Molecular strategies to reduce unnecessary repeat prostate biopsies of men with elevated serum PSA

By Karl H. Pang

MRCS (The Royal College of Surgeons of England, UK)
MBChB (The University of Glasgow, UK)
MSc in Clinical Research (The University of Sheffield, UK)
Intercalated BSc in Immunobiology (King’s College London, UK)

A thesis submitted for the award of a Doctorate in Philosophy

University of Sheffield
The Department of Oncology and Metabolism and
Academic Urology Unit

January 2018
I hereby declare that no part of this thesis has previously been submitted for any degree or qualification at this, or any other University or institute of learning.
# Table of Contents

TABLE OF CONTENTS 3

ABSTRACT 9

FOREWORD 10

ACKNOWLEDGMENTS 11

LIST OF ABBREVIATIONS 12

GRANT FUNDING FOR SUPPORT OF THIS WORK 16

LIST OF ACHIEVEMENTS ARISING FROM THE WORK UNDERTAKEN IN THIS THESIS 17

Peer-reviewed publications 17

Abstract publications 17

Other publications 18

Oral presentations 18

Poster presentations 19

Prizes 19

LIST OF TABLES 20

LIST OF FIGURES 21

CHAPTER 1: INTRODUCTION 23

## 1.1 Cancer background

1.1.1 Pathogenesis of cancer 24

1.1.2 Hallmarks of cancer 25

1.1.2.1 Sustaining proliferative signalling 27

1.1.2.2 Evading growth suppressors 27

1.1.2.3 Activating invasion and metastasis 27

1.1.2.4 Enabling replicative immortality 28

1.1.2.5 Inducing angiogenesis 28

1.1.2.6 Resisting cell death 28

1.1.3 Emerging hallmarks and characteristics 29

1.1.3.1 Reprogramming energy metabolism 29

1.1.3.2 Evading immune destruction 29

1.1.4 Sustaining core and emerging hallmarks 29
1.5.2.4 Somatic Copy Number Aberrations and gene expression 71
1.5.2.5 Cytokine signalling

1.6 Epigenetics of prostate cancer 72
1.6.1 DNA methylation 72
1.6.2 Histone and chromatin modifications 73
1.6.3 Non-coding and MicroRNAs 74
1.6.4 MicroRNA and prostate cancer 76
   1.6.4.1 MicroRNA and cellular pathways 78
   1.6.4.2 MicroRNA and androgen-signalling 78
   1.6.4.3 MicroRNA as biomarkers 79
   1.6.4.4 RNA methylation 81

1.7 Aims 84
1.8 Significance 86

CHAPTER 2: MATERIALS AND METHODS 87

2.1 ProtecT and ProMPT studies 88
   2.1.1 Patient recruitment, consent and ethical approval 88

2.2. General laboratory equipment and reagents 91
   2.2.1 Laboratory equipment 91
   2.2.2 Plastic and disposable equipment 92
   2.2.3 General laboratory chemicals and reagents 92

2.3 General materials and methods 93
   2.3.1 Cell Lines and cell cultures 93
   2.3.2 Urinary sample preparation 94
   2.3.3 Prostate tissue collection
      2.3.3.1 Prostate tissue fixation and embedding 94
      2.3.3.2 FFPE sectioning 95

2.4 Specific materials and methods 95
   2.4.1 RNA Extraction 95
      2.4.1.1 Cell lines and urinary pellets 96
      2.4.1.2 Prostate biopsy FFPE 99
   2.4.2 Reverse transcription 100
   2.4.3 Real-time PCR 102

2.5 Immunoprecipitation and sequencing of m6A 103
   2.5.1 Immunoprecipitation of m6A
      2.5.1.1 mRNA purification and RNA fragmentation 104
      2.5.1.2 Immunoprecipitation 105
      2.5.1.3 Elution 106
      2.5.1.4 RNA purification 106
      2.5.1.5 Immunoprecipitation quality control 107
   2.5.2 Library preparation and sequencing 107

