approaches. Fingerprinting is the generation of a metabolic profile incorporating all metabolites in a system under investigation. Footprinting is the analysis of metabolites that are external to the system to evaluate metabolic exchange. Profiling is the analysis of a large class of molecules with the inclusion of standards. Flux analysis studies production and consumption rates of isotopically or radioactively metabolites in a biological system, it also determines the metabolic fate of the molecule. Target analysis compares two closely related target metabolites in a system using one particular analytical technique for the best performance. Modified with permission from Armitage *et al.*, 2013 (104).

1.8.4 Proteomics

Proteomics is a comprehensive study of proteins that quantifies the products of gene expression to characterize different biological states and delineate mechanisms of gene expression control (110). Initiation and progression of cancer is a multi-step process characterized by abnormal protein expression. These proteins can be detected in biological fluids such as saliva, blood or urine and analysed to identify PTMs or expression levels associated with a pathological state. Several high throughput proteomic technologies such as 2-DE/MALDI-TOF, SELDI/Protein Chip and most recently liquid chromatography tandem mass spectrometry (LC-MS/MS) powered by hybrid mass spectrometers are being utilized to identify novel protein biomarkers (111).

1.9 Proteomics as a viable platform for biomarker discovery

Proteomics is a study that focuses on characterizing the entire protein complement of a cell, tissue or organism (112). While DNA microarrays have previously been used as the principle tool to analyse protein expression (113), mRNA expression levels are not an appropriate proxy for protein abundance. Shortcomings of this genomic approach are apparent when considering that there are approximately 25 000 genes in comparison to almost a million proteins (114). The poor correlation between mRNA and protein can be attributed to dynamic mechanistic and regulatory processes which the proteins are subjected to (111,112). These include proteolytic degradation, PTM and complex formation. The wide array of post translational biochemical alterations that proteins undergo means that their expression is temporal and varies markedly in different tissues (115).

The proteins expressed by a cell or tissue are termed the proteome (116). Analysis of any given proteome poses a major challenge due to the diversity of protein species that arise from the afore mentioned processes (116). Proteomic studies therefore rely on state-of-the-art high-throughput technologies to address this complexity. Proteomic techniques uses a variety of techniques including various pre-fractionation methods, liquid chromatography, MS and bioinformatics for the accurate identification and characterization of proteomes (116). Proteomic approaches can be used for proteome profiling or to study protein-protein interactions, the identification of PTMs or for comparative protein expression studies and the elucidation of system wide protein networks (111,114,115) .

MS based proteomics has established itself as a viable platform for cancer biomarker discovery (119). Proteomic methods can generate protein profiles of tumour or normal cells in human clinical samples such as plasma, serum or urine (116,117). Since proteins are the functional units of expressed genes their analysis can offer insights into the pathophysiology of neoplastic transformation (120). The transformation of a normal cell into a cancer cell is reflected by distinct features at the protein level such as differential expression, altered specific activity, differential protein modification and altered cellular localization (61,117). Proteomics can therefore be utilized to elucidate proteins which can potentially serve as diagnostic tools for the treatment of cancer.

1.9.1 Mass Spectrometry

Mass spectrometry is an analytical technique that measures the mass to charge ratio (*m/z*) of ionized gaseous molecules (121). The mass to charge ratio of each molecule is plotted against the relative abundance to yield a mass spectra (122). A basic mass spectrometer consists of an ion source, mass analyser and detector (122). The ion source simultaneously ionizes and converts analyte molecules to the gas phase. A mass analyser subsequently resolves the analyte molecules according to their m/z ratio. The final component of the mass spectrometer is the detector which quantifies gaseous ion species corresponding to a m/z value (119,120). The analytical power of the mass analyser rests on three principal factors namely, sensitivity, resolution and mass accuracy (119,120). High performance mass analysers enable the measurement of analytes down to the femtogram level (122).

Proteomic research usually employs four types of mass analysers: ion trap IT/LTQ), time-of-flight (ToF), quadrupole and Fourier transform ion cyclotron (FTICR-MS) analysers (123). Each type of mass analyser has its own mass filtering method and some hybrid mass spectrometers combine two or more types to increase the analytical capabilities of the instrument (123).

1.9.1.1 Hybrid mass spectrometers in proteomics

Hybrid mass spectrometers combine the analytical power off two or more mass analysers into one instrument. These combinations increase the resolving power and mass accuracy of peptide spectral analysis in proteomic investigations (124). MS-based proteomic experiments typically analyse highly complex protein mixtures which presents a great analytical challenge. Proteins undergo proteolytic digestion to produce peptides which further add to the complexity of the experimental sample. Following proteolytic digestion peptides are separated by LC, converted into gaseous ions and analysed by tandem mass spectrometry. Due to the great abundance of peptides that are eluted into the mass spectrometer and the large order of magnitude difference of the ion signals, an instrument with high resolving power, duty cycle and sensitivity is required for proteomic experiments (125). One such instrument is the Orbitrap Fusion™ Tribrid™ mass spectrometer (figure 1.6) used in this study. The Orbitrap Fusion is a hybrid mass spectrometer that combines a quadrupole, linear ion trap (LIT) and orbitrap mass analyser (discussed below) (126). The greatest attribute of the Orbitrap Fusion is the ability to perform high resolution multi-stage tandem experiments and simultaneously analyse ions in the orbitrap or LIT[6]. Sensitivity and ion selection are considerably enhanced by the quadrupole mass analyser which selects precursor ions of a particular m/z enabling parallel mass analysis[7]. Additionally the functionality of this hybrid mass spectrometer enables increased scan rate and several fragmentation techniques which when taken together with its other features provide high performance in proteomic investigations[8].

---

[6] http://bioinformatics.cesb.uky.edu/bin/view/RCSIRM/AnalyticalCore

[7] http://www.kobis.si

[8] https://www.thermofisher.com/blog/proteomics/orbitrap-fusion-tribrid-mass-spectrometer
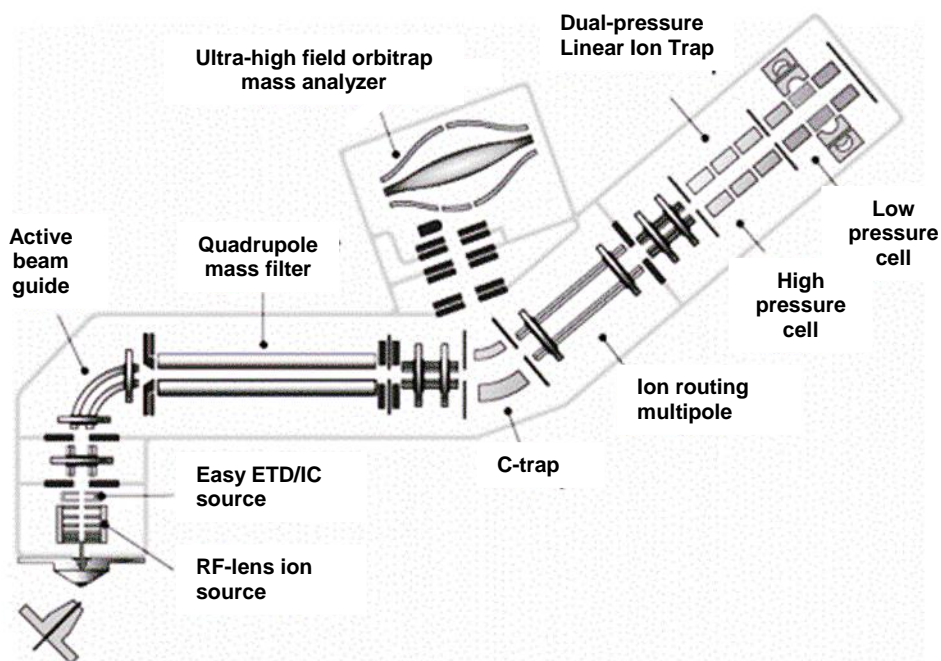
Figure 1.6: Schematic of the orbitrap fusion tribrid**.** The 'T' configuration of the instrument minimizes distance between the ion source and the mass analysers thereby reducing ion loss. This spatial configuration also provides sufficient space for ion fragmentation in the LIT which improves parallelized analysis. Modified with permission from Senko *et al.*, 2013 (126).

Quadrupole

Quadrupole mass analysers consist of four cylindrical rods arranged parallel to each other with opposite pairs connected electrically (127). A direct current (DC) voltage and superimposed radio frequency (RF) voltage is applied to each pair of rods to create a symmetrical electric field (128). A positive DC voltage is placed on the first pair of rods while the second pair of rods carries a negative DC and the RF field is phase shifted by 180° (129). Ions entering the analyser are separated by the electric field which causes them to oscillate and attain one of two types of trajectory. At a particular DC and RF potential only ions of a particular m/z attain a stable trajectory along the axis of the quadrupole. These ions possess a "bounded oscillation" trajectory and ultimately pass on to the detector. Conversely ions of different m/z attain an unstable trajectory that is perpendicular to the path of flight. Ions with this unstable oscillatory pattern are discharged either on the rods or a pre-filter that may be fitted in front of the analyser. By altering the DC – RF voltage ratio (scan function) ions of different m/z pass onto the detector producing mass spectra of analyte ions. Quadrupoles are classified as scanning mass analysers and as such their scan function progresses in a linear manner such that obtaining spectra takes a long time. The limitation of the acquisition rate results in spectral bias. A positive aspect of this analyser is its ability to select and analyse an ion with a particular m/z. This is known as single ion monitoring (SIM) and the analyser is very sensitive at this chosen m/z (127).

Liner Ion Trap

Linear ion traps consist of a set of quadrupoles to which an RF voltage is applied producing an oscillating ion trajectory (128). In principle, linear ion traps have a similar operating principle to quadrupole mass analysers. A distinguishing feature of the linear trap is how ions are confined in 2D along the axis of a segmented quadrupole. The ions are confined radially and alternately travel between both ends of the quadrupole. Another feature of the ion trap is the incorporation of low pressure helium gas. The gas acts as a buffer which prevents repulsion between like charged ions and consequently increases mass accuracy and resolution (130).

At the commencement of each trapping cycle the potential field is modulated to incorporate ions of a broad m/z spectrum into the trap (130). Following the trapping of ions in the analyser the field is further adjusted to confer a stable trajectory to all the ions present. Mass analysis is effected by a gradual increase in the amplitude of the RF trapping voltage which causes ions to have an unstable trajectory (125,127). This is known as resonant ejection and culminates in ions being ejected from the trap in order of increasing m/z and proceeding into the detector (125,127). In this manner, a mass spectrum is generated by a trapping cycle and subsequent resonant ejection into the detector.

Ion traps are renowned for sensitivity which relies on high efficiency of two aspects of the mass analyser. Sensitivity is dependent on efficient ion transmission from the mass analyser to the detector and maximal ion utilization (128). Efficient ion transmission may be attributed to the compact size of the analyser. Ion traps are designed in such a way that the trapping region is only a few centimetres from the detector (128). Also adding to the sensitivity of this instrument is the efficient utilization of ions produced by the ion source. This may be attributed to the ion trap-resonant ejection system of the mass analyser. Axial trapping of ions in the linear ion trap increases ion capacity, scan rate as well as detection and trapping efficiency. Consequently, these mass analysers are instrumental to proteomic studies as they can identify peptides in highly complex biological mixtures with remarkable sensitivity. Practicality convenience, low maintenance and the relatively cheap cost of this mass analyser are a few other factors that have made it popular in proteomic studies (131).

Orbitrap

An orbitrap mass analyser consists of a central (axial) spindle like electrode encapsulated by a coaxial outer barrel like electrode (128,129). This analyser does not rely on RF or magnetic fields like typical ion traps (133). Instead ions are orbitally trapped by an electrostatic force of attraction to the central spindle (133). Initial tangential force of ions generates a centrifugal force which balances this electrostatic field (130,131). The electrostatic fields produced by the axially symmetric electrodes of the orbitrap create a quadro-logarithmic electrostatic potential. In this electrostatic field trapped ions acquire a complex spiralling trajectory which consists of rotation around the central electrode and simultaneous harmonic oscillation (axial) along its length (130,133). The frequency of harmonic oscillation is independent of ion velocity and is inversely proportional to the square root of m/z (136). The axial oscillation of the ions creates an image current which is detected by the outer barrel electrode (137).

The outer electrodes are split in half at z = 0 and each ion current is amplified differentially (138). The resulting signal is processed analogous to the FT-ICR to produce time domain image current transients (138). Mass spectrum are produced by Fourier transforming these transients (134).

The orbitrap is a powerful mass analyser with renowned resolving power and high mass accuracy. It possesses a mass accuracy of 2-5 ppm, resolving power of 450 000, m/z range of 6000 and a dynamic range of 10^6 (137). In proteomic studies such high resolution allows the mass spectrometer to resolve near isobaric peptides in the same time window and helps prevent co-fragmentation. This is a prerequisite for accurate protein identification and quantification (139).

1.9.2 MS based proteomic strategies for protein identification

Analysis of the proteome utilises several technical methods that incorporate either of two strategies for the identification of proteins. The first strategy utilizes whole-intact proteins (top-down proteomics) while the second strategy relies on enzymatic or chemical proteolysis of proteins to derive peptides (bottom-up proteomics or shotgun proteomics) for identification (121). A common feature of both top-down and bottom-up methods is the use of front end separation techniques that reduce sample complexity. Multiple stages of separation implement chromatographic techniques or polyacrylamide gel electrophoresis which increase overall sensitivity and reproducibility of results for the given experiment (109,120).

1.9.2.1 Top-down proteomics

The top down approach entails ionization of whole proteins, preliminary mass analysis, fragmentation and subsequent protein identification (121). Prior to sample infusion into the mass spectrometer, several fractionation strategies are used to reduce the great sample complexity. Some of these protein separation techniques include capillary electrophoresis, tube gel electrophoresis and LC. The majority of these separation strategies can be carried out offline which consequently increases data acquisition time and greatly reduces fractionation method-MS compatibility stringency (140). Sample separation of intact proteins poses a challenging task due to the co-elution of isotopic variants. This can however be circumvented by using mass analysers capable of high mass accuracy and resolution such as the FT ICR and orbitrap (112).

Fragmentation methods commonly used in the top down approach include electron capture dissociation (ECD) and electron transfer dissociation (ETD) (112). These are the preferred fragmentation methods as they provide excellent sequence coverage and reliable characterization of labile PTMs (141). The top-down approach in combination with ETD/ECD fragmentation of intact proteoforms greatly simplifies protein quantification and increases resolution of protein isoforms due to greater sequence coverage (112). Furthermore reduced molecular complexity of fractioned and intact proteoforms minimizes inconclusive protein identification seen in peptide to protein mapping of the bottom-up approach (112).

1.9.2.2 Bottom-up proteomics

The bottom-up approach is the method of choice for high throughput analysis of highly complex samples. Using this approach, proteins undergo proteolysis to produce peptides prior to mass spectrometry (112). The peptide fragment mass spectra produced by ion fragmentation are searched against protein databases by search engines such as Mascot or X!Tandem (111). Protein identities are assigned by comparing experimental fragment spectra to theoretical spectra generated by *in silico* proteolysis of a protein database. Peptides may be uniquely assigned to a single protein or several proteins (111). A protein identified by several unique peptide spectra improves the confidence of identification whereas a protein identified by few and redundant peptide spectra reduces the confidence of identification (142). Some advantages of the bottom-up method include better front end separation of peptides and increased sensitivity relative to the top-down method (121). Despite the effectiveness of the bottom-up approach in large scale analysis of different proteomes its greatest drawback lies in the increased sample complexity created by protein proteolysis. Peptides produced by proteolysis can sometimes be incorrectly assigned to proteins especially when they possess PTMs. Additionally, due to limitations in instrument dynamic range only the high abundant peptides generate a detectable ion signal while the ion signal of low abundant peptides is masked (121).

1.9.3 Quantitative proteomics

Quantitative proteomics is a high throughput method used to measure the relative or absolute abundance of proteins. The relative abundance of proteins between different experimental conditions such as the normal and disease state, or perturbed and control samples is an important feature of biomarker studies (139). Absolute protein quantification measures the exact amount of protein in a sample and is therefore better suited to targeted analysis of a specific protein (143). There are two approaches used to quantify proteins in proteomic studies, namely label free methods and isotope labelling methods (139).

Label free quantification methods do not require the chemical modification of peptides as a prerequisite to protein quantification (144). The label free approach can follow either of two methods which are spectral counting and peptide peak intensity measurement (143). Spectral counting correlates the protein abundance to the amount of MS/MS spectra assigned to peptides or a specific protein (140,141). This is based on the principle that high abundant proteins produce more MS/MS spectra relative to low abundant proteins. Also the peptides of high abundant proteins undergo more sampling in fragment ion scans relative to low abundant peptides (143). In peptide peak intensity measurement protein abundance is correlated to spectral ion signal intensity (144). The peptide ion peaks from an LC-MS run are integrated and compared to other experimental values of the same ion species to determine the relative quantity (143). Implementation of label free quantification methods is becoming increasingly popular due to the following factors: low cost, less labour intensive and accommodation of a higher dynamic range (143).

Isotope labelling quantification methods rely on the incorporation of a mass tag of known molecular weight either by a chemical or metabolic method. The incorporation of a differential mass tag does not

adversely alter the chemical properties of the protein or peptide (143). Downstream LC separation and MS analysis can therefore proceed unhindered. The principle of this quantification method is based on evaluation of the mass difference between equal amounts of peptides due to differential isotope labelling (145). Peptide samples are typically tagged with a combination of a heavy and light isotope/left unlabelled. The samples are pooled, analysed by MS and the ratio of the isotopic tag variants is used to determine protein relative abundances (145). In metabolic labelling a heavy isotope, typically ^{15}N is introduced into proteins by culturing cells in media supplemented with a ^{15}N nitrogen source (144). Cells grown in ^{15}N enriched media naturally integrate this heavy nitrogen isotope into protein during growth and division (139). Alternatively, labelling can be carried out by incorporation of heavy and light isoforms of either arginine or lysine into proteins via cell culture in a method called 'stable isotope labelling by amino acids in cell culture' (SILAC). Metabolic labelling methods are however only compatible with a limited number of biological studies.

Chemical labelling methods offer a robust alternative making use of several methods such as isotope-coded affinity tags (ICAT), isobaric tag for relative and absolute quantification (iTRAQ) and tandem mass tag (TMT) (144). The ICAT method is used to compare protein expression levels between two different samples. Using ICAT a chemical label is added to cysteine residues of peptides subsequent to protein trypsinization (145). This method reduces sample complexity since it only tracks peptides with cysteine residues which coincidentally occur with low frequency. ICAT labelling is not used in biomarker studies due to the bias in quantification introduced by tagging of low abundant cysteine residues and the fact that it can only analyse two samples/states at a given time. Biomarker discovery studies typically employ the isobaric tagging methods iTRAQ and TMT which can analyse up to 8 (10 with TMT 10) samples in parallel (143). In iTRAQ and TMT variant isotopic tags are chemically attached to the primary amine groups on the peptides (144). Peptide fragmentation during MS analysis produces reporter ions. The relative intensity of these ions can then be used to determine the relative abundance of proteins (136,140).

1.9.4 Secretome proteomics for biomarker discovery

The term secretome refers to proteins that are released by a cell, tissue or organism. This also includes extracellular proteins and proteins that are shed from the cell membrane (146). Protein secretion is a cellular process carried out by either of two secretory pathways, the classical secretory pathway or the non-classical secretory pathway (147). The classical secretory pathway is well characterized and begins with synthesis of protein precursors in the rough ER. These protein precursors are transported to the Golgi apparatus and subsequently released into the microenvironment by fusion of secretory vesicles with the plasma membrane. Conversely, the non-classical secretory pathway is poorly characterized and it is thought that a relatively small proportion of proteins are secreted via this pathway (108,140) . Several mechanisms are responsible for protein secretion in the non-classical pathway and all lack the N-signal peptide sequence characteristic of proteins secreted by the classical pathway. In the first mechanism proteins are taken into endosomal compartments. Through the process of endosomal recycling in which the endosomes fuse with the plasma membrane, proteins are released into the extracellular space. The second mechanism is mediated by a ping-pong mechanism that

directly transports proteins into the extracellular space. A precondition for this mechanism is that the proteins are anchored to the membrane by dual acylation. Lastly, proteins may be secreted by exosomes via fusion of the plasma membrane and vesicle (147).

Despite only 10-15% of the human genome encoding secreted proteins they play a vital physiological role in cell-cell communication, cell signalling and matrix remodelling (147). The secretome constitutes an important sub-class of the proteome as it is reflective of the dynamic inter-cellular crosstalk in an organism. The cancer secretome is therefore a viable source of biomarkers as it is involved in processes associated with tumour progression such as angiogenesis, metastasis, differentiation and invasion (148). Tumour progression is also dependent on the interaction between malignant cells and stromal cells cell in the tumour microenvironment. A host of signalling proteins such as cytokines, chemokines and growth factors mediate communication between the different cell types which favours progression of the malignant phenotype. Proteomic analysis of both the tumour and stromal cell secretome can thus provide a more comprehensive panel of biomarkers (147).

MS based proteomic analysis of blood, conditioned media (CM) or tumour proximal fluid is the most widespread approach of mining the cancer secretome for biomarkers (149). Human blood by far presents the largest repository of potential biomarkers. However, LC-MS/MS analysis of blood has proven challenging due to the large diversity of proteins with a dynamic range of 9-10 orders of magnitude. High abundant proteins such as albumin, immunoglobulin and transferrin tend to have a masking effect on the low abundant secreted proteins. Separation methods to reduce blood complexity such as immunoaffinity chromatography and Protein G resins have been implemented to remove high abundant protein species. These protein depletion strategies however lower the accuracy of the investigation due to non-specific binding and proteins of interest forming complexes with the filtered high abundant proteins (149).

An alternative source of secreted biomarkers lies in the CM of cancer cell lines. Cancer cell lines secrete and shed proteins into the CM which is collected and subsequently analysed. CM is considerably less complex compared to blood and tumour proximal fluid such that low abundant proteins can be easily identified. Furthermore, cancer cell lines are grown under standardized conditions which increases reproducibility of results. Different cell lines are representative of varied tumour histotypes and stage thus increasing the scope of the study and accounting for the heterogeneity of cancer. The drawback of CM analysis however, lies in the fact that tissue culture does not completely mimic the different factors and physiological interactions that account for the complexity of cancer. Also, altered protein expression in cell lines can be attributed to stressful growth conditions encountered in tissue culture (147).