2.6 Statistical Analysis 110

CHAPTER 3: A RETROSPECTIVE ANALYSIS OF REPEAT PROSTATE BIOPSY OUTCOMES 111
CHAPTER 4: THE IDENTIFICATION AND ROLE OF PCA3-SHRNA2 IN PROSTATE CANCER

3.1 Background

3.2 Methods

3.3 Results

3.4 Discussion

4.1 Background

4.2 Methods

4.2.1 Design and patient population

4.2.2 Data collection and analysis

4.2.3 Statistical analysis

4.3 Results

4.3.1 Patient population within the Sheffield cohort

4.3.2 Outcomes of repeat prostate biopsy

4.3.3 Prostate cancer identified by prostate biopsy

4.3.4 The risks of being diagnosed with cancer on repeat biopsy

4.3.5 Comparison between men with and without cancer

4.3.6 Uni- and Multivariable analysis of predictors

4.3.7 Review of the international data

4.4.1 Diagnostic rate of prostate cancer on repeat biopsies

4.4.2 Grades of prostate cancer

4.4.3 Predictors of prostate cancer

4.4.4 Limitations

4.4.5 Generalisability

4.5 Conclusions
3.5 Conclusions

CHAPTER 5: PCA3-SHRNA2 EXPRESSION AND EVENTUAL DIAGNOSIS OF PROSTATE CANCER

5.1 Background

5.2 Methods
5.2.1 Expression of PCA3-shRNA2 in FFPE samples
5.2.2 RNA extraction from prostate biopsies (FFPE)

5.3 Results
5.3.1 Patients and FFPE samples
5.3.2 Expression of PCA3-shRNA2 in prostate biopsies (FFPE)
5.3.3 PCA3-shRNA2 expression and eventual diagnosis
5.3.4 The ability of PCA3-shRNA2 to identify disease (PBx FFPE)

5.4 Discussion

5.5 Conclusion

CHAPTER 6: N6-ADENOSINE METHYLATION AND CANCER

6.1 Background

6.2 Methods
6.2.1 In-silico analysis of m6A
6.2.1.1 Selection and annotation of RNA transcriptomic datasets
6.2.1.2 RNA Selection
6.2.1.3 Statistical analysis
6.2.2 Immunoprecipitation and sequencing of m6A

6.3 Results
6.3.1 Microarray datasets and sample population
6.3.2 The proportion of N6-adenosine methylated RNAs
6.3.3 Fold changes of m6A susceptible RNAs
6.3.4 Fold changes of the most differentially expressed RNAs
6.3.5 Functional annotation of m6A susceptible RNAs
6.3.5.1 Individual cancers
6.3.5.2 Between cancers
6.3.6 Immunoprecipitation of m6A
6.3.6.1 Validation of post-fragmentation RNA size
6.3.6.2 Immunoprecipitation quality control
6.3.7 RNA-sequencing of m6A-IP libraries
6.3.7.1 MeRIP-seq read quality control
6.3.7.2 MeRIP-seq read mapping
6.3.7.3 Identification of m6A sites
6.3.8 Expression of m6A transcripts in prostate cancer cell lines

6.4 Discussion
6.4.1 In-silico analysis of N6-methyladenosine
6.4.2 N6-methyladenosine profiling in prostate cancer cell lines
6.4.3 N6-methyladenosine machinery

6.5 Conclusion
Abstract

Prostate cancer (PCa) is the most common cancer in men. It is a heterogeneous disease and currently there are no reliable biomarkers available to stratify men for prostate biopsy (PBx) and treatment. Hence, there is a risk of over-diagnosing insignificant disease, or under-diagnosing significant disease. We aimed to evaluate Prostate Cancer gene 3 (PCA3, FDA approved) and N6-methyladenosine (m6A) for diagnostic properties.