Tumour interstitial fluid is yet another viable source of secreted protein biomarkers. Tumour proximal fluid is a biological fluid that surrounds both stromal and tumour cells. It is a rich source of biomarkers as tumour cells secrete signalling molecules involved in metastasis and progression directly into this

fluid. An advantage of this biomarker source is that it is a direct source of secreted proteins from tumour cells *in vivo* but with relatively less complexity in comparison to blood (150). The main drawback of tumour proximal fluid lie in its extraction which can result in contamination by serum proteins (147).

Tumour progression is a multi-step process that incorporates a host of signalling factors and different cell types. It has become evident that tumour cells recruit proximal stromal cells in their microenvironment that promote metastasis and invasion of distant organs. Therefore, analysis of the tumour secretome can offer insights into cancer pathobiology and present targets for therapeutic intervention.

1.10 Aims and objectives

This study aim to identify novel biomarkers which can be used to distinguish BPH from PCa by comparative proteomic analysis of the secretome and proteome of prostate cancer cells. This study used a tissue culture based model to generate profiles for the secretome and proteome of prostate cell lines. For the investigation of the secretome, cells were grown in serum free media which was concentrated and analysed by LC-MS/MS. For the proteome study, cells were grown, collected and analysed by LC-MS/MS. The proteins that were identified by MS were compared between the four cell lines and the differentially expressed proteins in BPH-1 cell line were further investigated as potential biomarkers. The specific objectives of this study were as follows:

1. To optimize the growth conditions of the prostate cell lines (LNCaP, PC3, PNT2 and BPH-1) in order to reduce cell death and maximize the levels of secreted proteins.

2. To optimize purification methods of secreted proteins in the conditioned media.

3. To investigate the secretome and proteome of four prostate cell lines (BPH-1, LNCaP, PC3 and PNT2) using mass spectrometry based proteomics.

Chapter 2

Materials and methods

2.1 Reagents

LNCaP, PC-3 and PNT2 cells were purchased from the European Collection of Authenticated Cell Cultures (ECACC). BPH-1 cells were a generous gift from Simon W Hayward (Vanderbilt University of Medical Center, Nashville, USA). Ham's F12K medium, RPMI-1640 medium, acetone (HPLC grade) and 3.5 kDa dialysis membrane were all purchased from Sigma-Aldrich (St. Louis, MO, USA). Chinese hamster ovary medium (CDCHO), L-glutamine, penicillin-streptomycin, fetal calf serum (FCS) and trypsin-EDTA were obtained from Gibco-BRL (Gaithersburg, MD, USA). Bicinchoninic acid (BCA) protein determination kits and lactate dehydrogenase cell viability assays were purchased from Pierce (Rockford, IL, USA) and Roche Diagnostics (Mannheim Germany), respectively. Trypan blue stain (0.4%) and cell count plates were purchased from Invitrogen (Eugene, USA). Corning® CellBIND® T-175 flasks were purchased from Corning® Life Sciences (NY, USA) and ammonium sulphate was purchased from Protea Chemicals (Western Cape, South Africa). All other chemicals were of the highest analytical grade and purchased from scientific supply houses. All protocols were carried out according to the manufacturer's instructions unless otherwise stated.

2.2 Prostate cell line tissue culture

2.2.1 Cell culture general growth conditions

BPH-1 and PNT2 cells were grown in RPMI-1640 media supplemented with 10% fetal calf serum 1.5 g $NaHCO_3$/L (pH 7) and 1% penicillin-streptomycin. PC-3 cells were grown in Ham's F12K media supplemented with 10% FCS, 1.5 g $NaHCO_3$/L (pH 7) and 1% penicillin-streptomycin. LNCaP cells were grown in RPMI-1640 media supplemented with 10% fetal calf serum (FCS), 1.5 g $NaHCO_3$/L (pH 7), 2.5 g D-(+)-Glucose, 1% 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), 1% sodium pyruvate and 1% penicillin-streptomycin. All cell lines were grown in a humidified incubator at 37°C and 5% CO_2. Cell count and viability were determined using trypan blue and a Countess® Automated cell counter (Invitrogen).

2.2.2 Cell culture growth conditions for analysis of the proteome

All cell lines were maintained as described in section 2.2.1 above. On the third passage all cell lines were plated in 75 cm^2 tissue culture flasks at a density of $5x10^4$ cells/mL(10mL) and were grown to 80% confluency. Once the cells had reached the desired confluence the media was aspirated from the culture flasks and the cells were gently washed with 20 mL of phosphate buffered saline (PBS) (pH 7.4). The cells were detached from the flasks using 4 mL pre-warmed trypsin. The detached cells were transferred into a conical tubes containing 10 mL of conditioned media and centrifuged at 600 x g for 5 minutes. Following centrifugation the conditioned media was aspirated, the cell pellet was re-suspended

in 10 mL PBS and again centrifuged at 600 x g for 5 minutes. The cells were then re-suspended in 1 mL of PBS and centrifuged at 600 x g for 5 minutes. This step was repeated once more in order to remove serum proteins from the FCS used to supplement conditioned media. The PBS was subsequently aspirated, the cells were weighed and the mass was adjusted to 50 mg for each biological repeat by pipetting off excess cells. The samples were stored at -80˚C till until use. Three independent biological experiments were performed for each cell line.

2.2.3 Cell culture growth conditions for analysis of the cell secretome

All cell lines were maintained as described in section 2.2.1 above. On the third passage BPH-1, PNT2 and PC-3 cells were plated into T175 flasks at varying densities in 30 mL of their respective CM. BPH-1 cells were seeded at 2, 5 and 8 million cells per flask; PNT2 cells were seeded at 4, 6 and 8 million cells per flask; PC-3 cells were seeded at 5, 8 and 11 million cells per flask; and LNCaP cells were seeded at 11, 16.5 and 22 million cells per flask. The cells were then incubated for 48 hours, with the exception of the LNCaP cells which were incubated for 72 hours. Following this incubation period, the conditioned media was aspirated and cells were gently washed twice with 20 mL of PBS. Next, the cells were incubated in 30 mL of CDCHO supplemented with L-glutamine (8 mmol/L) for 48 hours. After incubation in CDCHO, the conditioned media was collected and centrifuged at 600 x g for 5 minutes to remove cellular debris. Aliquots of the conditioned media were then collected for the measurement of total protein and LDH while the remainder was stored at -80˚C until further use.

2.3 Measurement of total protein and LDH

Total protein in the conditioned media was measured by the Bradford assay. Twenty microliters of each secretome sample was thawed and added to a 96 well transparent microplate (Greiner Bio-One International) in triplicate, together with a range of bovine serum albumin (BSA) standards which were prepared in deionized water (table 2.1.). Samples were incubated at room temperature for 30 minutes following the addition of 200 µL of Bradford reagent into each well. Absorbance at 540 nm was subsequently measured (Biotek Powerwave 340 Plate Reader). The average absorbance measurement for the blank (CDCHO) replicates was subtracted from the measurements of all other individual standard and unknown sample replicates. The standard curve was then generated by plotting the blank-corrected measurement for each BSA standard versus concentration. The standard curve was subsequently used to determine the protein concentration for each unknown sample.

Table 2.1: Preparation of protein standards for Bradford protein assay.

Dilution	Volume of deionized water (µL)	Volume and source of BSA (µL)	Final BSA concentration (µg/mL)
A	0	300 of stock (2mg/mL)	2000
B	125	375 of stock (2mg/mL)	1500
C	325	325 of stock (2mg/mL)	1000
D	175	175 of B dilution	750
E	325	325 of C dilution	500
F	325	325 of E dilution	250
G	325	325 of F dilution	125
H	400	100 of G dilution	0
I	400	0	0 = Blank

LDH concentration in the conditioned media was measured by an LDH assay kit (Roche Diagnostics) that reduces NAD to NADH. This reaction was detected by a colorimetric assay that produces a characteristic red formazan product which was measured spectrophotometrically at a wavelength of 492 nm.

One hundred microliters of each secretome sample was thawed and added to a 96 well transparent microplate (Greiner Bio-One International) in triplicate together with a range of LDH (table 2.2) standards prepared in CDCHO medium (table 2.2). The LDH stock solution was prepared by adding 2 µL of LDH stock (ScienCell) to 498 µL of PBS (final concentration 1000 mU/mL). Samples were incubated at room temperature in the dark for 30 minutes following the addition of 100 µL of working reagent. Absorbance at 492 nm was subsequently measured (Biotek Powerwave 340 Plate Reader). The average absorbance measurement of the blank (CDCHO medium) replicates was subtracted from the measurements of all other individual standard and unknown sample replicates. The standard curve was then generated by plotting the blank-corrected measurement for each LDH standard versus concentration. The standard curve was subsequently used to determine the LDH concentration for each unknown sample.

Table 2.2: Preparation of LDH standards for LDH cell viability assay.

Dilution	Volume of CDCHO added (µL)	Volume and source of LDH (µL)	Final LDH concentration (mU/mL)
1	450	450 of stock (1000mU/mL)	500
2	450	450 (of dilution 1)	250
3	450	450 (of dilution 2)	125
4	450	450 (of dilution 3)	62,5
5	450	450 (of dilution 4)	31,8
6	450	450 (of dilution 5)	15,6
7	450	450 (of dilution 6)	7,8
8	450	0	0 = Blank

2.4 Optimization of protein precipitation from conditioned media

Samples were prepared by pooling three 25 mL secretome samples from the BPH cell line and stirring for 15 minutes to achieve homogeneity. The pooled samples were then aliquoted equally into three conical tubes. Each conical tube containing 25 mL of the homogenous secretome sample underwent one of three protein precipitation techniques (described below). Protein concentration was determined using the Pierce BCA method subsequent to protein precipitation to evaluate the effectiveness of each technique. The experiment was repeated using pooled PNT2 secretome samples.

2.4.1 Acetone precipitation

Acetone was cooled overnight to -20˚C while the protein samples where thawed at 4°C. Four volumes (100 mL) of cooled acetone was subsequently added to a precooled conical flask followed by one volume of the secretome sample (25 mL). The acetone-sample mixture was then stirred for 5 minutes and incubated overnight at -20°C. Next, the mixture was centrifuged at 15000 x g for 10 minutes. The supernatant was discarded and the protein pellet was air dried. Following this, the protein pellet was re-suspended in 5 mL of deionized water and lyophilized overnight by a speedy-vac (LabConco Centrivap). The samples were re-suspended in 2 mL of deionized water and the protein concentration was determined using the BCA assay. All samples were stored at -80˚C after the protein concentration determination for later use.

2.4.2 Ammonium sulphate precipitation

A 25 mL of a secretome aliquot was poured into a 50 mL centrifuge tube and a pre-determined amount of ammonium sulphate (13.32 g) was added in order to obtain 80% saturation. The mixture was then incubated overnight at 4°C. Next, the protein-salt mixture was centrifuged at 10000 x g for 20 minutes. The resulting supernatant was discarded and the protein pellet was re-suspended in 5 mL deionized water. The re-suspended protein was dialyzed for 24 hours using a 3.5 kDa cut-off dialysis tubing at 4°C in 5 L of deionized water. The deionized water was changed twice after 4 hour intervals. Following dialysis, the sample was lyophilized overnight by a speedy-vac (LabConco Centrivap) to concentrate the protein and then re-suspended in 1 mL of deionized water. The protein concentration was subsequently determined by the BCA assay. Samples were then stored at -80˚C until required for further use.

2.4.3 Acetone-Phosphotungstic Acid (PTA) precipitation

Acetone was cooled overnight to -20˚C while the protein samples where thawed at 4°C. Four volumes (100 mL) of cooled acetone were subsequently added to a precooled conical flask followed by one volume of the secretome sample (25 mL). PTA was then added to a final concentration of 1.2% in the acetone-sample mixture. The acetone-sample mixture was stirred for 5 minutes and left to stand overnight at -20°C. Next, the mixture was centrifuged at 15000 x g for 10 minutes. The supernatant was discarded and the protein pellet was air dried. Following this, the protein pellet was re-suspended in 5 mL of deionized water and lyophilized overnight by a speedy-vac (LabConco Centrivap). The samples were re-suspended in 2 mL of deionized water and the protein concentration was determined by the BCA assay. Samples were then stored at -80˚C until required for further use.

2.4.4 BCA protein determination

Protein concentration following precipitation was measured using the Pierce BCA assay (Thermo Scientifc). The BCA assay was preferred to the Bradford method due to the incompatibility of the latter method with residual acetone present in the precipitated samples. Twenty five microliters of each secretome sample was thawed and added to a 96 well transparent microplate (Greiner Bio-One International) in triplicate together with a range BSA standards prepared in deionized water (table 2.1.). Samples were incubated at room temperature for 30 minutes at 37°C following the addition of 200 µL of working reagent into each well. The absorbance was subsequently measured at 562 nm (Biotek Powerwave 340 Plate Reader). The average absorbance measurement of the blank (deionized water) replicates was subtracted from the measurements of all other individual standard and unknown sample replicates. The standard curve was then generated by plotting the blank-corrected measurement for each BSA standard versus concentration. The standard curve was subsequently used to determine the protein concentration for each unknown sample

Table 2.3: Preparation of protein standards for BCA protein assay.

Dilution	Volume of deionized water (µL)	Volume and source of BSA (µL)	Final BSA concentration (µg/mL)
A	0	300 of stock (2mg/mL)	2000
B	125	375 of stock (2mg/mL)	1500
C	325	325 of stock (2mg/mL)	1000
D	175	175 of B dilution	750
E	325	325 of C dilution	500
F	325	325 of E dilution	250
G	325	325 of F dilution	125
H	400	100 of G dilution	0
I	400	0	0 = Blank

2.5 Preparation of cell pellets for Mass Spectrometry

Cells were thawed in 100 µL of extraction buffer containing 100 mM NaCl, 2 mM EDTA, 6 M guanidine-HCl, 1% OGP and 5 mM TCEP in 100 mM triethylammonium bicarbonate (TEAB, pH8). The cell-extraction buffer suspensions were sonicated and centrifuged at 12000 x g for 5 minutes and the supernatant collected. The pellet was re-suspended in 50 µL of extraction buffer and subjected to a second round of extraction. The resulting supernatants were pooled with the supernatants generated in the initial round of extraction. Proteins were subsequently precipitated with ice cold acetone in a 1:5 sample to acetone ratio. Proteins were allowed precipitate overnight before pelleting by centrifugation at 12 000 x g for 10 minutes. The resulting supernatant was discarded and the pellet air dried before being re-suspended in 100 µL of 100 mM TEAB containing 4 M guanidine-HCl and 1% octylgluco-mano-pyranoside. The disulphide bonds in the protein samples were then reduced uby adding 10 µL of 50 mM TCEP (final concentration of 5 mM) in 100 mM TEAB for 30 minutes at room temperature. Following this the reduced cysteine residues were subsequently modified using 5 µL of 200 mM methyl methanethiosulfonate (MMTS) (final concentration of 20 mM) in 100 mM TEAB buffer for 30 minutes. The samples were then diluted to 98 µL with 100 mM TEAB. Next, proteins were digested by adding 10 µL trypsin (Pierce) solution (1 µg/µL) and incubating for 6 hours at room temperature followed by a

further 16 hours at 37 °C. Samples were subsequently dried and re-suspended in 100 μL of 2% acetonitrile containing 0.1% formic acid and subsequently de-salted using C18 stage tips. Three independent biological experiments were performed for each cell line. Individual samples were analysed in triplicate using LC-MS.

2.6 LC-MS analysis of the cell proteome

Peptides were separated by liquid chromatography using a Thermo Scientific Ultimate 3000 RSLC equipped with a C18 trap column (300 μm x 5 mm (Thermo Scientific)) and an in house manufactured analytically column (75 μm x 40 cm). The mobile phases consisted of solvent A (2% acetonitrile:water; 0.1% formic acid) and solvent B (100% acetonitrile:water). Flow rate was set to 350 nL/minute and the gradient as follows: 2% A for 5 minutes; 2 - 4% B from 5-10 minutes, 4 - 10% from 10-20 minutes, 10 - 40% from 20-95 minutes using Chromeleon non-linear gradient 7, 40 - 80% B and from 95-100 minutes. Thereafter the column was washed for 10 minutes with 80% B followed by equilibration. Chromatography was performed at 50°C and the outflow delivered to the mass spectrometer through a stainless steel nano-bore emitter. Eluted peptides were analysed by an Orbitrap Fusion Tribrid mass spectrometer (Thermo Scientific) equipped with a Nanospray Flex ionization source. Data was collected in positive mode with spray voltage set to 2 kV and ion transfer capillary set to 275°C. Spectra were internally calibrated using polysiloxane ions at m/z = 445.12003 and 371.10024. MS1 scans were performed using the orbitrap detector set at 12 000 resolution over the scan range 350-1650 with AGC target at 3 E5 and maximum injection time of 40ms. Data was acquired in profile mode. MS2 acquisitions were performed using monoisotopic precursor selection for ion with charges +2- +6 with error tolerance set to +/- 10ppm. The raw files generated by the MS analysis were imported into Proteome Discoverer v1.4 (Thermo Scientific) and processed using the Mascot and SequestHT algorithms. Peptide validation was set to search against a false decoy rate (FDR) of 1%. Additional analyses were performed using the X!Tandem Sledgehammer algorithm. The files generated by data analysis of the spectra were combined and analysed using Scaffold 1.4 (Proteome software).

Chapter 3

Results

3.1 Optimization of secretome samples

The total protein and LDH concentration in the conditioned chemically defined serum free media obtained from the BPH-1, PNT2, LNCaP and PC3 cell lines was measured. The ratio of LDH to total protein concentration was determined for each seeding density for all cell lines in order to determine the optimal seeding density. The optimal seeding density was considered to be that density that had the highest amount of secreted proteins, but lowest amount of intracellular proteins as measured by the LDH assay. Intracellular proteins have a masking effect on relatively low abundant secreted proteins, therefore optimization maximises the amount of secreted proteins and increases the depth of secretome coverage.

The protein concertation was determined using the Bradford protein determination method as described in section 2.3. Unsurprisingly, the total protein concentration tended to increase with the seeding density, however, this trend was not significant in all cases. Statistically significant differences included a 2.5-fold increase in protein concentration between the lowest seeding density and 5×10^6 cells and a 3.0-fold increase between the lowest seeding density and 8×10^6 cells for the BPH-1 cell line. A 2.0-fold increase in protein concentration was observed between the lowest and highest seeding density in the PC3 cell lines, while in the LNCaP cell line a 2.10-fold increase was observed between the lowest and highest seeding density. No statistically significant differences were observed in the PNT2 cell line. The total protein concentrations for the cell lines (figure 3.1) are listed below in order of increasing seeding density: BPH-1- 13.93 µg/mL, 34.27 µg/mL and 41.88 µg/mL; PNT2- 19.08 µg/mL, 24.97 µg/mL and 31.12 µg/mL; LNCaP- 15.10 µg/mL, 22.89 µg/mL and 32.17 µg/mL; PC3- 11.52 µg/mL, 18.72 µg/mL and 23.17 µg/ml.
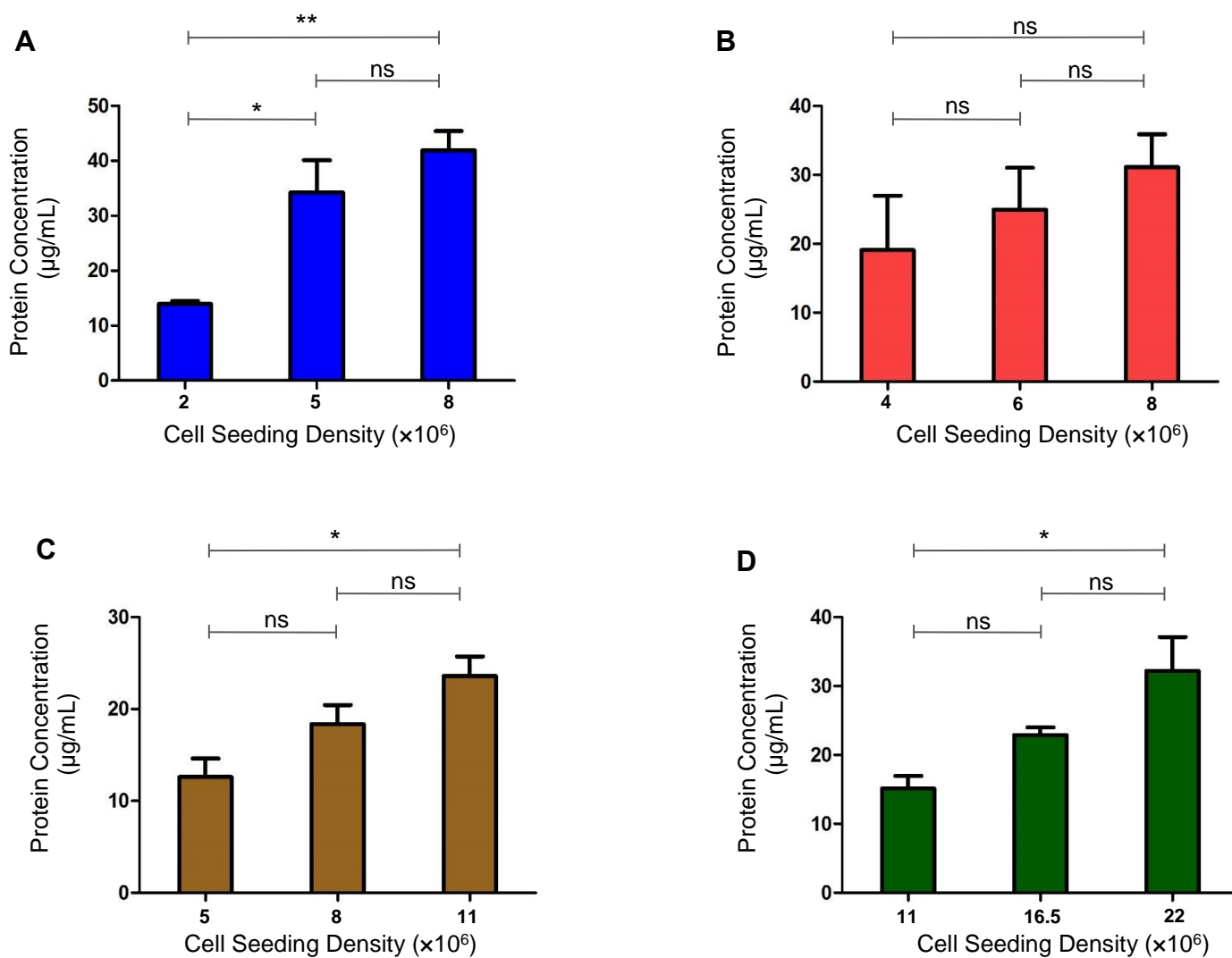
Figure 3.1: Total protein content in the serum free conditioned media from (A) BPH-1, (B) PNT2, (C) PC3 and (D) LNCaP cell lines. Results are shown as means ± SEM of three independent experiments each performed in triplicate. One-way analysis of variance (ANOVA) with Tukey's post-test were used for group comparisons. *P < 0.05; **P < 0.01; ***P < 0.001; ns= not significant