PCA3 is a long non-coding RNA (ncRNA) that is unstable, has an unclear biological role and is expensive to chemically treat to prevent degradation prior to analysis. Long ncRNAs are degraded into shorter forms, we explored whether this was the fate for PCA3. We identified a short segment of RNA within intron 1 of PCA3 bioinformatically which we termed PCA3 short RNA2 (PCA3-shRNA2). The expression of this short RNA correlated to that of PCA3 in PCa cell lines, urinary samples and PBx tissue. PCA3-shRNA2 was overexpressed in urinary samples obtained from men with PCa compared to BPH, was regulated by testosterone and had a diagnostic accuracy similar to that of PCA3. We identified oncogenic mRNA targets of PCA3-shRNA2 and found that COPS2 was underexpressed in cancerous urinary samples.

There are over a hundred RNA modifications described and methylation of N6-adenosine base is the most common methylated site. m6A is reversible and may be involved in oncogenesis. We profiled m6A in PCa cell lines by immunoprecipitation and RNA sequencing and found oncogenic RNAs (e.g. PARG) that were differentially expressed in LNCaP-LN3 cells.

We identified a novel RNA within PCA3 that is easy to measure, overexpressed in PCa samples and appeared to target oncogenic mRNAs. We profiled m6A in PCa cell lines and have identified N6-adenosine methylated RNAs associated with PCa development. In conclusion PCA3-shRNA2 and m6A have evolving roles in cancer and may function well as biomarkers.
Foreword

Urological Surgery was introduced to me formerly when I was a Foundation doctor at Oxford University Hospitals. My Academic Urology career started when I was employed through national selection by Professor James Catto.

Each day I see elderly men in the assessment unit, on the wards, or in the outpatients department presenting with signs and symptoms related to prostate cancer. The disease is diverse and can have a high impact on quality of life and life expectancy. Current diagnosis and treatment is not based on any molecular stratification, resulting in over- or under-diagnosis/treatment of the disease.

To improve diagnosis and stratification for treatment, I decided to go back and study prostate cancer at the molecular level. Hence, I undertook this PhD.

Within Sheffield, around 10,400 men participated in the ProtecT study and 1,500 men participated in the ProMPT study. They came with no obligations and were fully committed to help us unravel new knowledge to help the next generation of patients. I am indebted to these participants, some of whom have passed away in the three years it took to complete this PhD. I hope my work adds to the scientific community, and hope that we will improve patient care with the implementation of this knowledge.
Acknowledgments

First and foremost, I would like to thank my supervisor Professor James Catto for selecting me through hurdles of applications and interviews to become an Academic Urology trainee. He gave me opportunities to conduct both clinical and basic science research, and guided me through series of successful and unsuccessful grant applications. The day he offered me the job, he said that I have a glittering career ahead and that he would hope to shape and channel this. He has not failed to do so.

I would also like to thank my second supervisor Dr Ishtiaq Rehman for giving me advice in the lab and reviewing my thesis. I thank my personal tutor Professor Stephen Wharton for keeping me on track with completing my thesis. Special thanks go to Professor Stuart Wilson for allowing me to conduct experiments in his laboratory under the supervision of himself and his team. I am grateful for the assistance and support provided by my colleagues and lab mates Saiful, Ross, Marcus, Simon, Anil, Sheila, Maggie and Helen.

I thank Urology Consultant and Senior Lecturer Mr Derek Rosario for supervising the clinical part of this thesis, and as Urology Training Programme Director, for making the required arrangements for me to go out of, and back into the Urology Training Programme.

I would like to thank all the patients who consented in participating in our research studies and who provided the samples analysed in this PhD.

I am grateful for the financial support I received from the Royal College of Surgeons of England and The Urology Foundation, which allowed me to complete all the work necessary for this thesis.

Finally, I owe all my successes to my Mother Stella and my Father John. I thank them for my upbringing and facilitating the steps I took in becoming a doctor.
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADT</td>
<td>Androgen deprivation therapy</td>
</tr>
<tr>
<td>APC</td>
<td>Adenomatous polyposis coli gene</td>
</tr>
<tr>
<td>AR</td>
<td>Androgen receptor</td>
</tr>
<tr>
<td>AS</td>
<td>Active surveillance</td>
</tr>
<tr>
<td>ASAP</td>
<td>Atypical small acinar proliferation</td>
</tr>
<tr>
<td>BAK</td>
<td>Bcl antagonist/killer</td>
</tr>
<tr>
<td>BAX</td>
<td>Bcl-associated X</td>
</tr>
<tr>
<td>Bcl-2</td>
<td>B-cell lymphoma 2</td>
</tr>
<tr>
<td>BPH</td>
<td>Benign prostate hyperplasia</td>
</tr>
<tr>
<td>BRAF1</td>
<td>v-raf murine sarcoma viral oncogene homolog B1</td>
</tr>
<tr>
<td>BSA</td>
<td>Bovine serum albumin</td>
</tr>
<tr>
<td>CDK</td>
<td>Cyclin dependent kinase</td>
</tr>
<tr>
<td>CES</td>
<td>Cauda equina syndrome</td>
</tr>
<tr>
<td>CHD</td>
<td>Chromodomain helicase DNA binding protein</td>
</tr>
<tr>
<td>COSMIC</td>
<td>Catalogue of Somatic Mutations in Cancer</td>
</tr>
<tr>
<td>CRPC</td>
<td>Castration-resistant prostate cancer</td>
</tr>
<tr>
<td>CSAP</td>
<td>Cryosurgical ablation of prostate</td>
</tr>
<tr>
<td>CTA</td>
<td>Cancer testes antigen</td>
</tr>
<tr>
<td>CT scan</td>
<td>Computerised tomography scan</td>
</tr>
<tr>
<td>CYPs</td>
<td>Cytochrome P450 enzymes</td>
</tr>
<tr>
<td>DHT</td>
<td>Dihydrotestosterone</td>
</tr>
<tr>
<td>DMSO</td>
<td>Dimethyl-sulfoxide</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>DNMT</td>
<td>DNA methyltransferase</td>
</tr>
<tr>
<td>DRE</td>
<td>Digital rectal examination</td>
</tr>
<tr>
<td>DTX2</td>
<td>Deltex E3 Ubiquitin Ligase 2</td>
</tr>
<tr>
<td>EAU</td>
<td>European association of urology</td>
</tr>
<tr>
<td>EBRT</td>
<td>External beam radiotherapy</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>EMT</td>
<td>Epithelial-mesenchymal transition</td>
</tr>
<tr>
<td>EP300</td>
<td>E1A binding protein p300 gene</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>ERSPC</td>
<td>European Randomised Study of Screening for Prostate Cancer</td>
</tr>
<tr>
<td>ETS</td>
<td>E26 transformation-specific family</td>
</tr>
<tr>
<td>ETV</td>
<td>ETS variant</td>
</tr>
<tr>
<td>EZH2</td>
<td>Histone-lysine N-methyltransferase</td>
</tr>
<tr>
<td>FCS</td>
<td>Fetal calf serum</td>
</tr>
<tr>
<td>FLI1</td>
<td>Friend leukaemia virus integration 1</td>
</tr>
<tr>
<td>FOXA1</td>
<td>Forkhead-box A1</td>
</tr>
<tr>
<td>FTO</td>
<td>Fat mass and obesity-associated protein</td>
</tr>
<tr>
<td>GSTP1</td>
<td>Glutathione S-transferase P1</td>
</tr>
<tr>
<td>GWAS</td>
<td>Genome-wide association studies</td>
</tr>
<tr>
<td>HDAC</td>
<td>Histone deacetylase</td>
</tr>
<tr>
<td>HDM</td>
<td>Histone demethylase</td>
</tr>
<tr>
<td>HGPIN</td>
<td>High-grade prostatic intraepithelial neoplasia</td>
</tr>
<tr>
<td>HIFU</td>