The LDH concentration was measured in the conditioned chemically defined serum free media of prostate cell lines using a commercially purchased LDH kit as described in section 2.3. Similar to the protein concentration, the LDH levels tended to increase with increased seeding densities (figure 3.2), though this trend was again not significant in all cases. In BPH-1 all the observed trends were statistically significant. The fold increase in LDH concentration were 2.4 (5×10^6 cells) and 3.0 (8×10^6 cells) with respect to the lowest seeding density. The trends in PNT2 LDH concentrations were as follows: a 2.6-fold increase in LDH concentration between the lowest seeding density and 6×10^6 cells (P< 0.05) and a 4.0-fold increase between the lowest seeding density and 8×10^6 cells (P< 0.001). PC3 showed a 1.6-fold increase in LDH concentration between 5×10^6 cells and 11×10^6 cells (P< 0.001) and a 1.3-fold increase between 8×10^6 cells and 11×10^6 cells (P< 0.05). The trends observed in LNCaP were not statistically significant. The LDH concentrations for the cell lines are listed below in order of increasing seeding density: BPH-1- 381.80 mU/mL, 917.99 mU/mL, 1137.57 mU/mL; PNT2- 132.34 mU/mL, 345.29 mU/mL and 531.22 mU/mL; LNCaP- 266.83 mU/mL, 445.39 mU/mL and 802.78 mU/mL; PC3- 269.51 mU/mL, 338.72 mU/mL and 441.96 mU/mL.
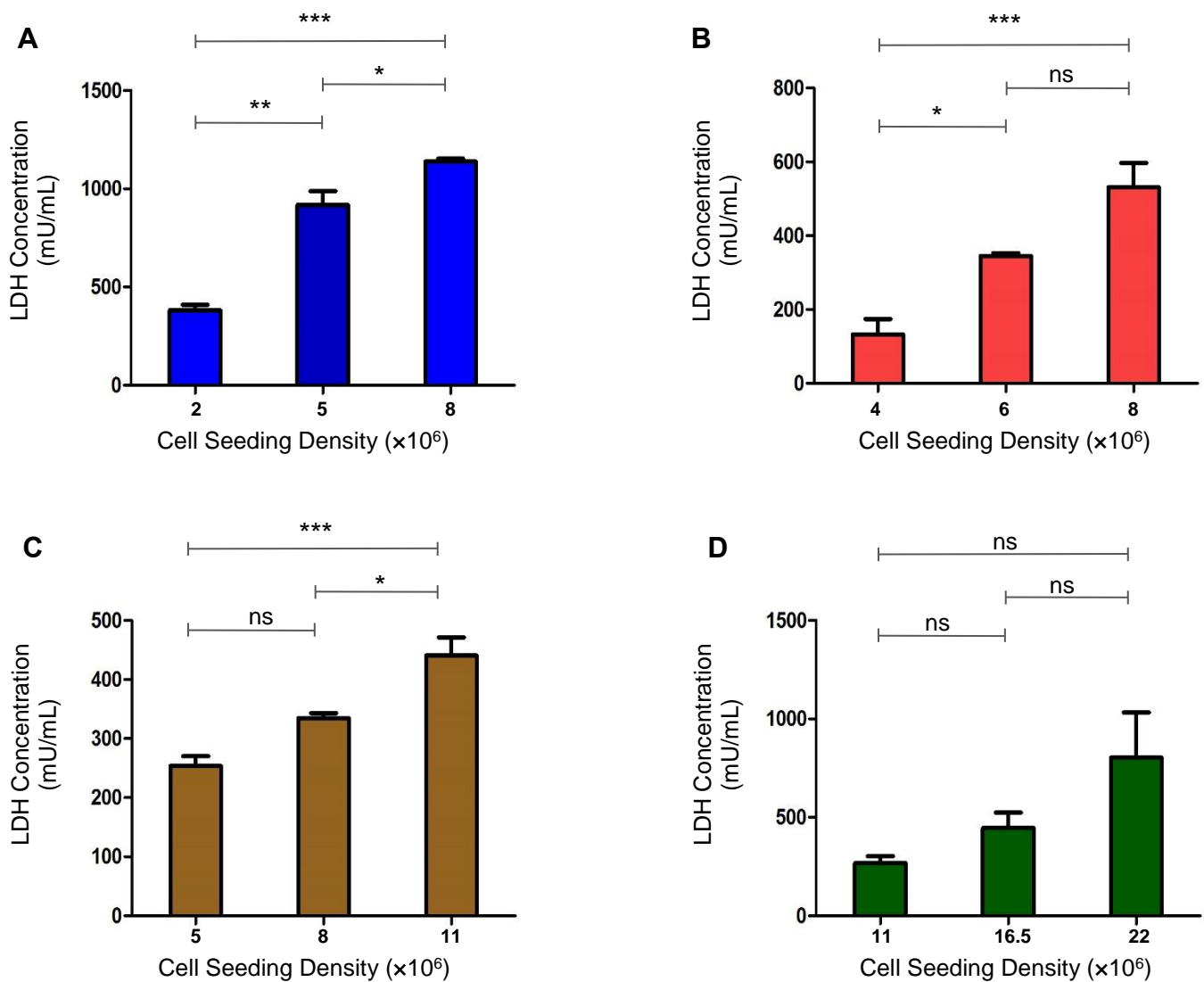
Figure 3.2: LDH concentrations in the serum free media of four prostate cell lines. Concentration of LDH (A) BPH-1 (B) PNT2 (C) PC3 and (D) LNCaP at different seeding densities. Results (A), (B), (C) and (D) are shown as means ± SEM of three independent experiments performed in triplicate. One-way analysis of variance (ANOVA) with Tukey's post-test were used for group comparisons. *P < 0.05; **P < 0.01; ***P < 0.001; ns= not significant.

The most informative way to assess the seeding densities may be to compare the ratio of LDH to protein concentration (figure 3.3). The smallest LDH/protein ratio would indicate the optimal samples. The ratio of LDH/protein for the cell lines are listed below in order of increasing seeding density: BPH-1- 27.44 mU/µg, 29.07 mU/µg and 27.57 mU/µg; PNT2- 7.54 mU/µg, 15.29 mU/µg and 17.21 mU/µg; LNCaP- 17.67 mU/µg, 19.23 mU/µg and 23.81mU/µg; PC3- 19.43 mU/µg, 17.63 mU/µg and 17.47mU/µg. No statistically significant differences were observed in the LDH/Protein ratios with the exception of PNT2 (4×10^6 cells and 8×10^6 cells). As a result of the observation that the LDH/Protein ratios were similar across the seeding densities in most cases it was decided to instead consider the LDH values as LDH is a reliable indicator of the amount of intracellular proteins in each sample and intracellular proteins are known to have a masking effect on low abundant secreted proteins during LC-MS/MS. For this reason the lowest seeding density for each cell line was selected as the optimal seeding density (2×10^6, 4×10^6, 11×10^6 and 5×10^6 cells for BPH, PNT2, LNCaP and PC3 cell lines, respectively).
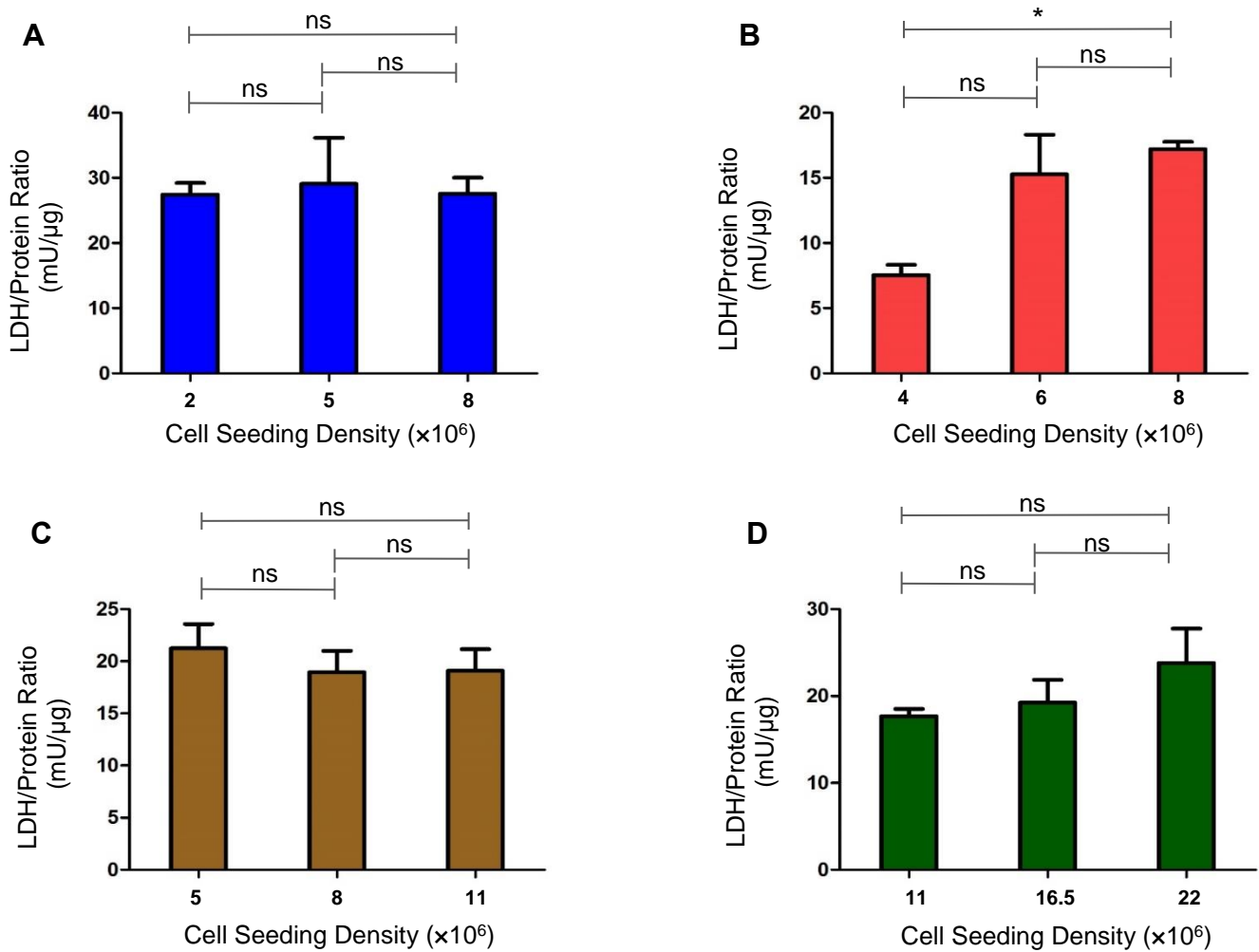
Figure 3.3: Ratio of LDH to total protein concentration of conditioned media from four prostate cell lines at different seeding densities. Ratios were as follows: (A) BPH-1 (B) PNT2 (C) PC3 and (D) LNCaP at different seeding densities.   Results (A), (B), (C) and (D) are shown as means ± SEM of three independent experiments performed in triplicate. One-way analysis of variance (ANOVA) with Tukey's post-test were used for group comparisons. *P < 0.05; ns= not significant.

3.2 Optimization of protein precipitation methods

Protein precipitation is commonly employed as an initial protein purification step when investigating secreted proteins in conditioned media. It was therefore important to evaluate different protein precipitation techniques to determine the method of choice for maximizing protein recovery from conditioned media. The ammonium sulphate, methanol-chloroform, acetone and acetone-phosphotungstic acid protein precipitation methods were compared in this study. These methods exert different modes of action that facilitate varying amounts of protein precipitation for downstream LC-MS/MS.

Methanol-chloroform precipitation extracts proteins from solution by exploiting the principle of phase separation of a polar (methanol) and non-polar (chloroform) solvent. The addition of water to this solvent mixture induces phase separation into a hydrophobic and hydrophilic layer. Precipitated proteins collect at the interface of the two solvents known as the interphase. Ammonium sulphate precipitation removes the hydration layer of water molecules associated with charged side chains and polar groups on the protein surface. The hydration layer is crucial for maintaining solubility and correct conformation of proteins in solution. The addition of salt increases the surface tension of water increasing the interaction of hydrophobic groups on the protein and water. This results in hydrophobic groups associating leading to precipitation (151). Acetone precipitation relies on a similar principle of hydrophobic segregation to achieve protein precipitation. The addition of increasing amounts of acetone reduces the dielectric constant of the solvent thereby reducing its polarity. The reduction in solvent polarity coincides with a reduction in solubility of the protein which effects precipitation. In addition to disruption of the solvation layer, acetone also partially denatures the protein which exposes more hydrophobic groups resulting in increased protein aggregation. Acetone can also be used in combination with phosphotungstic acid (PTA) for protein precipitation. PTA is a heteropoly acid ($H_3PW_{12}O_{40}$) used as an alkaloidal protein precipitating agent. It lowers the pH of a medium resulting in the majority of proteins carrying a net positive charge. Protein cations in the medium form complexes with negatively charged acid molecules to form a protein-tungstate flocculent precipitate.

Protein precipitation experiments were carried out on pooled secretome samples from the BPH-1 and PNT2 cell lines (described in section 2.4). The methanol-chloroform technique failed to recover detectable levels of protein from the conditioned media. Acetone based techniques delivered greater protein recovery in comparison to the ammonium sulphate and acetone-PTA precipitation methods (figure 3.4). Differences between the amounts of protein recovered by each precipitation method for both BPH-1 and PNT2 were statistically significant. In the conditioned media of BPH-1 there was 28-fold increase in the amount of protein recovered by acetone relative to ammonium sulphate ($P < 0.001$), and a 19-fold increase in the protein recovered by acetone-PTA relative to ammonium sulphate ($P < 0.001$). This trend in protein recovery was similar when the experiment was repeated on the conditioned media of PNT2. There was a 16-fold increase in the amount of protein recovered by acetone relative to ammonium sulphate ($P < 0.001$), and a 13-fold increase in the amount of protein recovered by acetone-PTA relative to ammonium sulphate ($P < 0.001$). The amount of protein recovered from the conditioned

media of BPH-1 cell line was 22.5 mg, 15 mg and 0.78 mg for acetone, acetone-PTA and ammonium sulphate precipitation method, respectively. The amount of protein recovered from the conditioned media of PNT2 cell line was 18.23 mg, 15 mg and 1.13 mg for acetone, acetone-PTA and ammonium sulphate precipitation methods, respectively.
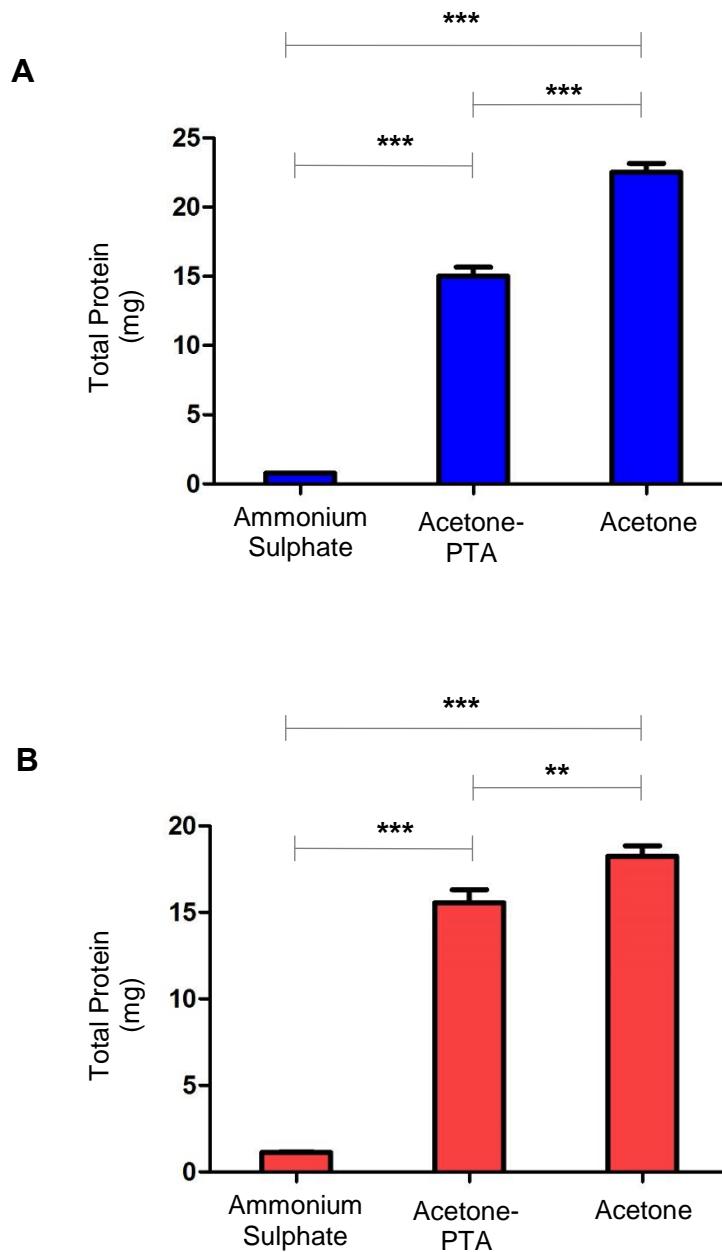

Figure 3.4: Total protein (mg) precipitated from conditioned media of the (A) BPH-1 and (B) PNT2 cell lines. Results are shown as means ± SD of a single experiments measured triplicate. One-way analysis of variance (ANOVA) with Tukey's post-test were used for group comparisons. *P < 0.05; **P < 0.01; ***P < 0.001.

3.3 Characterization of the proteome

The proteome of four prostate cell lines was analysed using LC-MS/MS. Three independent biological replicates were prepared for each cell line and each replicate was analysed in triplicate. X!Tandem and Sequest search engines were used to identify individual peptides. This data was subsequently imported into Scaffold (version Scaffold_4.4.8, Proteome Software Inc., Portland, OR) which was used for subsequent protein assignment. Positive peptide identification was accepted at a cut-off threshold of 90% probability. Peptide spectrum matches were generated in X!Tandem and subsequently validated and filtered by the Peptide Prophet algorithm. Peptide spectrum matches generated by Sequest were subsequently processed by Peptide Prophet and Protein Prophet. The peptides identified by both search engines were consolidated in Scaffold and used for protein identification. Positive protein identifications were accepted at an FDR of less than 1.0% and if they contained at least 2 peptides. The FDR was determined by searching the peptide assignments against a concatenated target decoy database using the same search parameters and Scaffold cut-offs. The total number of proteins assigned per cell line were as follows: 2079 in BPH-1, 2081 in LNCaP, 1853 in PC3 and 2137 proteins in the PNT2 cell line. A total of 655 proteins shown in figure 3.6 (A) were common to all cell lines and the number of unique proteins identified in each cell line were as follows: BPH-1; 473 (23%), LNCaP; 553 (30%), PC3; 472 (22%), PNT2; (565) 26%.
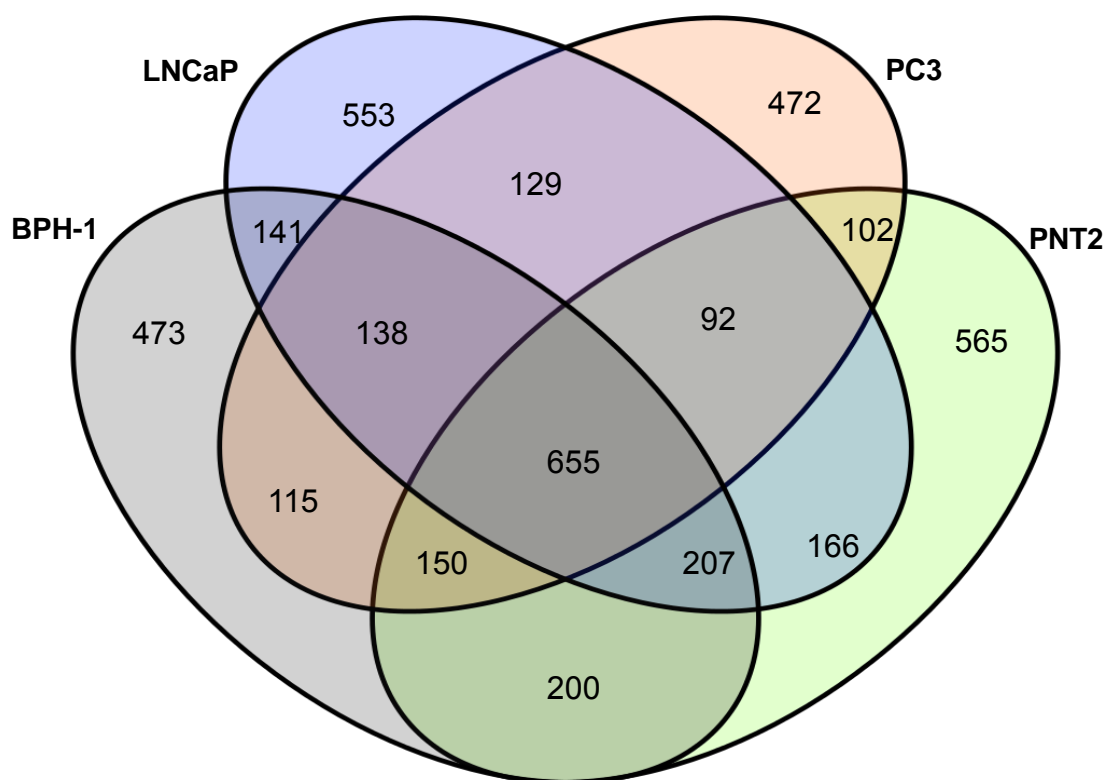
Figure 3.5: Venn diagram showing the overlap of all proteins identified in the BPH-1, LNCaP, PC3 and PNT2 cell lines. Proteins were positively identified with an FDR of 1.0% and a minimum of two peptides.

3.3.1 Subcellular localization of identified proteins

The proteins identified in all cell lines were cross referenced with the Gene Ontology database. The cellular localizations that the proteins were designated included extracellular, intracellular, endoplasmic reticulum, Golgi apparatus, nucleus, endosome, and membrane (figure 3.6). Proteins that did not possess a Gene Ontology classification or did not belong to any of the given cellular localizations were listed as being unclassified. Interestingly, the number of unclassified proteins accounted for a large percentage of the identified proteins: BPH-1; 44%, PC3; 12%, PNT2; 42%, LNCaP; 35%. Intracellular and nuclear proteins were among the top three annotated proteins ranging from 14-18% for the nucleus and 19-26% for intracellular proteins. Extracellular proteins accounted for 7% of annotated proteins in BPH-1, PNT2 and LNCaP, and 19% in PC3. The distribution of membrane proteins was as follows: BPH-1; 9%, PC3; 15%, PNT2; 10%, LNCaP; 11%. The endoplasmic reticulum, Golgi apparatus, mitochondria and endosome accounted for a relatively small percentage of annotated proteins ranging from 1-5% across the cell lines. It should also be noted that there may be some redundancy in the classification of proteins by subcellular location as proteins can be listed under more than one sub cellular location.

BPH-1

Endosome; 1%
Mitochondria; 3%
Golgi Apparatus; 1%
Endoplasmic reticulum; 1%
Unclassified; 44%
Nucleus; 14%
Intracellular; 19%
Membrane; 9%
Extracellular; 7%

PNT2

Endosome; 1%
Mitochondria; 3%
Golgi Apparatus; 2%
Endoplasmic reticulum; 2%
Unclassified; 42%
Nucleus; 14%
Intracellular; 19%
Membrane; 10%
Extracellular; 7%

PC3

Golgi Apparatus; 1%
Endosome; 1%
Mitochondria; 5%
Endoplasmic reticulum; 2%
Unclassified; 12%
Nucleus; 18%
Extracellular; 19%
Intracellular; 26%
Membrane; 15%

LNCaP

Endosome; 1%
Mitochondria; 3%
Golgi Apparatus; 2%
Endoplasmic reticulum; 2%
Nucleus; 15%
Unclassified; 35%
Intracellular; 22%
Extracellular; 7%
Membrane; 11%

Figure 3.6: The subcellular cellular distribution of proteins positively identified in the BPH-1, LNCaP, PC3 and PNT2 cell lines.

53

Extracellular and membrane bound proteins are the most likely proteins to be detected in the secretome and were thus analysed further to identify candidate biomarkers for BPH (figure 3.7). Analysis of this sub proteome identified 26 proteins common to all cell lines. The number of unique proteins were as follows: BPH-1; 110 (27%), LNCaP; 86 (33%), PC3; 180 (70%), PNT2; 25 (8%). A literature review of the 110 proteins that were unique to BPH-1 identified mesencephalic astrocyte-derived neurotrophic factor (MANF) as a candidate biomarker. All other proteins were excluded as potential biomarkers for BPH as they had previously been associated with prostate cancer and/or other cancer types.. Refer to table G in the appendix for a comprehensive list of the positively identified proteins unique to BPH-1 in this sub proteome.
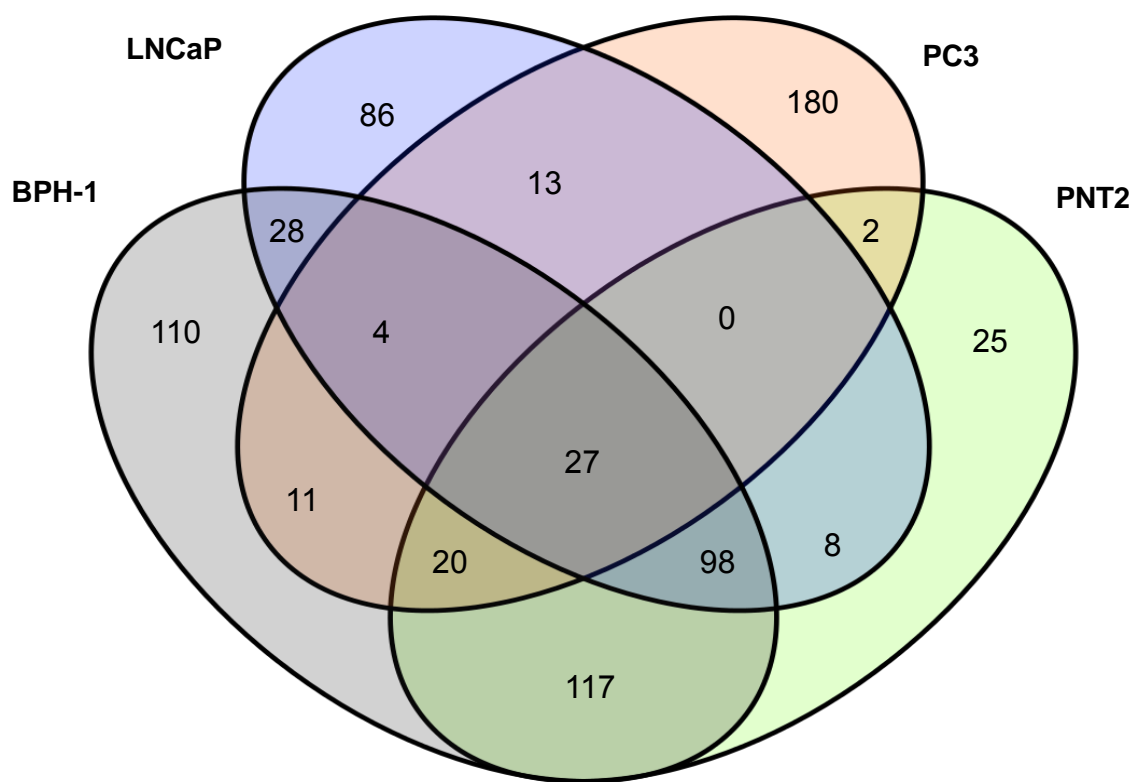

Figure 3.7: Venn diagram showing the overlap of extracellular and membrane bound proteins identified in the BPH-1, LNCaP, PC3 and PNT2 cell lines. Proteins were positively identified with an FDR of 1.0% and a minimum of two peptides.

3.3.2 Pathway analysis

Comparative analysis of the cell secretome was the first approach that was used to identify biomarkers in this study and only yielded one candidate tumour marker, namely MANF. The second approach focussed on identifying pathways that were altered in the BPH disease state. Pathway analysis was carried out using the Ingenuity Pathway Analysis (IPA) software package http://www.ingenuity.com). IPA is a database that serves a repository for biological, functional annotations and pathway models for genes, proteins, cell lines and tissue. Using this software, the experimental data set was cross referenced with the Ingenuity Knowledge base and mapped to the most relevant canonical pathways. A comparative analysis was used to evaluate the relative differences in pathways between two disease states of the prostate with normal prostate cells represented by the PNT2 cell line acting as a reference for fold changes in the other prostate cell lines. Fold changes were first determined in Scaffold prior to IPA analysis and only fold changes with a p-value of less than 0.05 were included in the subsequent IPA analysis. Protein identities from the experimental datasets were mapped against the IPA database and the most highly matched canonical pathways were displayed in order of statistical significance. The statistical significance is expressed as $-\log(p)$ which is calculated by the Fisher's exact test right-tailed. Therefore, the greater the $-\log(p)$ value, the more statistically significant the matched canonical pathway.

The two most variant canonical pathways in BPH are shown in figure 3.8. The overlap between dataset proteins and these canonical pathways where as follows: pyrimidine ribonucleotides de novo synthesis, 21.1%; antiproliferative role of TOB in T cell signalling, 16.7%. Refer to tables A-E in the appendix for log ratios and p-values for proteins mapped to the 5 canonical pathways shown in figure 3.8.
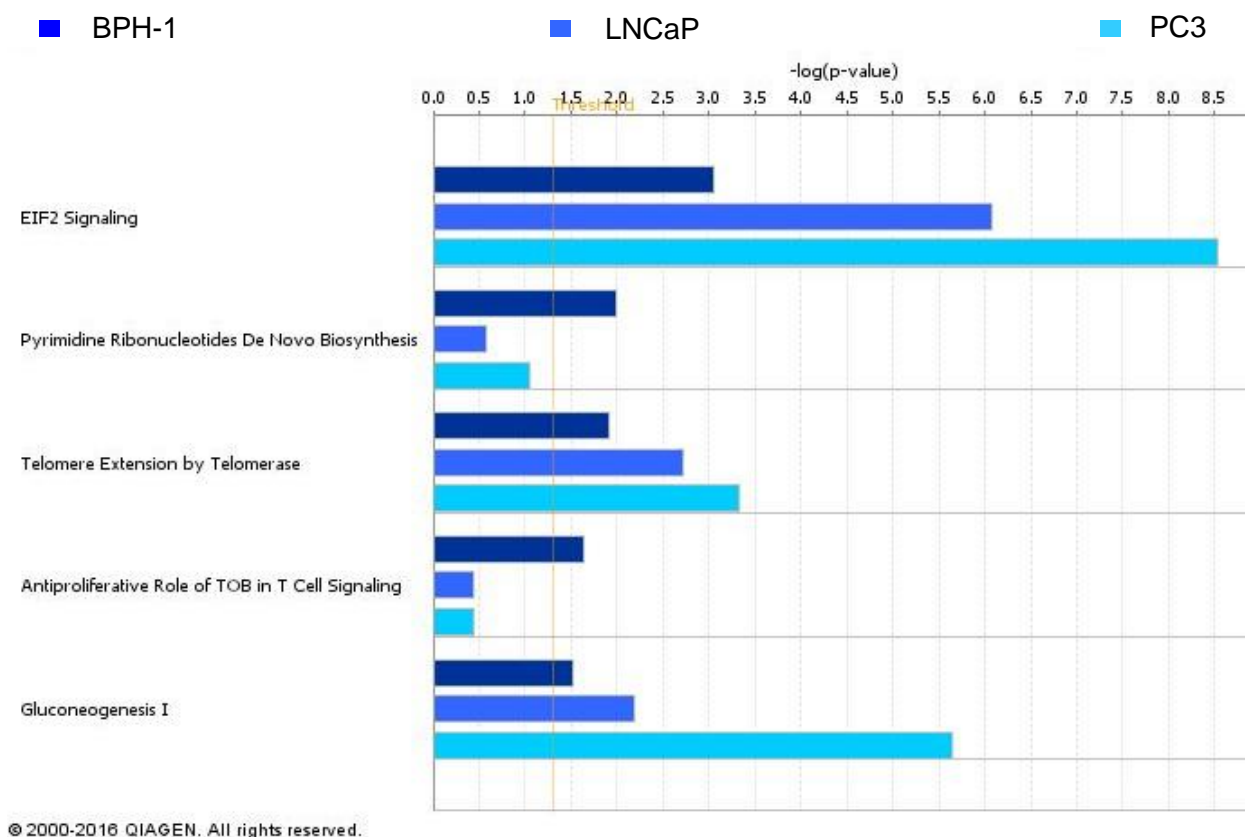
Figure 3.8: Canonical pathways from the IPA database that showed the greatest difference between BPH-1, PC3 and LNCaP datasets. The significance of this overlap was measured as a ratio of the number of molecules from the dataset that map to the pathway divided by the total number of proteins associated with that particular pathway. Fisher's exact test was used to calculate the statistical significance of proteins from the datastet that were matched to the canonical. A p-value of 0.01 was considered significant. Threshold p < 0.01 is shown as yellow line. Bars that are above the line indicate significant enrichment of a pathway.

3.3.3 Disease and function pathway analysis

Following the analysis of canonical pathways, disease and functions associated with BPH were analysed by IPA as described in section 3.3.2. A Fisher's exact test, set at a threshold of 0.01, was then carried out to determine the probability of overlap between the input gene set and the genes present in each annotated disease/function. Vitamin and mineral metabolism, and cell signalling are cell functions that showed significant association with BPH-1 exclusively (figure 3.9). The HSPA5 gene was mapped to both functions. HSPA5 encodes Heat shock 70-kDa protein 5 which is involved in calcium homeostasis and acts as signal receptor localized in the plasma membrane (152,153).
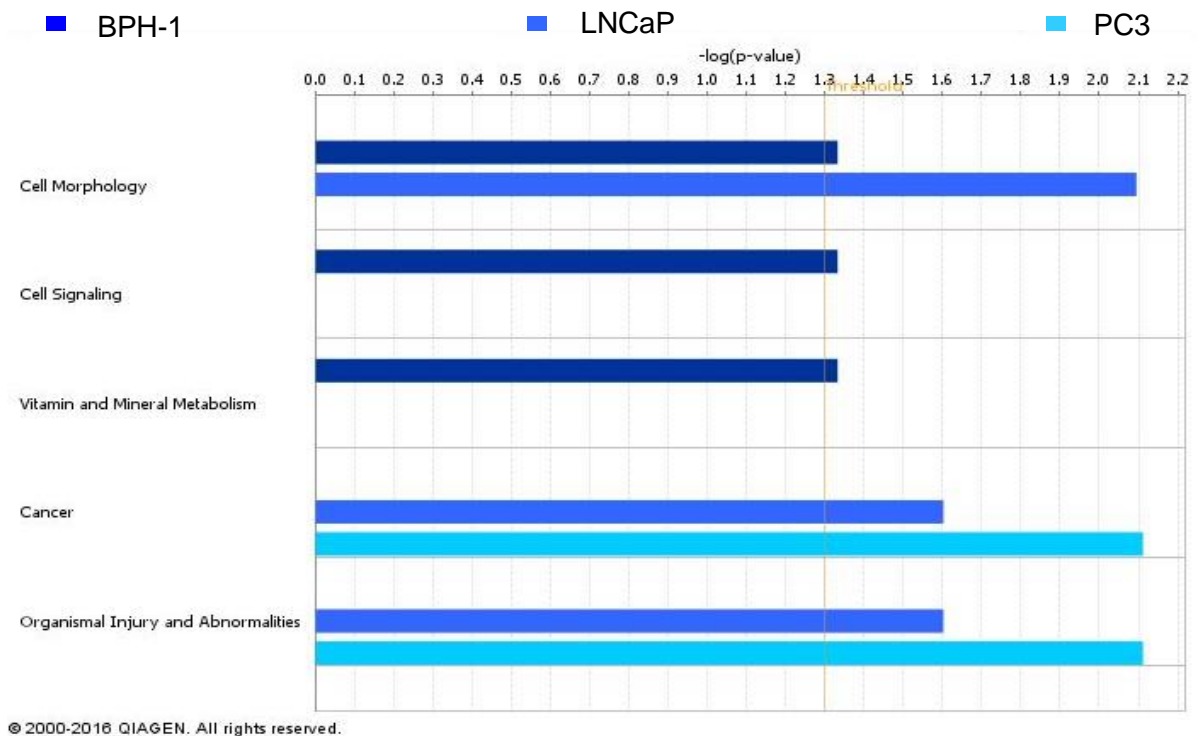
Figure 3.9: Comparison of disease and functions from the BPH-1, PNT2 and LNCaP dataset. The significance of this overlap was measured as a ratio of the number of molecules from the dataset that map to the pathway divided by the total number of proteins associated with that particular pathway. Fisher's exact test was used to calculate the statistical significance of proteins from the datastet that were matched to the canonical pathwyas. A p-value of 0.01 was considered significant.

3.3.4 Network analysis

Canonical pathway analysis together with disease and function analysis revealed the pathways that were altered in the BPH disease state. In order to visualize the interaction of proteins in the BPH-1 dataset and gain an understanding of how they contribute to the pathobiology of BPH IPA network analysis was utilized. Network analysis using IPA graphically displays nodes (gene products) and edges (biological relationships between the nodes). The colour intensity of the nodes indicates the extent of up or down-regulation. IPA identified networks associated with proteins in the datasets of the three cell lines. These networks were scored and displayed in order of highest score (table F in appendix). Each network was listed with its respective top functions and diseases. To gain an understanding of the interaction of these molecules IPA generated networks display the connections between the related proteins.

The top scoring molecular networks in the BPH-1 dataset shown in figure 3.10 below. Diseases and functions associated with the highest scoring network include: cellular growth and proliferation, gene expression (figure 3.10A). This network overlapped with 12 proteins from the BPH-1 dataset and had a

57

score of 7. Upregulated molecules included CDCA5, KRT18, NCOR2 and PRDX1. Downregulated molecules included CALU, CTNNBIP1, ENO1, HNRNPA1, HNRNPK, HSP90AB1, HSPA8, STK4 and TACC2. The second scoring molecular network was associated with cell death and survival, cell development, cellular growth and proliferation (figure 3.10B). This network overlapped with 13 proteins from the BPH-1 dataset and had a score of 6. Upregulated molecules included AGR2, MIF, NUMA1, PDLIM7, TPD52 and TPX2. Downregulated molecules included CD44, CDH1, CDKN1B, IL1RN, MCMC and PA2G4. The third scoring BPH-1 network was mapped to the following diseases and functions: cell death and survival, organismal injury and abnormalities, neurological disease (figure 3.10C). This network overlapped with 6 proteins from the BPH-1 dataset and had a score of 5. Downregulated molecules included HSPA5, HSPD1, ID1, ITGA6, ITGB4 and XRC66.

A
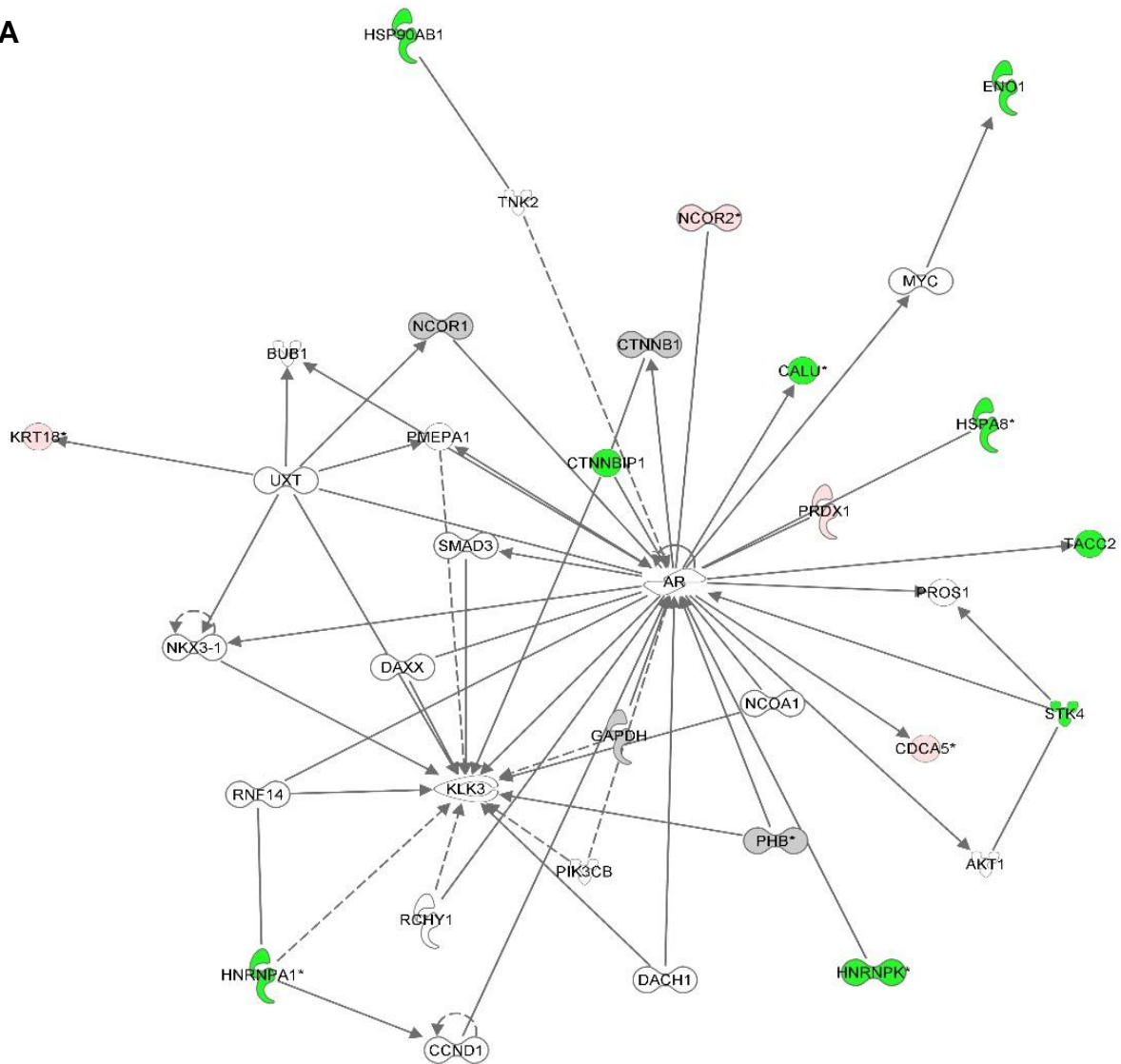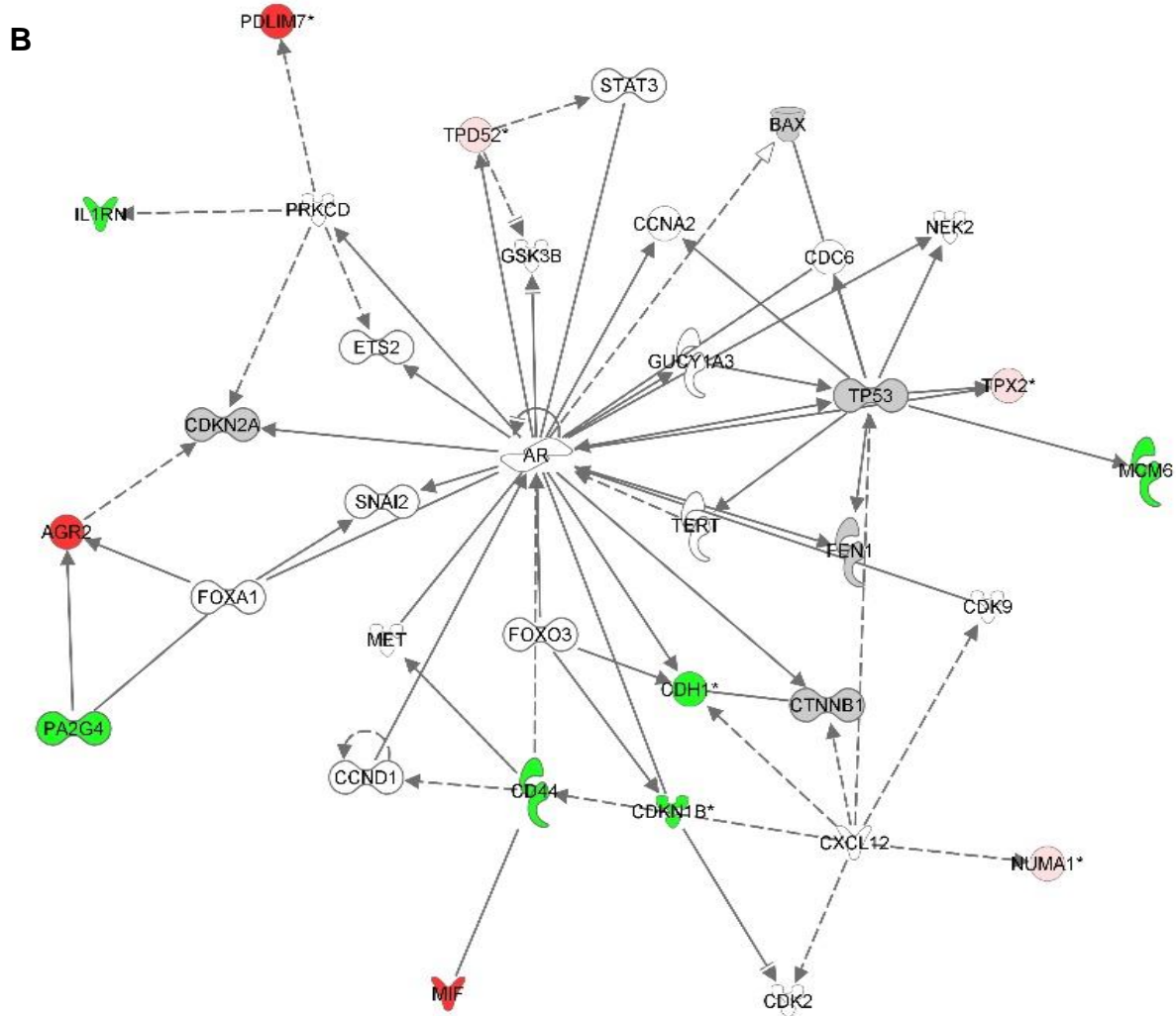
B

C
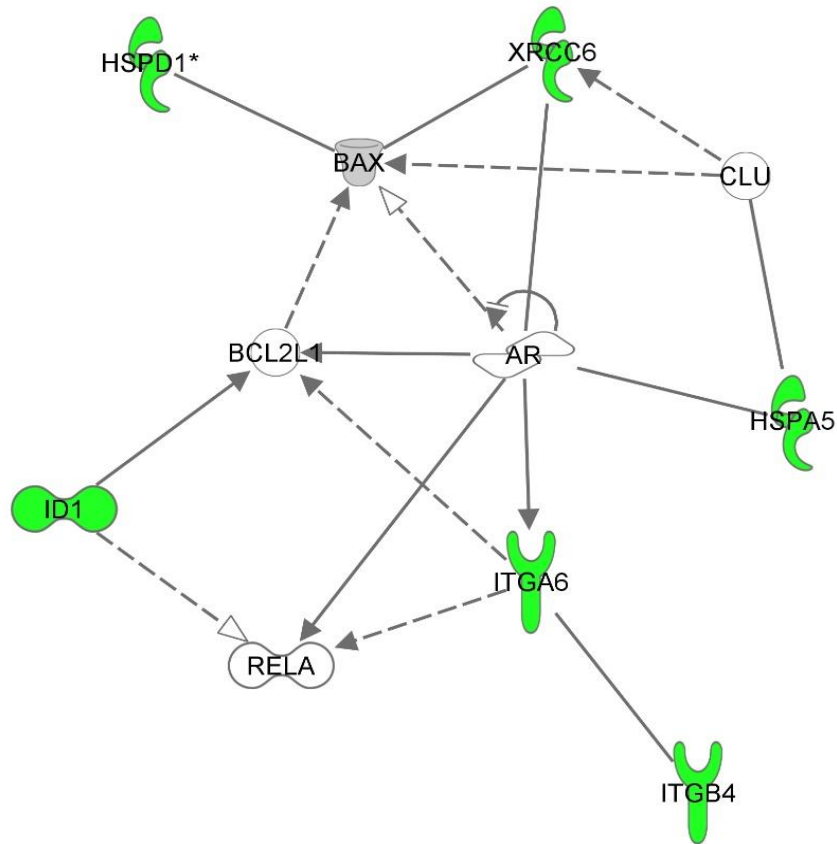
Figure 3.10: Biological network analysis of BPH-1 gene products showing molecular interactions in (A) AR- KLK-3 related pathway (B) AR related pathway (C) AR- BCL-2 related pathway. Lines indicate relationships and the arrows at the end of these lines show the direction of the interaction. Lines without arrowheads show a binding interaction. Dotted lines indicate an indirect interaction. Up-regulated molecules are coloured red while down regulated molecules are coloured green. Molecules coloured grey do not meet the cut-off threshold while molecules coloured white are added from the Ingenuity Knowledge Base.

3.4 Analysis of the secretome

Secretome samples were collected from four prostate cell lines following the optimisation of seeding density described above in section 2.2.3. The samples, which included three independent biological replicates were all prepared using acetone precipitation. Unfortunately, these samples could not be analysed by LC-MS/MS in time for the submission of this thesis due to technical difficulties. Construction in the building in which the mass spectrometer is housed resulted in the instrument being taken off line until such time that the renovations are completed and the instrument serviced. However, secretome preliminary secretome samples from the BPH-1 and PC3 cell lines were previously analysed by LC-MS/MS in order to investigate the proof of concept. It should, however be noted that these only included a single biological replicate and the proteins were precipitated by ammonium sulphate. A total of 120 proteins were identified in BPH-1 and 173 proteins were identified in PC3 secretome. Of these 46 proteins (table 3.1) were unique to BPH-1 while 99 proteins were unique to PC3. The proteins that were differentially expressed in BPH-1 were uploaded on IPA and analysed using the biomarker filter feature to identify possible biomarkers for BPH. However, the biomarker search did not yield any novel or existing putative biomarkers for BPH. Several putative PCa biomarkers such as galectin-3, IL-6 and α-fetoprotein were identified in the BPH-1 secretome which potentially renders them unsuitable as diagnostic tools for the detection of PCa.
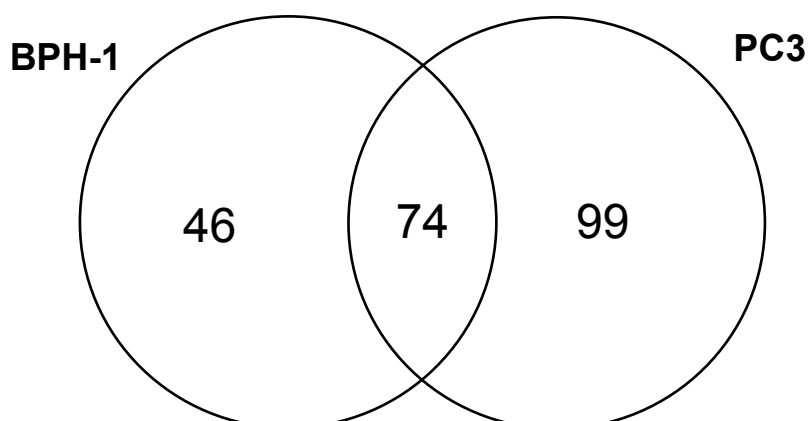

Figure 3.11: Comparison of the cell secretome between BPH-1 and PC3 cell line. Each circle in the Venn diagram represents a cell line and the relative abundance of secreted protein unique to each cell as well as the overlap of proteins between the two cell lines.

61

Table 3.1: Proteins identified in the BPH-1 cell secretome, but not the PC3 secretome.

Uniprot ID	Protein Name	Gene Name
Q9UBP4	Dickkopf-related protein 3	DKK3 REIC UNQ258/PRO295
P00533	Epidermal growth factor receptor	EGFR ERBB ERBB1 HER1
Q12904	Aminoacyl tRNA synthase complex-interacting multifunctional protein 1	AIMP1 EMAP2 SCYE1
Q9GZN4	Brain-specific serine protease 4	PRSS22 BSSP4 PRSS26 SP001LA UNQ302/PRO343
Q8IUL8	Cartilage intermediate layer protein 2	CILP2
O95084	Serine protease 23	PRSS23 ZSIG13 UNQ270/PRO307
Q29960	HLA class I histocompatibility antigen, Cw-16 α chain	HLA-C HLAC
P05155	Plasma protease C1 inhibitor	SERPING1 C1IN C1NH
Q07021	Complement component 1 Q subcomponent-binding protein, mitochondrial	C1QBP GC1QBP HABP1 SF2P32
Q15582	Transforming growth factor-β-induced protein ig-h3	TGFBI BIGH3
P05543	Thyroxine-binding globulin	SERPINA7 TBG
P61812	Transforming growth factor β-2	TGFB2
Q14767	Latent-transforming growth factor β-binding protein 2	LTBP2 C14orf141 LTBP3
Q96EE4	Coiled-coil domain-containing protein 126	CCDC126 UNQ786/PRO1605
Q92563	Testican-2	SPOCK2 KIAA0275 TICN2 UNQ269/PRO306
P25311	Zinc-α-2-glycoprotein	AZGP1 ZAG ZNGP1
Q13751	Laminin subunit β-3	LAMB3 LAMNB1
Q14508	WAP four-disulfide core domain protein 2	WFDC2 HE4 WAP5
P00813	Adenosine deaminase	ADA ADA1
Q16557	Pregnancy-specific β-1-glycoprotein 3	PSG3

Q6UY14	ADAMTS-like protein 4	ADAMTSL4 TSRC1 PP1396 UNQ2803/PRO34012
P01024	Complement C3	C3 CPAMD1
Q12805	EGF-containing fibulin-like extracellular matrix protein 1	EFEMP1 FBLN3 FBNL
P14780	Matrix metalloproteinase-9	MMP9 CLG4B
P13497	Bone morphogenetic protein 1	BMP1 PCOLC
Q8NCW5	NAD	NAXE AIBP APOA1BP YJEFN1
P15151	Poliovirus receptor	PVR PVS
O60687	Sushi repeat-containing protein SRPX2	SRPX2 SRPUL
Q13421	Mesothelin	MSLN MPF
Q9Y337	Kallikrein-5	KLK5 SCTE UNQ570/PRO1132
P17936	Insulin-like growth factor-binding protein 3	IGFBP3 IBP3
P20908	Collagen α-1	COL5A1
P17931	Galectin-3	LGALS3 MAC2
A6NKQ9	Choriogonadotropin subunit β-variant 1	CGB1
Q99523	Sortilin	SORT1
O00300	Tumor necrosis factor receptor superfamily member 11B	TNFRSF11B OCIF OPG
P05231	Interleukin-6	IL6 IFNB2
Q92876	Kallikrein-6	KLK6 PRSS18 PRSS9
P36952	Serpin B5	SERPINB5 PI5
P01008	Antithrombin-III	SERPINC1 AT3 PRO0309
Q76M96	Coiled-coil domain-containing protein 80	CCDC80 DRO1 URB HBE245
P02765	α-2-HS-glycoprotein	AHSG FETUA PRO2743
Q15828	Cystatin-M	CST6
P02771	α-fetoprotein	AFP HPAFP
Q8WVN6	Secreted and transmembrane protein 1	SECTM1 K12
P13521	Secretogranin-2	SCG2 CHGC

Chapter 4

Discussion

Prostate cancer is the second most prevalent non-cutaneous malignancy affecting men in the world (34). The current method for screening and diagnosis of prostate cancer relies on the measurement of serum levels of kallikrein 3 (KLK3) in conjunction with the digital rectal exam. However, serum KLK3, or prostate specific antigen (PSA) as it is better known, is also elevated by BPH (2). Consequently, a major challenge in the treatment and management of this disease is the ability to discriminate between prostate cancer and benign tumours such as BPH. Misdiagnosis of prostate cancer entails unnecessary treatment for patients which culminates in psychological stress and physical discomfort (2). Although PSA has been shown to be ineffective as a true biomarker, it is still used as the primary diagnostic tool for the detection of prostate cancer. This is largely due to lack of an alternative and it is therefore crucial to find a new biomarker which can distinguish malignant prostate cancer from benign prostatic hyperplasia.

4.1 Secretome optimization

In this study the proteome and secretomes of four prostate cell lines, representing different conditions, were investigated with the aim of identifying novel biomarkers which could distinguish BPH from PCa. The PC3 and LNCaP cell lines were used to represent malignant prostate tumours, while the non-transformed epithelial prostate cell line, BPH-1, was used to represent BPH. The normal prostate epithelial cell line, PNT2, was used as a healthy control for all experiments.

Analysis of conditioned serum free media from different prostate cell lines offers an attractive method for biomarker discovery for several reasons. Firstly, serum free culture media is relatively less complex when compared to human serum, which is highly complex with dynamic protein concentrations ranging up to 12 orders of magnitude. The highly abundant proteins often prevent the ionization and subsequent identification of low abundant secretome proteins during MS. Secondly, the biological homogeneity of each cell line translates to reproducible results relative to tissue or serum samples. Lastly, prostate cell lines are easy to grow and maintain which greatly reduces sample preparation time. The latter property of prostate cell lines coupled to the simple complexity of serum free media results in rapid MS profiling of the cancer proteomes and secretomes.

Despite the use of serum free conditioned media offering several advantages for biomarker studies, this strategy it is not without its challenges. Intracellular proteins are released into the conditioned media at the onset of cell death when the cell membrane ruptures. These intracellular proteins are several times more abundant than secreted proteins and therefore can have a masking effect on the secretome during MS analysis. Growth conditions of the prostate cell lines must therefore be optimized in order to reduce cell death and maximize secreted proteins. LDH and total protein levels in the secretome were

therefore measured from three different seeding densities for each cell line (figures 3.1 and 3.2). A low LDH concentration indicates less cell death which correlates to less intracellular proteins in a sample. The ratio of LDH/Protein was then calculated to determine the sample with less intracellular proteins and more secreted proteins. Therefore, both the LDH/Protein ratio and LDH concentration were evaluated and taken into consideration for the determination of the optimal seeding density to increase protein coverage during MS analysis. Despite the lowest seeding densities not always having the lowest ratio value they were selected for LC-MS/MS analysis due to their relatively low LDH concentration (figure 3.3). The selected seeding densities were as follows: BPH-1; 2×10^6, LNCaP; 11×10^6, PNT2; 4×10^6, PC3; 5×10^6. In a previous study of the cell secretome by Sardana *et al.* the optimal seeding densities for PC3 and LNCaP were 7.5×10^6 and 22×10^6 cells respectively (2). It must be noted the latter study used a roller bottle cell culturing method which greatly reduced the amount of cell death and therefore allowed for higher seeding densities.

4.2 Protein purification optimization

Four different protein precipitation techniques were evaluated for their effectiveness in recovering protein from the conditioned serum free medium (figure 3.4). The chloroform-methanol method yielded poor results due to difficulty in isolating small amounts of protein from the solvent interphase as well as the difficulty of handling the large solvent volumes required by this method. This method was therefore excluded from further analyses. When comparing the remaining methods, the acetone based precipitation techniques (acetone and acetone-PTA) resulted in up to 20-fold higher protein recoveries than the ammonium sulphate precipitation technique. This large disparity in the effectiveness of the methods can be attributed to the biochemical modes of precipitation (described in section 3.2). Acetone based techniques remove solvation layers around proteins as well as denature the proteins leading to greater association of hydrophobic groups and larger protein aggregates which are able to trap smaller proteins when they associate. Conversely, the ammonium sulphate method relies solely on the removal of solvation around proteins. Therefore, the dual mode of action of acetone precipitation clearly results in greater protein recovery from conditioned media. Acetone precipitation also offers a more a more attractive method as it minimizes keratin contamination which is an inevitable consequence of sample preparation and reduces protein identification. Interestingly, acetone used in combination with PTA resulted in a smaller amount of recovered protein relative to the PTA free acetone precipitation method. The smaller protein yield is likely due to the protein being difficult to re-suspend in deionized water subsequent to precipitation and residual PTA interfering with the BCA protein determination assay. Cell secretome studies by Sardana *et al.* and Kulasingham *et al.* both employed ammonium sulphate precipitation to isolate protein from the conditioned media (2,63)

4.3 Characterization of the proteome

The proteome refers to the total ensemble of protein that constitute a cell. Therefore, the comparative study of prostate cell line proteomes allowed us to decipher the differences between two disease states. The total number of proteins identified per cell line were as follows: 2079 in BPH-1, 2081 in LNCaP, 1853 in PC3 and 2137 proteins in the PNT2 cell line. Approximately 25% of the identified proteins were unique to each cell line. This difference in protein composition among the cell lines highlights the

heterogeneity of these cell lines, even though they all originate from the prostate epithelium. Tumour heterogeneity arises from multiple non-identical mutations occurring in a given tumour. It may therefore be more feasible to identify a panel of biomarkers as opposed to a single biomarker in order to account for this heterogeneity. Another factor taken into consideration in this study was that *in vitro* conditions differ considerably from the physiological environment. This greatly impacts on the composition of the proteome and some of the proteins identified in this study may not be relevant under *in vivo* conditions. It is therefore imperative to subsequently conduct validation studies of candidate tumour markers.

The positively identified proteins were cross referenced on the Gene Ontology database to determine their subcellular localization (figure 3.6). Approximately 49% of annotated proteins were found to be intracellular. Unclassified protein accounted for approximately 40% of the total identified proteins. Although the percentage of membrane and extracellular proteins identified from the proteome were relatively low, these still accounted for (7%) and (11%) of the total number of proteins identified, respectively. Extracellular and membrane bound proteins are most likely to be secreted or shed into circulation and therefore present attractive targets for biomarker discovery as they can be easily obtained from blood serum and subsequently analysed. Candidate biomarkers for BPH were therefore identified by comparative analysis of extracellular and membrane bound proteins in all prostate cell lines used in this study. Of the 110 proteins that were unique to BPH-1 (table G in the appendix) only one protein, mesencephalic astrocyte-derived neurotrophic factor (MANF) was identified as a potential biomarker for BPH. The remainder of the proteins unique to BPH-1 have previously been implicated in PCa and other cancers, and are being investigated as potential biomarkers for those respective diseases. These proteins were disqualified as potential BPH tumour markers as there is a high likelihood they will fail to discriminate between the malignant and benign prostate tumours. Proteins identified in other cancer types were also disqualified as it is not uncommon for malignant cells, regardless of localization or tissue type, to exhibit similar secretome profiles or even share biomarkers.

MANF is a secreted protein that is produced in response to ER related stress and protects somatic cells from inflammatory response damage. In normal mammalian cells the ER triggers the activation of a pathway known as unfolded protein response (UPR) in response to the accumulation of misfolded proteins. This pathway can also be activated in response to ER stress caused by inflammation, and MANF has been shown to be upregulated in patients with inflammatory diseases (154). Inflammation is a key factor in the development of BPH and PCa. The prostate is an immunocompetent organ with over 90% of leucocytes present being CD8+ T lymphocytes (155). The inflammatory response in BPH is thought to be immune mediated by these T-cells in response to external stimuli such as a bacterial infections, viral infections, chemical irritation and sexually transmitted diseases (156). External stimuli trigger an immune response characterized by T-cells releasing cytokines and growth factors which in turn cause damage to prostate cells, chronic inflammation and abnormal remodelling of fibromuscular tissue. IL-8 is one such cytokine released by T-cells and triggers the expression of FGF-2 which stimulates abnormal stromal and epithelial cell growth. Another cytokine implicated in BPH is TGF-β which has been shown to be elevated in BPH tissue (155). Further evidence of the connection between inflammation and BPH was provided by IPA disease and function analysis of BPH-1 which mapped a

T-cell mediated inflammatory response to BPH-1 (discussed in section 4.4.5). Mediators of the inflammatory response in the prostate can therefore potentially serve as a source of biomarkers in BPH. However, while chronic inflammation in the prostate may be a causative factor for the development of BPH it has also implicated in prostate carcinogenesis. In PCa, two inflammatory cytokines have been identified that are thought to have a link to prostate carcinogenesis. Macrophage inhibitory cytokine (MIC) is one such cytokine that is upregulated in PCa. The levels of MIC in circulation are elevated in PCa and it is linked to a poor prognosis in PCa patients. The second cytokine is IL-6 along with its receptor IL-6R which are known to be upregulated in PCa. The levels of IL-6 in circulation are elevated in both metastatic and hormone refractory PCa and it has already been identified as a promising biomarker for PCa. IL-6 is thought to have a contributory role in the initiation and progression of PCa. It has also been detected in malignant tissue and high grade PIN (157). Chronic inflammation is also thought to lead to the development of proliferative inflammatory atrophy (PIA) (158). PIA is characterized by regions of the prostate regenerating at a high frequency in response to tissue damage (157). These atrophic lesions often merge with high grade PIN and are therefore considered to be precursors of prostate cancer (157). Although inflammation has been implicated in both BHP and PCa, there are several delineating factors in the inflammatory response of both diseases that may be able to distinguish them. Firstly, local hypoxia induced in the prostate during BPH triggers prostate cells to secrete growth factors such as FGF-7, TGF-β, IL-8 and FGF-2 (159). Localized hypoxia in the central zone of the prostate is characteristic of BPH and does not commonly occur in PCa. Secondly, proinflammatory cytokines such as IL-17 (overexpressed by T-cells in 79% of BPH patients), IL-5 in stromal cells and IL-7 in infiltrating T-cells have been reported in BPH tissue (155). Elevated levels of these cytokines are unique to BPH while elevated serum IL-6 is unique to PCa. Lastly, the overexpression of cyclooxygenase-2 by PIA which is characteristic of PCa (158). Therefore, the role of chronic inflammation and its infiltrates in the pathogenesis of the prostate could be further explored to develop strategies to detect and treat BPH.

In this study, approximately 75% of the BPH-1 secretome profile was shared with the secretome profiles of either LNCaP, PC3 or PNT2. This indicates that all prostate cells share similar modes of cell signalling to mediate intercellular communication. Furthermore, the similarity in the secretome profile greatly increases the challenge of finding novel biomarkers to distinguish between the BPH and PCa. To circumvent this challenge, we decided to investigate aberrant pathways (described in section 3.3.2) associated with BPH as these can offer insights into the molecular mechanisms of disease progression and present targets for therapeutic intervention.

4.4 IPA pathway analysis

In this study a comparative analysis was used to evaluate the relative differences in pathways between two disease states (figure 3.8 and figure 3.9). All fold changes were measured relative to the normal prostate cell line, PNT2, which was used as a control. Fold changes were first determined in Scaffold prior to IPA analysis and only fold changes with a p-value of less than 0.05 were included in the subsequent IPA analysis. Because proteins do not act in isolation in a biological system network analysis was needed to analyse their complex interactions and therefore provide a more holistic image

of the physiological environment (160). The major advantage of network analysis is that it can be used to integrate different omic technologies. Implementing two or more omic strategies in a biomarker study can essentially cross validate candidate biomarkers. Additionally, integration can reveal otherwise latent biological interactions that can only be elucidated after data analysis of different biochemical entities (161). However, despite the widespread use of network analysis in proteomic studies there is one major drawback. Since the human proteome has not yet been fully characterized the network knowledgebase (in this study IPA) that is used for the analysis only contains a fraction of the full human proteome complement. It is therefore probable that the sample datasets that were uploaded into IPA for Core Analysis and the subsequent networks that were generated differ quite considerably from the global network at the physiological level (162).

Canonical pathways that showed significant alteration in BPH-1 included the antiproliferative role of transducer of ERB2 (TOB) in T cell signalling and pyrimidine ribonucleotide *de novo* biosynthesis pathway (figure 3.8). Diseases and functions that showed significant alteration in BPH included cell signalling and vitamin and mineral metabolism (figure 3.9). Each of these pathways is briefly discussed below.

4.4.1 Antiproliferative role of TOB in T cell signalling

T cells are leucocytes that play a major role in the immune response. An effective T cell response involves cytokine secretion and clonal expansion. The regulation of T cells and their maintenance in an unresponsive state in the absence of antigens is crucial for correct functioning of the immune system. Transducer of ERB2 is a negative regulator of T cell proliferation. It exerts its effect by arresting cell cycle progression and inhibiting transcription of interleukin-2[9]. Chronic inflammation in the prostate gland is known to contribute to the development of BPH (as discussed in section 4.3). Though IPA could not predict the activity state of the antiproliferative role of TOB in T cell signalling it is highly probable it is down regulated due to the elevated expression of T cells in the prostate in both BPH and PCa. The antiproliferative role of TOB in T cell Signalling together with MANF provide further evidence of the association of BPH and inflammation and the prostate inflammatory response should therefore be investigated to identify biomarkers.

4.4.2 Pyrimidine ribonucleotide *de novo* biosynthesis, cell signalling and vitamin and mineral metabolism

The pyrimidine ribonucleotide biosynthesis pathway is an energy consuming pathway that synthesizes precursors for DNA and RNA in the cell. Pyrimidines are six carbon cyclic compounds with a nitrogen atom attached at the 1 and 3 position. Conjugation of a ribose sugar at carbon position 1 on the nitrogenous base results in the formation of a ribonucleoside. Subsequent attachment of one or more phosphate groups to the number 5 position of the ribose sugar forms a pyrimidine ribonucleotide. *De novo* pyrimidine synthesis occurs in the cytoplasm and results in the formation of CTP and UTP. Both nucleotides serve as energy sources with CTP being involved in lipid metabolism while UTP activates

---

[9] https://www.qiagen.com/es/shop/genes-and-pathways/pathway-details/?pwid=443

glucose and galactose (163). BPH-1 showed greater association with this pathway relative to the PCa cell lines. PCa is a slow developing disease relative to other cancer types that mostly exhibits symptoms in its advanced stages[10]. In addition to this malignant prostate cells are localized in the peripheral zone of the prostate and are involved in a host of interactions with neighbouring cells. Conversely, BPH cells are localized in the central zone of the prostate where they replicate indefinitely at a relatively faster rate (164). As a result of the more homogenous cell population and increased cell proliferation BPH cells may require more energy to sustain cell replication as well as nucleic acids which provide the framework for this replication. Network analysis of BPH-1 (figure 3.10) provided further evidence of this hyperproliferation through the following network enriched functions: cell death and survival, cell development, cellular growth and proliferation. The AR was a central node in all networks generated from BPH-1. Dysregulation of the AR is a feature mostly associated with PCa but as both diseases are androgen dependent it is plausible that this also occurs in BPH. This dysregulation of the AR in combination with aberrant cell signalling, and vitamin and mineral metabolism (reported in the disease and function analysis, section 3.3.4) may account for the increased rate of cell proliferation in BPH. The HSPA5 gene was mapped to both cell signalling and vitamin and mineral metabolism, and encodes Heat shock 70-kDa protein 5 (HSPA5). HSPA5 is involved in calcium homeostasis and acts as a signal receptor localized in the plasma membrane. Intracellular calcium (Ca^{2+}) is involved in the regulation of cell processes such as gene expression and cell proliferation. The Ca^{2+} signalling cascade is considerably altered during tumorigenesis to facilitate evasion of apoptosis and increased cell proliferation (152,153) .

4.5 Preliminary secretome analysis

The secretome comprises of proteins that are shed or secreted by cell, tissue or organism. The secretome of prostate cell lines provides a rich source of biomarkers as these proteins are often implicated in signalling processes promoting tumour progression such as proliferation, metastasis, angiogenesis and evasion of apoptosis. The secretome samples that were prepared as described in section 2.2.3 could not be analysed by LC-MS/MS due to technical difficulties beyond our control. Construction in the building in which the mass spectrometer is housed resulted in the instrument being taken off line until such time that the renovations were completed and the instrument serviced. Preliminary comparative analysis of the PC3 and BPH-1 secretome profiles did however identify 46 proteins unique to the BPH-1 cell line. These proteins were uploaded in IPA and analysed by the biomarker feature to identify potential biomarkers. A literature search was also conducted in support of the IPA biomarker filter analysis. The search did not yield any potential biomarkers as most of the proteins were either already identified in similar biomarker studies as potential PCa biomarkers or were implicated in PCa development and progression and would therefore not be suitable to distinguish BPH from PCa. The lack of any viable biomarkers from this preliminary investigation can be attributed to the ammonium sulphate precipitation method that was used for protein purification. As reported in section 3.2 there was a significant fold increase in the amount of protein recovered by acetone precipitation relative to ammonium sulphate. Therefore, the poorer protein purification method that was used in this

---

[10] https://old.cancer.org/acs/groups/cid/documents/webcontent/003182-pdf.pdf

preliminary study may have drastically reduced the amount of secreted proteins that could be recovered and subsequently identified by MS. This preliminary secretome study did however provide proof of concept. Furthermore, the identification of putative PCa biomarkers such as galectin-3, α-fetoprotein and IL-6 in the secretome of BPH-1 cells suggests that many of these proteins would not be suitable biomarkers for PCa as, like PSA, they may not distinguish between PCa and BPH. This therefore illustrates the use of testing secretome samples from representative cell lines in combination with secretomes obtained from serum or urine.

4.6 Conclusion

After optimization of prostate cell line growth conditions, the lowest seeding densities were selected for all samples due to the lower amount of intracellular proteins in these samples as abundant intracellular proteins have a masking effect during MS analysis. Optimization of protein purification identified acetone precipitation as the best method for isolating proteins from serum free media due to its dual mode of action that involved denaturing and removing solvation layers from proteins. Although secretome samples were prepared for analysis by LC-MS/MS, these samples could not be analysed due to technical difficulties and will be will be completed in future.

The proteome of four prostate cell lines were instead analysed by LC-MS/MS to identify potential biomarkers which can be used to distinguish BPH from PCa. The secreted protein MANF was identified as candidate tumour marker in this regard. MANF is a secreted protein associated with inflammation and this further pointed towards the association of BPH with a chronic inflammatory response. It is, however, unlikely that a single prognostic biomarker can be used across different populations due to tumour heterogeneity. The candidate biomarker identified in this study could therefore potentially be used in tandem with current or emerging biomarkers of PCa and BPH. Prior to implementation as a prognostic tool the protein would have to undergo stringent validation trials in animal models and patient serum.

In addition to analysis of differentially expressed proteins in the proteome of the prostate cell lines, we investigated pathways that were affected in BPH relative to PCa. These included Pyrimidine ribonucleotides *de novo* biosynthesis and the antiproliferative role of TOB in T cell signalling. Interestingly both the latter mentioned pathway and candidate biomarker are involved in the inflammatory response. Further investigation of the role of this pathway in BPH as well as an investigation of pathways involved in the immune mediated inflammatory response in prostate cells could yield targets for pharmacological intervention. Pyrimidine ribonucleotide *de novo* biosynthesis and vitamin and mineral metabolism pointed to an increased rate of cell proliferation in BPH relative to PCa and may account for the greater prostate mass measured in BPH patients relative to PCa. Targeting nucleic acid synthesis via the pyrimidine ribonucleotide *de novo* biosynthesis in BPH cells may deprive tumour cells of a valuable energy source and reduce the rate of cell proliferation. Abnormal vitamin and mineral metabolism is linked to dysregulated Ca^{2+} signalling. Since Ca^{2+} signalling is involved in cell survival and proliferation, inhibiting the mitochondrial uptake may provide an effective remedy in BPH treatment.

Preliminary analysis of the secretome did not identify any candidate biomarkers but provided proof of concept. Differential analysis of the secretome may provide a more promising source of biomarkers relative to differential analysis of the cell proteome. This is due to optimization of cell culture growth conditions which effectively concentrates secreted proteins and allows for greater secreted protein coverage by MS.

4.7 Future studies

The comparative analysis of the proteome identified MANF as candidate biomarker for BPH. The clinical utility of this protein as a diagnostic tool to distinguish the benign from malignant prostate cancer is at this stage unknown. Future studies must therefore focus on validating the presence and effectiveness of MANF in patient serum by conducting immunoassays. MS quantification methods could also be used to measure the amount of serum MANF and correlate this to disease stage. The role of MANF in BPH could further be investigated by gene knockout experiments. Pathways that were found to be altered in the BPH disease state by IPA can be further analysed to yield biomarkers by analysis of their metabolic profile. This would entail integration of proteomics and metabolomics. Interpretation of proteomic data through *in silico* network analysis can identify pathways involved in pathogenesis. These pathways can then be further investigated by metabolomics to analyse the interaction of gene and protein downstream products. Furthermore, BPH showed an association with chronic inflammation in the prostate and this can also be further investigated to understand the role this inflammatory response and identify possible biomarkers for BPH.

Due to a technical fault the secretome of all prostate cell lines could not be analysed. Future investigations must augment the analysis of the cell secretome by the addition of several experimental conditions. Firstly, additional prostate cancer cell lines, such as DU145, 22Rv1, VCaP and PPC-1, can be added to the study in order to further enrich the secretome profile of PCa. Secondly, as this study was conducted in the absence of androgens the effect of androgens must be investigated to understand their influence on protein expression. Lastly, prostate cells may be co-cultured with stromal cells and CD8+ cells to in a 3D-culture model to better match physiological conditions and gain insights on the immune mediated inflammatory response in disease aetiology.

References

1. Perez C., Halperin E., Brady L. & Wazer D. 2013. *Principles and Practice of Radiation Oncology.* Lippincott Willaiams & Wilkins. Philadelphia.

2. Girish S. 2008. *Proteomic Analysis of Prostate Cancer Cell Line Conditioned Media for the Discovery of Candidate Biomakers*. University of Toronto

3. Epstein J. & Netto G. 2007. *Biopsy Interpretation of the Prostate.* Philadelphia. Lippincott Willaiams & Wilkins.

4. Nickel C. 1999. *Textbook of Prostatitis.* Oxford: Isis Medical Media.

5. Bolla M. & Van Poppel H. 2012. *Management of Prostate Cancer: A Multidisciplinary Approach.* New York. Springer.

6. Muniyan S, Chaturvedi N.K, Dwyer J.G & Lagrange C.A. 2013. Human Prostatic Acid Phosphatase : Structure, Function and Regulation. *International Journal of Molecular Sciences*, 14(5): 10438-64.

7. Pestell R. & Nevalainen M. 2008. *Prostate Cancer: Signaling Networks, Genetics, and New Treatment Strategies.* New Jersey. Human Press.

8. Warmkessel J. 2000. *Contemporary Issues in Prostate Cancer: A Nursing Perspective.* Boston. Jones and Bartlett.

9. Cunha G.R., Hayward W.S, Wang Y.Z. & Ricke W.A. 2003. Role of the Stromal Microenvironment in Carcinogenesis of the Prostate. *International Journal of Cancer* **107**,1–10.

10. Zhu M-L & Kyprianou N. 2008. Androgen Receptor and Growth Factor Signaling Cross-talk in Prostate Cancer Cells. *Endocrine Related Cancer* **15**,841–9.

11. Pin E., Fredolini C. & Petricoin E.F. 2013. The Role of Proteomics in Prostate Cancer Research: Biomarker discovery and validation. *Clinical Biochemistry* **46**, 524–38.

12. Alruwaili J.A. 2011. *Serum Proteomic Analysis of Prostate Cancer Progression*. University of Portsmouth.

13. Mydlo J. (ed) & Godec C. (ed.) 2003. *Prostate Cancer: Science and Clinical Practice.* San Diego. Academic Press.

14. Gronberg H. 2003. Prostate Cancer Epidemiology. *Lancet* **361**, 859–864.

15. Bostwick D.G., Burke H.B., Djakiew D,, Euling S., Ho S.M., Landolph J, *et al.* 2004. Human Prostate Cancer Risk Factors. *Cancer* **101**, 2371–490.

16. Schrecengost R. & Knudsen K.E. 2013. Molecular pathogenesis and progression of prostate

cancer. *Seminars in Oncology* **40,** 244-258.

17. Montironi R., Mazzucchelli R., Lopez-Beltran A., Cheng L. & Scarpelli M. 2007. Mechanisms of Disease: High-Grade Prostatic Intraepithelial Neoplasia and other Proposed Pre-neoplastic Lesions in the Prostate. *Nature Clinical Practice Urology* **4**,321–32.

18. Demarzo A.M., Nelson W.G., Isaacs W.B. & Epstein J.I. 2003. Pathological and Molecular Aspects of Prostate Cancer. *Lancet* **361**, 955–64.

19. De Marzo A.M., Putzi M.J., Nelson W.G., New concepts in the Pathology of Prostatic Epithelial Carcinogenesis. *Urology* **57**,103–14.

20. Sun J., Turner A., Xu J., Grönberg H. & Isaacs W. 2007. Genetic Variability in Inflammation Pathways and Prostate Cancer Risk. *Urologic Oncology: Seminars and Original Investigations* **25**, 250–9.

21. Hsing A.W., Sakoda L.C., Chen J., Chokkalingam A.P., Sesterhenn I, Gao Y.T., Xu J. & Zheng S.L. 2007. MSR-1 Variants and the Risks of Prostate Cancer and Benign Prostatic Hyperplasia: A Population Based Study in China. *Carcinogenesis* **12**, 2530–2536.

22. Waugh D.J.J. & Wilson C. 2008. The Interleukin-8 Pathway in Cancer. *Clinical Cancer Research* **14**, 6735–41.

23. Isaacs W. & Kainu T. 2001. Oncogenes and Tumor Suppressor Genes in Prostate Cancer. *Epidemiologic Reviews* 23, 36–41.

24. Sharma A., Yeow W., Ertel A., Coleman I., Clegg N., Thangavel C., Morrissey C., Zhang X., Comstock C.E., Witkiewicz A.K. & Gomella L. 2010. The Retinoblastoma Tumor Suppressor Controls Androgen Signaling and Human Prostate Cancer Progression. *The Journal of Clinical Investigation* **120**, 4478-92

25. Phin S., Moore M.W. & Cotter P.D. 2013. Genomic Rearrangements of PTEN in Prostate Cancer. *Frontiers in Oncology* 3: 240.

26. Croce C.M. 2008. Oncogenes and Cancer. *The New England Journal of Medicine* **358**, 502-11.

27. Dasgupta S., Srinidhi S. & Vishwanatha J. 2012. Oncogenic Activation in Prostate Cancer Progression and Metastasis Molecular Insights and Future Challenges. *Journal of Carcinogenesis* 11: 4.

28. Catz S.D.& Johnson J.L. 2003. BCL-2 in Prostate Cancer. *Apoptosis* **8**, 29–37.

29. Tan E., Li J., Xu HE., Melcher K. & Yong E. 2014. Androgen receptor: Structure, Role in Prostate Cancer and Drug Discovery. *Nature Publishing Group* **36**, 3–23.

30. Chang C. & Heinlein C.A. 2016. Androgen Receptor in Prostate Cancer. *Endocrine Reviews*

25, 276–308.

31. Grossmann M.E., Huang H. & Tindall D.J. 2001. Androgen Receptor Signaling in Androgen-Refractory Prostate Cancer. *Journal of the National Institute of Cancer* **22**,1687–97.

32. Turcu A., Smith J.M., Auchus R. & Rainey W.E. 2014. Adrenal Androgens and Androgen Precursors—Definition, Synthesis, Regulation and Physiologic Actions. *Comprehensive Physiology* (2014).

33. Lavin N. 2002. *Manual of Endocrinology and Metabolism*. Philadelphia. Lippincott Willaiams & Wilkins.

34. Velonas V.M., Woo H.H., Dos Remedios C.G. & Assinder S.J. 2013. Current Status of Biomarkers for Prostate Cancer. *International Journal of Molecular Sciences* **14**, 11034–60.

35. Nargund V., Raghavan D. & Sandler H. 2015. *Urological Oncolog*y. London. Springer.

36. Lajiness M.(ed.) & Quallich S.(ed.) 2016. *The Nurse Practitioner in Urology*. London. Springer.

37. Humphrey P.A. 2004.Gleason grading and Prognostic Factors in Carcinoma of the Prostate. *Modern Pathology* **17**,292–306.

38. Denis L., (ed.) Griffiths K., (ed.) Kaisary A.(ed.) & Murphy G.(ed) 1999. *Textbook of Prostate Cancer: Pathology, Diagnosis and Treatment*. London. Martin Dunitz

39. Scher H.I., Scardino P.T. & Zelefsky M.J. 2015. Chapter 68: Cancer of the Prostate. *DeVita, Hellman, and Rosenberg's Cancer: Principles and Practice of Oncology*. Philadelphia. Lippincott Willaiams & Wilkins.

40. Kadmon D. 2002. Radiation Therapy After Radical Prostatectomy: Strike Early, Strike Hard! The Case for Adjuvant Radiation Therapy. *Reviews in Urology* **4**, 87–9

41. Jonler M., Riehmann M. & Brinkmann R.1994. Benign prostatic hyperplasia. *Endocrinology Metabolism* **23**(4), 795–807.

42. Thiruchelvam N. Surgical therapy for benign prostatic hypertrophy/bladder outflow obstruction. *Indian journal of urology,* (2):202.

43. McNicholas T. & Swallow D. 2011. Benign Prostatic Hyperplasia. *Surgery* **29**, 282–286

44. Ziada A., Rosenblum M. & Crawford E.D. 1999. Benign Prostatic Hyperplasia: An Overview. *Urology*. **53**, 1–6.

45. Bravi F., Bosetti C., Dal Maso L., Talamini R., Montella M., Negri E., Ramazzotti V., Franceschi S. & La Vecchia C. 2006 Food Groups and Risk of Benign Prostatic Hyperplasia. *Urology* **67**, 73–9.

46. Ilic D. & Misso M. 2012. Lycopene for the Prevention and Treatment of Benign Prostatic

Hyperplasia and Prostate Cancer: A Systematic Review. *Maturitas* **72**, 269–76.

47.   Salam M.T., Ursin G., Skinner E.C., Dessissa T. & Reichardt J.K. 2005. Associations Between Polymorphisms in the Steroid 5α-Reductase Type II (SRD5A2) Gene and Benign Prostatic Hyperplasia and Prostate Cancer. *Urologic Oncology* **23**, 246–53.

48.   Briganti A., Capitanio U., Suardi N., Gallina A., Salonia A., Bianchi M, et al. 2009. Benign Prostatic Hyperplasia and Its Aetiologies. *European Urology Supplements* **8**, 865–71.

49.   Lee K.L. & Peehl D.M. 2004. Molecular and Cellular Pathogenesis of Benign Prostatic Hyperplasia. *The Journal of Urology* **172**, 1784–91.

50.   Shabbir M. & Mumtaz F. Benign Prostatic Hyperplasia. *Journal- Royal Society For The Promotion Of Health* **124**, 222–227.

51.   Crawford E.D., Ventii K. & Shore N.D. 2014. New Biomarkers in Prostate Cancer. *Oncology Cancers* **28**, 135–41`

52.   Dias S. 2012. Benign Prostatic Hyperplasia : Clinical Manifestations and Evaluation. *Techniques In Vascular and Interventional Radiology* **15**, 265–269

53.   Thorpe A. & Neal D. 2003. Benign Prostatic Hyperplasia. *Lancet (London, England)* **361**,1359–67.

54.   Parnham A. & Haq A. 2013. Benign Prostatic Hyperplasia. *Journal of Clinical Urology* **24**, 24–31.

55.   Mayeux R. 2004. Biomarkers : Potential Uses and Limitations.Th*e Journal of the American Society for Experimental NeuroTherapeutics* **1**,182–8.

56.   Wagner D.P. Verma M. & Srivastava S. 2004.Challenges for Biomarkers in Cancer Detection. *Annals of the New York Academy of Sciences* **1022**, 9–16.

57.   Bhatt A.N, Mathur R., Farooque A., Verma A. & Dwarakanath B.S. 2010. Cancer biomarkers - Current Perspectives. *Indian Journal Of Medical Research* **132**, 129–49.

58.   Henry N.L. & Hayes D.F. 2012. Cancer biomarkers. *Molecular Oncology* **6**, 140–6.

59.   Mishra A. & Verma M. 2010. Cancer Biomarkers: Are We Ready for the Prime Time? *Cancers* **2**, 190–208.

60.   Menni C., Fauman E., Erte I., Perry J.R., Kastenmüller G., Shin S.Y., Petersen A.K., Hyde C., Psatha M., Ward K.J. & Yuan W. 2013. Biomarkers for type 2 diabetes and impaired fasting glucose using a non-targeted metabolomics approach. *Diabetes*, p.DB_130570.

61.   Srinivas P.R., Kramer B.S. & Srivastava S. 2001. Trends In Biomarker Research For Cancer Detection.*The Lancet. Oncology* **2**, 698–704.

62.   Flepisi B.T., Bouic P., Sissolak G. & Rosenkranz B. 2014. Biomarkers of HIV-associated

Cancer. *Biomarkers In Cancer* **6**, 11–20.

63.   Kulasingam V. 2008. *Identification and Validation of Candidate Breast Cancer Biomarkers: A Mass Spectrometric Approach*. University of Toronto

64.   Chatterjee S.K. & Zetter B.R. 2005. Cancer biomarkers:Knowing the Present and Predicting the Future. *Future Oncology* **1**,37–50.

65.   Gold P. & Freedman S.O. 1965. Demonstration of Tumor-Specific Antigens in Human Colonic Carcinomata By Immunological Tolerance and Absorption Techniques. *The Journal Of Experimental Medicin*e **121**, 439–62.

66.   Stephan C., Jung K., Lein M. & Diamandis E.P. 2007. PSA and Other Tissue Kallikreins for Prostate Cancer Detection. *European Journal of Cancer* **43**, 1918–26.

67.   Cook E.D. & Nelson A.C. 2011. Prostate Cancer Screening. *Current Oncology Reports* **13**, 57–62.

68.   Catalona W., Smith D., Dodds K., Coplen D., Yuan J. & Petros J. 1991. Measurement of Prostate Specific Antigen In Serum As A Screening Test For Prostate Cancer. T*he New England Journal Of Medicine* **324**, 1156–61.

69.   Kulasingam V. & Diamandis E.P. 2008. Strategies for Discovering Novel Cancer Biomarkers Through Utilization of Emerging Technologies. *Nature Clinical Practice Onclology* **5**, 588–99.

70.   Manne U., Srivastava R-G. & Srivastava S. Keynote review: Recent Advances in Biomarkers for Cancer Diagnosis and Treatment. *Drug Discovery Today* **10**, 965–76.

71.   Haj-ahmad T.A. 2013. *Prostate Cancer Diagnostics and The Search for Novel Biomarkers*. Brock University.

72.   Maruvada P. & Srivastava S. 2004. Biomarkers for cancer diagnosis: implications for nutritional research. *The Journal of Nutrition* **134**(6), 1640S-5S.

73.   Bleavins M., (ed.,) Carini C., (ed.,) Jurima-Romet M., (ed.,) Rahbari R. (ed.) 2011. *Biomarkers in Drug Development: A Handbook of Practice, Application, and Strategy*. New Jersey. John Wiley and Sons.

74.   Heckman-Stoddard B.M. 2012. Oncology Biomarkers: Discovery, Validation, and Clinical Use. *Seminars in Oncology Nursing* **28**, 93–8.

75.   Martin S.K., Vaughan T.B., Atkinson T., Zhu H. & Kyprianou N. 2012. Emerging Biomarkers of Prostate Cancer (Review). *Oncology Reports* **28**, 409–17.

76.   Mirabelli P. & Incoronato M. 2013. Usefulness of Traditional Serum Biomarkers for Management of Breast Cancer Patients. *BioMed Research International* **2013**, 1–9

77.   Filella X. & Foj L. Emerging Biomarkers in the Detection and Prognosis of Prostate Cancer.

Clinical Chemistry and Laboratory Medicinem **53**, 963–73.

78.   Huang J.G., Campbell N. & Goldenberg L. 2014. PSA and Beyond: Biomarkers in Prostate
       Cancer. *British Columbia Medical Journal* **56**, 334–41.

79.   Obort A.S., Ajadi M.B. & Akinloye O. 2013. Prostate-Specific Antigen: Any Successor In Sight?
       Reviews In Urology **15**, 97–107.

80.   Sardana G., Dowell B. & Diamandis E.P. 2008. Emerging Biomarkers for the Diagnosis and
       Prognosis of Prostate Cancer. *Clinical Chemistry* **54**, 1951–60.

81.   Hong S.K. 2014. Kallikreins as Biomarkers for Prostate Cancer. *BioMed Research
       International* **2014**.

82.   Ramsay A.J., Reid J.C., Adams M.N, Samaratunga H., Dong Y., Clements J.A. & Hooper J.D.
       Prostatic Trypsin-like Kallikrein-Related Peptidases (KLKs) and other Prostate-Expressed
       Tryptic Proteinases as Regulators of Signalling via Proteinase-Activated Receptors (PARs).
       Biological Chemistry **389**, 653–68.

83.   Fan L., Wang H., Xia X., Rao Y., Ma X., Ma D., Wu P. & Chen G. Loss of E-cadherin Promotes
       Prostate Cancer Metastasis via Upregulation of Metastasis-Associated Gene 1 Expression.
       Oncology Letters **4**, 1225–33.

84.   Stoeber K., Swinn R., Prevost A.T., Clive-Lowe P., Halsall I., Dilworth S.M., Marr J., Turner
       W.H., Bullock N., Doble A. & Hales C.N. 2010. Diagnosis of Genito-Urinary Tract Cancer by
       Detection of Minichromosome Maintenance 5 Protein in Urine Sediments. *British Journal Of
       Cancer* **103**,701–707.

85.   Zhao Z., Ma W., Zeng G., Qi D., Ou L. & Liang Y. 2011. Serum Early Prostate Cancer Antigen
       (EPCA) Level and Its Association with Disease Progression in Prostate Cancer in a Chinese
       Population. *PLoS One* **6** (5).

86.   Azevedo A., Cunha V., Teixeira A.L. & Medeiros R. 2011. IL-6/IL-6R as a Potential Key
       Signaling Pathway in Prostate Cancer Development. *World Journal Of Clinical Oncology* **2**,
       384–96.

87.   Erb H.H., Langlechner R.V., Moser P.L., Handle F., Casneuf T., Verstraeten K., Schlick B.,
       Schäfer G., Hall B., Sasser K. & Culig Z. 2013. IL-6 Sensitizes Prostate Cancer to the
       Antiproliferative Effect of IFNα2 through IRF9. *Endocrine-Related Cancer* **20**, 677–89.

88.   Bussemakers M.J., Van Bokhoven A., Verhaegh G.W., Smit F.P., Karthaus H.F., Schalken
       J.A., Debruyne F.M., Ru N. & Isaacs W.B. 1999. DD3: A New Prostate-Specific Gene, Highly
       Overexpressed in Prostate Cancer. *Cancer Research* **59**, 5975–9.

89.   Fiorentino M., Capizzi E. & Loda M. 2010. Blood and tissue biomarkers in prostate cancer:
       state of the art. *Urologic Clinics of North America* **37**(1), 131-41.

90.   Dong Z., Saliganan A.D., Meng H., Nabha S.M., Sabbota A.L., Sheng S., Bonfil R.D. & Cher M.L. 2008. Prostate Cancer Cell-Derived Urokinase-Type Plasminogen Activator Contributes to Intraosseous Tumor Growth and Bone Turnover. *Neoplasia* **10**, 439–49.

91.   Skovgaard D., Persson M. & Kjaer A. 2016. PET Imaging of Urokinase-Type Plasminogen Activator Receptor (uPAR) in Prostate Cancer: Current Status and Future Perspectives. *Clinical And Translational Imaging* **4**, 457-65.

92.   Sallam R.M. 2015. Proteomics in Cancer Biomarkers Discovery: Challenges and Applications. *Disease Markers* **2015**, 1–12.

93.   Ghosh D. & Poisson L.M. "Omics" Data and Levels of Evidence for Biomarker Discovery. *Genomics* **93**, 13–6.

94.   Bernard P.S.& Wittwer C.T. 2002. Real-Time PCR Technology for Cancer Diagnostics. *Clinical Chemistry* **48**, 1178–85.

95.   Mardis E.R. & Wilson R.K. Cancer Genome Sequencing: A Review. *Human Molecular Genetics* **18**,163–8.

96.   Kussmann M. & Affolter M. 2006. OMICS-Driven Biomarker Discovery in Nutrition and Health. *Journal of Biotechnology* **124**, 758–87.

97.   Bumgarner R. 2013. DNA Microarrays: Types, Applications and their Future.*Current Protocols In Molecular Biology* **6137**,1–17.

98.   Lee J-Y., Kim H.S., Suh D.H., Kim M-K., Chung H.H., Song Y-S. 2013. Ovarian Cancer Biomarker Discovery Based on Genomic Approaches. *Journal Of Cancer Prevention* **18**, 298–312.

99.   McDermott U., Downing J.R. & Stratton, M.R. 2011. Genomics and the Continuum of Cancer Care. *New England Journal of Medicine* **36,** 340-350.

100.   Quail M., Smith M.E., Coupland P., Otto T.D., Harris S.R., Connor T.R., Bertoni A., Swerdlow H.P. & Gu Y. 2012. A Tale of Three Next Generation Sequencing Platforms: Comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq Sequencers. *BMC Genomics* **13**, 341.

101.   Liu L., Li Y., Li S., Hu N., He Y., Pong R., Lin D., Lu L. & Law M. 2012. Comparison of Next-Generation Sequencing Systems. *Journal Of Biomedicine & Biotechnology* **2012,** 2012.

102.   Zhang A., Sun H. & Wang X. 2012. Serum Metabolomics as a Novel Diagnostic Approach for Disease: A Systematic Review. *Analytical and Bioanalytical Chemistry* **404**, 1239–45.

103.   Peng B., Li H. & Peng X. Functional Metabolomics: From Biomarker Discovery to Metabolome Reprogramming. *Protein & Cell* **6**, 628–37.

104.   Armitage E.G. & Barbas C. 2014. Metabolomics in Cancer Biomarker Discovery: Current Trends and Future Perspectives. *Journal Of Pharmaceutical And Biomedical Analysis* **87**, 1–

11.

105.   Allwood J.W., Ellis D.I. & Goodacre R. 2008. Metabolomic technologies and their application to the study of plants and plant-host interactions. *Physiologia Plantarum* **132**(2), 117–35.

106.   Soliman L.C., Hui Y., Hewavitharana A.K. & Chen D.D.Y. 2012. Monitoring potential prostate cancer biomarkers in urine by capillary electrophoresis-tandem mass spectrometry. *Journal of Chromatography* **1267,** 162–9.

107.   An H.J., Kronewitter S.R., de Leoz M.L. & Lebrilla C.B. Glycomics and Disease Markers. *Current Opinion In Chemical Biology* **13**, 601–607.

108.   Zaia J. 2008. Review Mass Spectrometry and the Emerging Field of Glycomics. *Chemistry And Biology* **15**, 881–92.

109.   Adamczyk B., Tharmalingam T. & Rudd P.M., 2012. Glycans as cancer biomarkers.

110.   Anderson N.L. & Anderson N.G. 1998. Proteome and proteomics: new technologies, new concepts, and new words. *Electrophoresis* **19**(11), 1853-61.

111.   Zhang Y., Fonslow B.R., Shan B., Baek M.C. & Yates J.R. 2013. Protein Analysis by Shotgun/Bottom-Up Proteomics. *Chemical Reviews* **113**, 2343–94.

112.   Yates J.R., Ruse C.I. & Nakorchevsky A. 2009. Proteomics by Mass Spectrometry: Approaches, Advances, and Applications. *Annual Review Of Biomedical Engineering* **11**, 49–79.

113.   Gstaiger M. & Aebersold R. Applying Mass Spectrometry-Based Proteomics to Genetics, Genomics and Network biology. *Nature Reviews Genetics* **10**, 617–27.

114.   Chandramouli K., Qian P-Y., 2009. Proteomics: Challenges, Techniques and Possibilities to Overcome Biological Sample Complexity. *Human Genomics And Proteomics* **1**(1).

115.   Sallam R.M. Proteomics in Cancer Biomarkers Discovery: Challenges and Applications. Disease Markers **2015**, 1–12.

116.   Van Oudenhove L. & Devreese B. 2013. A Review on Recent Developments in Mass Spectrometry Instrumentation and Quantitative Tools Advancing Bacterial Proteomics. *Applied Microbiology and Biotechnology* **97**, 4749–62.

117.   Kisluk J., Ciborowski M., Niemira M., Kretowski A. & Niklinski J. 2014. Proteomics Biomarkers for Non-Small Cell Lung Cancer. *Journal Of Pharmaceutical And Biomedical Analysis* **101**, 40–9.

118.   Ong S.E. & Mann M. 2005. Mass Spectrometry-Based Proteomics Turns Quantitative. *Nature Chemical Biology* **1**, 252–62.

119.   Tainsky M.A. 2009. Genomic and Proteomic Biomarkers for Cancer: A Multitude of

Opportunities. *Biochimica et Biophysica Acta (BBA)-Reviews On Cancer* **1796**, 176–93.

120. Ullah M.F. & Aatif M. 2009. The Footprints of Cancer Development: Cancer Biomarkers. *Cancer Treatment Reviews* **35**, 193–200.

121. Han X., Aslanian A. & Yates J.R. 2008. Mass Spectrometry for Proteomics. *Current Opinion In Chemical Biology* **12**, 483–90.

122. Paul D., Kumar A., Gajbhiye A., Santra M.K. & Srikanth R. 2013. Mass Spectrometry-Based Proteomics in Molecular Diagnostics: Discovery of Cancer Biomarkers using Tissue Culture. *BioMed Research International* **2013**, 1–16.

123. Aebersold R. & Mann M. 2003. Mass Spectrometry-Based Proteomics. *Nature* **422**, 198–207.

124. Glish G.L. & Burinsky D.J. 2008. Hybrid Mass Spectrometers for Tandem Mass Spectrometry. *Journal Of The American Society For Mass Spectrometry* **19**, 161–72.

125. Michalski A., Damoc E., Hauschild J.P., Lange O., Wieghaus A., Makarov A., et al. 2011. Mass Spectrometry-Based Proteomics Using Q Exactive, A High-Performance Benchtop Quadrupole Orbitrap Mass Spectrometer. *Molecular & Cellular Proteomics* **10**, M111-011015.

126. Senko M.W., Remes P.M., Canterbury J.D., Mathur R., Song Q., Eliuk M., Mullen C., Earley L., Hardman M., Blethrow J.D. & Bui H. 2013. Novel Parallelized Quadrupole/Linear Ion Trap/Orbitrap Tribrid Mass Spectrometer Improving Proteome Coverage and Peptide Identification Rates. *Analytical Chemistry* **85**(24), 11710-4.

127. Greaves J. & Roboz J. Mass Spectrometry for the Novice. Third Edit. New York: CRC Press; 2013. 309 p.

128. Kinter M. & Sherman N.E. 2005. P*rotein Sequencing and Identification Using Tandem Mass Spectrometry*. New York. John Wiley & Sons.

129. Siuzdak G. 1996. *Mass Spectrometry for Biotechnology*. San Diego. Academic Press.

130. Cseke J.L., Kirakosyan A., Kaufman B.P. & Westfall V.M. 2011. *Handbook of Molecular and Cellular Methods in Biology and Medicine*. Boca Raton. CRC Press

131. Toldra F. & Nollet L. editors., 2013. *Proteomics in Foods: Principles and Applications*. New York. Springer Science & Business Media.

132. March E.R. & Todd J.F. editors., 2010. *Practical Aspects of Trapped Ion Mass Spectrometry, Volume IV: Theory and Instrumentation*. Florida. CRC Press.

133. Whitelegge J. editor., 2008. *Protein Mass Spectrometry*. Oxford. Elsevier

134. Zhu M. & Lee S.M. editors., 2011. *Mass Spectrometry in Drug Metabolism and Disposition: Basic Principles and Applications*. New York. Wiley Interscience.

135. Scigelova M. & Makarov A. 2006. Orbitrap Mass Analyzer – Overview and Applications in

Proteomics. *Proteomics* **6**, 16–21.

136.   Kang J.S. 2012. Principles and Applications of LC-MS / MS for the Quantitative Bioanalysis of Analytes in Various Biological Samples. *Tandem Mass Spectrometry - Applications and Principles* **2012**, 441–492.

137.   Watson T.J. & Sparkman O.D. 2007. *Introduction to Mass Spectrometry: Instrumentation, Applications, and Strategies for Data Interpretation.* West Sussex. John Wiley and Sons.

138.   Dass C. 2007. *Fundamentals of Contemporary Mass Spectrometry*. New Jersey. John Wiley and Sons.

139.   Walther T.C. & Mann M. 2010. Mass Spectrometry-Based Proteomics in Cell Biology. *The Journal Of Cell Biology* **190**, 91–500.

140.   Kelleher N.L. 2004. Top-Down Proteomics. *Analytical Chemistry* **76**, 197A – 203A.

141.   Gregorich Z. & Ge Y. 2014. Top-Down Proteomics in Health and Disease: Challenges and Opportunities. *Proteomics* **14**, 1195–1210.

142.   Rohrbough J.G., Breci L., Merchant N., Miller S. & Haynes P.A. 2006. Verification of single-peptide protein identifications by the application of complementary database search algorithms. Journal of Biomolecular Technology **17**(5), 327–32.

143.   Wasinger V.C., Zeng M. & Yau Y. 2013. Current Status and Advances in Quantitative Proteomic Mass Spectrometry. *International Journal Of Proteomics* **2013**.

144.   Li Z., Adams R.M., Chourey K., Hurst G.B., Hettich R.L. & Pan C. 2012. Systematic Comparison of Label-Free, Metabolic Labeling, and Isobaric Chemical Labeling for Quantitative Proteomics on LTQ Orbitrap Velos. *Journal Of Proteome Research* **11**, 1582–90.

145.   Chahrour O., Cobice D. & Malone J. 2015. Stable Isotope Labelling Methods in Mass Spectrometry-Based Quantitative Proteomics. *Journal Of Pharmaceutical And Biomedical Analysis Elsevie* **113**, 2–20.

146.   Stastna M. & Van Eyk J.E. 2012. Secreted Proteins as a Fundamental Source for Biomarker Discovery. *Proteomics* **12**, 722–35.

147.   Karagiannis G.S., Pavlou M.P. & Diamandis E.P. 2010. Cancer Secretomics Reveal Pathophysiological Pathways in Cancer Molecular Oncology. *Molecular Oncology* **4**, 496–510.

148.   Pavlou M.P. & Diamandis E.P. 2010. The cancer Cell Secretome: A good Source for Discovering Biomarkers? *Journal Of Proteomics* **73**, 1896–906.

149.   Xue H., Lu B. & Lai M. 2008. The Cancer Secretome: A Reservoir of Biomarkers. *Journal Of Translational Medicine* **6** (1).

150.   Schaaij-Visser T.B., De Wit M., Lam S.W. & Jiménez C.R. 2013. The Cancer Secretome,

Current Status and Opportunities in the Lung, Breast and Colorectal Cancer Context. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics* **1834**, 2242–58.

151.   Wingfield P.T. Protein Precipitation Using Ammonium Sulfate. 2016;1–10.

152.   De Maio A. 2011. Extracellular heat shock proteins, cellular export vesicles, and the Stress Observation System: a form of communication during injury, infection, and cell damage. *Cell Stress and Chaperones* **16**(3), 235-49.

153.   Shi W., Xu G., Wang C., Sperber S.M., Chen Y., Zhou Q., Deng Y & Zhao H. Heat shock 70-kDa protein 5 (Hspa5) is essential for pronephros formation by mediating retinoic acid signalling. *Journal of Biological Chemistry* **290**(1), 577-89.

154.   Chen L., Feng L., Wang X., Du J., Chen Y., Yang W., Zhou C., Cheng L., Shen Y., Fang S & Li J. 2015. Mesencephalic astrocyte-derived neurotrophic factor is involved in inflammation by negatively regulating the NF-κB pathway. *Scientific reports* **5**.

155.   Gandaglia G., Zaffuto E., Fossati N., Cucchiara V., Mirone V., Montorsi F. & Briganti A. 2016. The role of prostatic inflammation in the development and progression of benign and malignant diseases. *Current Opinion in Urology* (2016).

156.   Elkahwaji J.E. 2012. The role of inflammatory mediators in the development of prostatic hyperplasia and prostate cancer. *Res Rep Urol* **5**(5), 1-10.

157.   Sfanos K.S. & De Marzo A.M. 2012. Prostate cancer and inflammation: the evidence. *Histopathology* **60**(1), 199-215.

158.   De Nunzio C., Kramer G., Marberger M., Montironi R,. Nelson W. Schröder F., Sciarra A. & Tubaro A. 2011. The controversial relationship between benign prostatic hyperplasia and prostate cancer: the role of inflammation. *European Urology* **60**(1), 106-17.

159.   Bostanci Y., Kazzazi A., Momtahen S., Laze J. & Djavan B. 2013. Correlation between benign prostatic hyperplasia and inflammation. Current Opinon in Urology **23**, 5–10.

160.   Charitou T., Bryan K. & Lynn D.J. 2016. Using biological networks to integrate, visualize and analyze genomics data. *Genetics Selection Evolution* **48**(1), 27.

161.   Shah S.H. & Newgard C.B. 2015. Integrated Metabolomics and Genomics*. Circulation: Cardiovascular Genetics* **8**(2), 410-9.

162.   Wanichthanarak K., Fahrmann J.F. & Grapov D. 2015. Genomic, proteomic, and metabolomic data integration strategies. *Biomarker Insights* **10**(Suppl 4):1.

163.   Powner M.W., Gerland B. & Sutherland J.D. 2009. Synthesis of activated pyrimidine ribonucleotides in prebiotically plausible conditions. *Nature* **459**, 239-42.

164.   Denis L. editor., 1991. *The Medical Management of Prostate Cancer*. New York. Springer

Appendix

Table A: Molecules mapped to the EIF2 signalling pathway from the experimental datasets and their respective fold change, p-value. The green arrow represents up regulation while the red arrow indicates down regulation. A1- BPH-1, A2- LNCaP and A3- PC3.

Symbol	Identifier			Expression Value						Location	Type(s)
	GenPept/UniProt	GenPept/UniProt	GenPept/UniProt	A1		A2		A3#			
				Expr Log Ratio	Expr p-value	Expr Log Ratio	Expr p-value	Expr Log Ratio	Expr p-value		
RPS6	RS6_HUMAN	RS6_HUMAN	RS6_HUMAN	↑∞	3.60E-01	↑4.700	1.20E-02	↓-4.644	4.10E-03	Cytoplasm	other
RPS3A	RS3A_HUMAN	RS3A_HUMAN	RS3A_HUMAN	↓-1.322	4.90E-02	↑1.807	7.90E-03	↓-2.322	2.10E-04	Nucleus	other
RPS3*	E9PL09_HUMAN	E9PQX2_HUMAN	RS3_HUMAN*	↓-1.000	1.30E-01	↑1.585	3.80E-01	↓-2.322	1.00E-04	Cytoplasm	enzyme
RPS18	RS18_HUMAN	RS18_HUMAN	RS18_HUMAN	↓-∞	4.70E-04	↑∞	6.30E-02	↓-∞	2.30E-03	Cytoplasm	other
RPL8	RL8_HUMAN	RL8_HUMAN	RL8_HUMAN	↓-3.322	1.00E-04	↑3.322	5.00E-03	↓-3.322	2.80E-03	Cytoplasm	other
RPL5	RL5_HUMAN	RL5_HUMAN	RL5_HUMAN	↓-1.000	1.30E-01	↑0.138	8.50E-01	↓-2.322	1.00E-04	Cytoplasm	other
RPL27	K7ELC7_HUMAN	RL27_HUMAN	RL27_HUMAN	↓-∞	3.00E-01	↑∞	1.50E-01	↓-∞	3.80E-04	Cytoplasm	other
RPS8	Q5JR95_HUMAN	Q5JR95_HUMAN	Q5JR95_HUMAN	↓-3.322	3.20E-04	↑4.248	1.50E-01	↓-4.322	1.00E-04	Cytoplasm	other
PPP1CB*	PP1B_HUMAN*	PP1B_HUMAN	PP1B_HUMAN	↑2.585	2.10E-04	↑0.138	4.30E-01	↑1.632	2.60E-03	Cytoplasm	phosphatase
GRB2*	P62993	P62993-2*	P62993-2*	↑1.070	3.70E-03	↓-∞	1.20E-04	↑∞	1.20E-04	Cytoplasm	kinase
RPS20	P60866-2	P60866-2	P60866-2	↑0.485	2.10E-02	↑0.263	1.70E-01	↑0.678	4.40E-03	Cytoplasm	other
EIF2AK2	P19525-2	P19525-2	P19525-2	↑0.585	2.30E-01	↓-2.322	2.60E-03	↑∞	1.00E-04	Cytoplasm	kinase
RPL17	P18621-3	P18621-3	P18621-3	↓-1.322	4.90E-03	↑2.036	3.60E-04	↓-1.737	1.00E-04	Cytoplasm	other
EIF1AX	P47813	P47813	P47813	↑0.263	1.20E-01	0.000	6.90E-01	↑1.585	1.00E-04	Cytoplasm	translation
EIF2S3	--	P41091	P41091			↑∞	6.70E-02	↓-∞	1.50E-03	Cytoplasm	translation
RPL12*	P30050-2*	P30050-2*	P30050*	↑∞	8.00E-02	↓-∞	6.30E-02	↑0.766	2.10E-04	Nucleus	other
RPL13*	P26373-2	P26373*	P26373	↓-3.322	4.10E-02	↓-∞	1.90E-03	↓-∞	1.00E-04	Nucleus	other
HSPA5	P11021	P11021	P11021	↓-1.737	1.00E-04	↑1.807	1.00E-04	↓-2.322	1.00E-04	Cytoplasm	enzyme
EIF1AY	O14602	O14602	O14602	↑0.138	2.60E-01	0.000	7.30E-01	↑1.433	1.00E-04	Other	translation
RPL23A*	H7BY10_HUMAN	H7BY10_HUMAN	K7EIV9_HUMAN	↓-0.152	6.10E-01	↑0.678	7.00E-03	↓-0.322	9.40E-03	Cytoplasm	other
RPL19	J3KTE4_HUMAN	J3KTE4_HUMAN	J3KTE4_HUMAN	↓-∞	5.30E-03	↑∞	5.60E-03	↓-∞	1.00E-04	Cytoplasm	other
IRS1	IRS1_HUMAN	IRS1_HUMAN	IRS1_HUMAN	↑3.053	1.90E-02	↓-∞	3.80E-03	↓-2.322	1.00E-04	Cytoplasm	enzyme
EIF5	IF5_HUMAN	IF5_HUMAN	IF5_HUMAN	0.000	8.00E-01	↑0.379	1.10E-01	↓-0.737	4.40E-03	Cytoplasm	translation
EIF2S2*	IF2B_HUMAN*	IF2B_HUMAN*	IF2B_HUMAN*	↑∞	1.00E-04	↓-∞	1.00E-04	↑∞	1.00E-04	Cytoplasm	translation
HNRNPA1*	P09651*	F8VZ49_HUMAN	F8W6I7*	↓-3.474	1.00E-04	↓-∞	3.30E-01	↑0.585	1.00E-04	Nucleus	enzyme
RPS10*	F6U211_HUMAN	F6U211_HUMAN	F6U211_HUMAN*	↓-0.152	9.10E-01	↑0.379	1.20E-01	↓-0.515	1.50E-03	Cytoplasm	other
RPL35	F2Z388_HUMAN	RL35_HUMAN	F2Z388_HUMAN	↓-3.322	2.10E-03	↑3.459	1.00E-04	↓-3.322	1.00E-04	Cytoplasm	other
PABPC1*	E7ERJ7*	E7ERJ7*	E7ERJ7	↓-0.515	2.20E-03	↑0.926	1.00E-04	↓-0.515	3.10E-04	Cytoplasm	translation
RPL24	C9JNW5_HUMAN	RL24_HUMAN	C9JNW5_HUMAN	↓-0.152	7.20E-01	↑1.000	2.70E-02	↓-1.000	1.00E-04	Cytoplasm	other
RPS7*	RS7_HUMAN	RS7_HUMAN	B5MCP9_HUMAN*	↓-1.322	8.50E-03	↑1.322	6.60E-02	↓-1.737	9.00E-03	Cytoplasm	other
EIF3J	B4DUI3	B4DUI3	B4DUI3	↓-∞	1.00E-04	↑∞	1.00E-04	↓-∞	1.00E-04	Cytoplasm	translation
ACTB	C9IZR7_HUMAN	ACTB_HUMAN	ACTB_HUMAN	↓-0.152	8.20E-01	↑0.485	8.90E-02	↓-1.322	1.00E-04	Cytoplasm	other
ACTA2	--	--	ACTA_HUMAN					↓-0.737	1.00E-04	Cytoplasm	other

Table B: Molecules mapped to the pyrimidine ribonucleotide de novo biosynthesis pathway from the experimental datasets and their respective fold change, p-value. The green arrow represents up regulation while the red arrow indicates down regulation. A1- BPH-1, A2- LNCaP and A3- PC3.

Symbol	Identifier			Expression Value						Location	Type(s)
	GenPept/UniProt/S	GenPept/UniProt/S	GenPept/UniProt/Sw	A1#		A2		A3			
				Expr Log Ratio	Expr p-value	Expr Log Ratio	Expr p-value	Expr Log Ratio	Expr p-value		
ANXA1	P04083	P04083	P04083	↓-1.322	1.00E-04	↓-∞	1.00E-04	0.000	8.50E-01	Plasma	enzyme
CMPK1*	Q5T0D2_HUMAN*	Q5T0D2_HUMAN*	Q5T0D2_HUMAN*	↓-2.322	1.00E-04	↑1.807	1.00E-04	↓-2.322	1.00E-04	Nucleus	kinase
HNRNPA1*	F8VZ49_HUMAN*	F8W6I7*	F8W6I7*	↓-3.474	1.00E-04	↓-∞	3.30E-01	↑0.585	1.00E-04	Nucleus	enzyme
NUDT5*	A6NFX8_HUMAN	A6NJU6_HUMAN*		↑∞	9.50E-03	↓-1.322	8.50E-02	↑∞	6.60E-02	Cytoplasm	phosphatase

Table C: Molecules mapped to telomere extension by telomerase from the experimental datasets and their respective fold change, p-value. The green arrow represents up regulation while the red arrow indicates down regulation. A1-BPH-1, A2-LNCaP and A3- PC3.

Symbol	Identifier		Expression Value						Location	Type(s)
	GenPept/UniProt/S	GenPept/UniPr	A1		A2		A3#			
			Expr Log Ratio	Expr p-value	Expr Log Ratio	Expr p-value	Expr Log Ratio	Expr p-value		
TERF2IP	Q9NYB0	Q9NYB0	↓-2.322	1.00E-04	↑1.632	1.00E-04	↓-1.322	9.40E-03	Nucleus	other
HNRNPA2B1	P22626	P22626	↑0.263	1.90E-02	↓-0.322	9.80E-03	↑0.766	1.00E-04	Nucleus	other
NBN	NBN_HUMAN	NBN_HUMAN	↑3.018	3.40E-02	↓-∞	8.80E-03	↑∞	8.70E-03	Nucleus	other
MRE11	F8W7U8_HUMAN	F8W7U8_HUMAN	0.000	8.70E-01	↓-1.000	3.70E-02	↑∞	1.00E-04	Nucleus	enzyme
HNRNPA1*	F8VZ49_HUMAN*	P09651*	↓-3.474	1.00E-04	↓-∞	3.30E-01	↑0.585	1.00E-04	Nucleus	enzyme

Table D: Molecules mapped to the antiproliferative role of TOB from the experimental datasets and their respective fold change, p-value. The green arrow represents up regulation while the red arrow indicates down regulation. A1#-BPH-1, A2-LNCaP and A3- PC3.

Symbol	Identifier			Expression Value						Location	Type(s)
	GenPept/UniP	GenPept/UniP	GenPept/UniP	A1#		A2		A3			
				Expr Log Ratio	Expr p-value	Expr Log Ratio	Expr p-value	Expr Log Ratio	Expr p-value		
CDKN1B*	E7ES52*	E7ES52*	E7ES52*	↓-1.737	1.10E-03	↓-1.322	3.40E-01	↑∞	6.60E-02	Nucleus	kinase
PABPC1*	E7ERJ7*	E7ERJ7*	E7ERJ7	↓-0.515	2.20E-03	↑0.926	1.00E-04	↓-0.515	3.10E-04	Cytoplasm	translation
PABPC4*	H0YEU6*	Q13310-2	H0Y6X6*	↓-1.737	2.30E-03	↑1.379	2.60E-01	↑∞	7.20E-02	Cytoplasm	translation
SKP1	P63208	P63208	P63208	↑0.766	3.50E-03	↓-2.322	1.00E-04	↑∞	1.00E-04	Nucleus	transcription

Table E: Molecules mapped to the gluconeogenesis pathway from the experimental datasets and their respective fold change, p-value. The green arrow represents up regulation while the red arrow indicates down regulation. A1- BPH-1, A2-LNCaP and A3- PC3.

Symbol	Identifier			Expression Value						Location	Type(s)
	GenPept/UniProt/S	GenPept/UniProt/S	GenPept/UniProt/	A1		A2		A3#			
				Expr Log Ratio	Expr p-value	Expr Log Ratio	Expr p-value	Expr Log Ratio	Expr p-value		
ALDOA	ALDOA_HUMAN	ALDOA_HUMAN	ALDOA_HUMAN	↓-1.000	1.00E-04	↑0.926	3.80E-03	↓-1.322	1.00E-04	Cytoplasm	enzyme
ALDOC	A8MVZ9_HUMAN	A8MVZ9_HUMAN	A8MVZ9	↓-∞	5.00E-02	↑∞	4.40E-02	↓-4.644	1.00E-04	Cytoplasm	enzyme
ENO1	P06733	P06733	P06733	↓-1.322	2.60E-04	↑1.000	4.70E-02	↓-1.322	1.00E-04	Cytoplasm	enzyme
GAPDH	P04406	P04406	P04406	↑0.138	6.40E-01	↑0.585	1.40E-01	↓-1.322	5.80E-04	Cytoplasm	enzyme
MDH1	P40925-3	P40925-3	P40925-3	↓-0.515	5.30E-01	↓-0.152	8.30E-01	↓-2.322	1.00E-04	Cytoplasm	enzyme
MDH2	P40926	P40926	P40926	↓-0.515	5.00E-04	↑1.202	1.00E-04	↓-1.000	1.00E-04	Cytoplasm	enzyme
PGAM1	P18669	P18669	P18669	↓-1.000	4.40E-02	↑1.585	1.90E-03	↓-1.737	3.50E-03	Cytoplasm	phosphatase
PGK1	PGK1_HUMAN	PGK1_HUMAN	PGK1_HUMAN	↓-0.737	5.50E-02	↑1.536	2.00E-03	↓-2.322	1.00E-04	Cytoplasm	kinase

Table F: Top three networks in the BPH-1 dataset generated by IPA. The network score is based on the hypergeometric distribution of network eligible molecules in the experimental dataset and is calculated with the right-tailed Fisher's Exact Test. The green arrows represents down regulation while the red arrows indicates up regulation.

Network	Molecules in Network	Score
A	AKT1, AR, BUB1, CALU*, CCND1, CDCA5*, CTNNB1, CTNNBIP1, DACH1, DAXX, ENO1, GAPDH, HNRNPA1*, HNRNPK*, HSPA8*, KLK3, KRT18*, MYC, NCOA1, NCOR1, NCOR2*, NKX3-1, PHB*, PIK3CB, PMEPA1, PRDX1, PROS1, RCHY1, RNF14, SMAD3, STK4, TACC2, TNK2, UXT	9
B	AGR2, AR, BAX, CCNA2, CCND1, CD44, CDC6, CDH1*, CDK2, CDK9, CDKN1B*, CDKN2A, CTNNB1, CXCL12, ETS2, FEN1, FOXA1, FOXO3, GSK3B, GUCY1A3, IL1RN, MCM6, MET, MIF, NEK2, NUMA1*, PA2G4, PDLIM7*, PRKCD, SNAI2, STAT3, TERT, TP53, TPD52*, TPX2*	6
C	AR, BAX, BCL2L1, CLU, HSPA5, HSPD1*, ID1, ITGA6, ITGB4, RELA, XRCC6	5

Table G: Extracellular and membrane proteins identified in the BPH-1 cell proteome, but not the other prostate cell lines.

Uniprot ID	Protein Name	Gene Name
Q7L7L0	Histone H2A type 3	HIST3H2A
Q16777	Histone H2A type 2-C	HIST2H2AC H2AFQ
Q9BTM1	Histone H2A.J	H2AFJ
Q8IUE6	Histone H2A type 2-B	HIST2H2AB
P16104	Histone H2AX	H2AFX H2AX
Q96QV6	Histone H2A type 1-A	HIST1H2AA H2AFR
P10412	Histone H1.4	HIST1H1E H1F4
Q01546	Keratin, type II cytoskeletal 2 oral	KRT76 KRT2B KRT2P
P37802	Transgelin-2	TAGLN2 KIAA0120 CDABP0035
P51858	Hepatoma-derived growth factor	HDGF HMG1L2
O95171	Sciellin	SCEL
P68036	Ubiquitin-conjugating enzyme E2 L3	UBE2L3 UBCE7 UBCH7
P48637	Glutathione synthetase	GSS
P53999	Activated RNA polymerase II transcriptional coactivator p15	SUB1 PC4 RPO2TC1
O75347	Tubulin-specific chaperone A	TBCA
P07108	Acyl-CoA-binding protein	DBI
P04179	Superoxide dismutase, mitochondrial	SOD2
P55145	Mesencephalic astrocyte-derived neurotrophic factor	MANF ARMET ARP
Q15843	NEDD8	NEDD8
Q9UBG3	Cornulin	CRNN C1orf10 DRC1 PDRC1 SEP53
P13646	Keratin, type I cytoskeletal 13	KRT13
P06702	Protein S100-A9	S100A9 CAGB CFAG MRP14
P50502	Hsc70-interacting protein	ST13 AAG2 FAM10A1 HIP SNC6
P63208	S-phase kinase-associated protein 1	SKP1 EMC19 OCP2 SKP1A TCEB1L
Q9H3K6	BolA-like protein 2	BOLA2 BOLA2A My016; BOLA2B

P61088	Ubiquitin-conjugating enzyme E2 N	UBE2N BLU
Q5JXB2	Putative ubiquitin-conjugating enzyme E2 N-like	UBE2NL
P31151	Protein S100-A7	S100A7 PSOR1 S100A7C
P19957	Elafin	PI3 WAP3 WFDC14
Q07065	Cytoskeleton-associated protein 4	CKAP4
P12277	Creatine kinase B-type	CKB CKBB
A6NHG4	D-dopachrome decarboxylase-like protein	DDTL
Q92817	Envoplakin	EVPL
H7C469	Uncharacterized protein	
Q04760	Lactoylglutathione lyase	GLO1
Q96FQ6	Protein S100-A16	S100A16 S100F AAG13
O75531	Barrier-to-autointegration factor	BANF1 BAF BCRG1
O14896	Interferon regulatory factor 6	IRF6
P08779	Keratin, type I cytoskeletal 16	KRT16 KRT16A
P01040	Cystatin-A	CSTA STF1 STFA
O15020	Spectrin β-chain, non-erythrocytic 2	SPTBN2 KIAA0302 SCA5
P29508	Serpin B3	SERPINB3 SCCA SCCA1
O95785	Protein Wiz	WIZ ZNF803
Q96NY8	Nectin-4	NECTIN4 LNIR PRR4 PVRL4
Q9HCY8	Protein S100-A14	S100A14 S100A15
O00115	Deoxyribonuclease-2-α	DNASE2 DNASE2A DNL2
O76070	Gamma-synuclein	SNCG BCSG1 PERSYN PRSN
P32926	Desmoglein-3	DSG3 CDHF6
Q9NZH8	Interleukin-36 gamma	IL36G IL1E IL1F9 IL1H1 IL1RP2 UNQ2456/PRO5737
P06744	Glucose-6-phosphate isomerase	GPI
P46940	Ras GTPase-activating-like protein IQGAP1	IQGAP1 KIAA0051
P47929	Galectin-7	LGALS7 PIG1; LGALS7B

Q9UMD9	Collagen α-1	COL17A1 BP180 BPAG2
Q9H0E2	Toll-interacting protein	TOLLIP
P12268	Inosine-5'-monophosphate dehydrogenase 2	IMPDH2 IMPD2
P46531	Neurogenic locus notch homolog protein 1	NOTCH1 TAN1
P62826	GTP-binding nuclear protein Ran	RAN ARA24 OK/SW-cl.81
P01772	Immunoglobulin heavy variable 3-33	IGHV3-33
P16152	Carbonyl reductase [NADPH] 1	CBR1 CBR CRN SDR21C1
P35232	Prohibitin	PHB
P62269	40S ribosomal protein S18	RPS18 D6S218E
P01767	Immunoglobulin heavy variable 3-53	IGHV3-53
Q86Y46	Keratin, type II cytoskeletal 73	KRT73 K6IRS3 KB36 KRT6IRS3
P78310	Coxsackievirus and adenovirus receptor	CXADR CAR
Q9Y265	RuvB-like 1	RUVBL1 INO80H NMP238 TIP49 TIP49A
Q96CF2	Charged multivesicular body protein 4c	CHMP4C SHAX3
P11279	Lysosome-associated membrane glycoprotein 1	LAMP1
P36952	Serpin B5	SERPINB5 PI5
Q9BW30	Tubulin polymerization-promoting protein family member 3	TPPP3 CGI-38
P50452	Serpin B8	SERPINB8 PI8
Q9H190	Syntenin-2	SDCBP2 SITAC18
P19012	Keratin, type I cytoskeletal 15	KRT15 KRTB
Q6ZSZ5	Rho guanine nucleotide exchange factor 18	ARHGEF18 KIAA0521
P22087	rRNA 2'-O-methyltransferase fibrillarin	FBL FIB1 FLRN
P05198	Eukaryotic translation initiation factor 2 subunit 1	EIF2S1 EIF2A
P53990	IST1 homolog	IST1 KIAA0174
Q5VT79	Annexin A8-like protein 1	ANXA8L1 ANXA8L2
P00918	Carbonic anhydrase 2	CA2
Q9NZT1	Calmodulin-like protein 5	CALML5 CLSP
P12004	Proliferating cell nuclear antigen	PCNA
O43175	D-3-phosphoglycerate dehydrogenase	PHGDH PGDH3

Q9UNN8	Endothelial protein C receptor	PROCR EPCR
P12830	Cadherin-1	CDH1 CDHE UVO
P39023	60S ribosomal protein L3	RPL3 OK/SW-cl.32
P63244	Receptor of activated protein C kinase 1	RACK1 GNB2L1 HLC7 PIG21
Q9UN37	Vacuolar protein sorting-associated protein 4A	VPS4A VPS4
P62280	40S ribosomal protein S11	RPS11
P20933	N(4)-(β-N-acetylglucosaminyl)-L-asparaginase	AGA
Q5JSG7	Proprotein convertase subtilisin/kexin type 5	PCSK5
P16333	Cytoplasmic protein NCK1	NCK1 NCK
Q8TD16	Protein bicaudal D homolog 2	BICD2 KIAA0699
Q9NP74	Palmdelphin	PALMD C1orf11 PALML
E7EPK1	Septin-7	SEPT7
O94906	Pre-mRNA-processing factor 6	PRPF6 C20orf14
Q9Y4H2	Insulin receptor substrate 2	IRS2
P48230	Transmembrane 4 L6 family member 4	TM4SF4 ILTMP
Q9NVP1	ATP-dependent RNA helicase DDX18	DDX18 cPERP-D
Q9C0B5	Palmitoyltransferase ZDHHC5	ZDHHC5 KIAA1748 ZNF375
Q8TAE6	Protein phosphatase 1 regulatory subunit 14C	PPP1R14C KEPI
E9PCW0	Prohibitin	PHB
Q53RT3	Retroviral-like aspartic protease 1	ASPRV1 SASP
Q9Y6M5	Zinc transporter 1	SLC30A1 ZNT1
P62760	Visinin-like protein 1	VSNL1 VISL1
P07949	Proto-oncogene tyrosine-protein kinase receptor Ret	RET CDHF12 CDHR16 PTC RET51
P01715	Immunoglobulin lambda variable 3-1	IGLV3-1
O43639	Cytoplasmic protein NCK2	NCK2 GRB4
Q9P2E7	Protocadherin-10	PCDH10 KIAA1400
Q6NZY7	Cdc42 effector protein 5	CDC42EP5 BORG3 CEP5
Q9NX40	OCIA domain-containing protein 1	OCIAD1 OCIA
P59551	Taste receptor type 2 member 60	TAS2R60