<td>High-intensity focused ultrasound</td>
</tr>
<tr>
<td>HMT</td>
<td>Histone methyltransferase</td>
</tr>
<tr>
<td>IL</td>
<td>Interleukin</td>
</tr>
<tr>
<td>IP</td>
<td>immunoprecipitation</td>
</tr>
<tr>
<td>KDM6A</td>
<td>Lysine(K)-specific demethylase 6A</td>
</tr>
<tr>
<td>KLK</td>
<td>Kallikrein</td>
</tr>
<tr>
<td>KRAS</td>
<td>v-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog</td>
</tr>
<tr>
<td>LUTS</td>
<td>Lower urinary tract symptoms</td>
</tr>
<tr>
<td>M6A</td>
<td>N6-methyladenosine</td>
</tr>
<tr>
<td>MAB</td>
<td>Maximal androgen blockade</td>
</tr>
<tr>
<td>MAPK</td>
<td>Mitogen-activated protein kinase</td>
</tr>
<tr>
<td>MeRIP-seq</td>
<td>Methylated RNA immunoprecipitation sequencing</td>
</tr>
<tr>
<td>METTL</td>
<td>Methyltransferase like</td>
</tr>
<tr>
<td>miRNA</td>
<td>MicroRNA</td>
</tr>
<tr>
<td>MLL2</td>
<td>Myeloid/lymphoid leukaemia 2 gene</td>
</tr>
<tr>
<td>(mp)MRI</td>
<td>(multiparametric) Magnetic resonance imaging</td>
</tr>
<tr>
<td>mTOR</td>
<td>Mammalian target of rapamycin</td>
</tr>
<tr>
<td>MYC</td>
<td>v-myc myelocytomatosis viral oncogene homolog gene</td>
</tr>
<tr>
<td>NaCl</td>
<td>Sodium Chloride</td>
</tr>
<tr>
<td>NCOA2</td>
<td>Nuclear receptor coactivator 2 gene</td>
</tr>
</tbody>
</table>
NCOR2  Nuclear receptor corepressors 2 gene
ncRNA  Non-coding RNA
NEAA  Non-essential amino acid
NRIP1  Nuclear receptor interacting protein 1 gene
PAR  Poly (ADP-ribose)
PARG  Poly (ADP-ribose) Glycohydrolase
PARP1  Poly (ADP-ribose) polymerase 1
PBS  Phosphate based saline buffer
(r)PBx  (repeat) Prostate biopsy
PCa  Prostate cancer
PCR  Polymerase chain reaction
PDCD4  Programmed cell death 4
PHI  Prostate Health Index test
PI3K  Phosphoinositide 3-kinase
PIVOT  Prostate Cancer Intervention Versus Observation Trial
PLCO  The Prostate, Lung, Colon and Ovaries screening trial
PROMIS  PROstate MRI Imaging Study
ProMPT  Prostate Cancer: Mechanisms of progression and treatment trial
ProtecT  Prostate testing for cancer and treatment Trial
PSA  Prostate-specific antigen
PSA-DT  PSA doubling time
PSAV  PSA velocity
PTEN  Phosphatase and tension homolog gene
QoL  Quality of life
qRT-PCR  Quantitative real-time polymerase chain reaction
RAF1  v-raf-1 murine leukaemia viral oncogene homolog 1
RARP  Robotic-assisted radical prostatectomy
RASSF1  Ras association domain-containing protein 1
RB  Retinoblastoma
RCT  Randomised controlled trial
REDUCE  Reduction by Dutasteride of Prostate Cancer trial
RNA  Ribonucleic acid
RP  Radical prostatectomy
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>RRP</td>
<td>Retropubic radical prostatectomy</td>
</tr>
<tr>
<td>RT</td>
<td>Radiotherapy</td>
</tr>
<tr>
<td>RVC</td>
<td>Ribonucleoside vanadyl complexes</td>
</tr>
<tr>
<td>SAM</td>
<td>S-adenosylmethionine</td>
</tr>
<tr>
<td>SCNA</td>
<td>Somatic copy number aberrations</td>
</tr>
<tr>
<td>SIRT1</td>
<td>Silent information regulator 1</td>
</tr>
<tr>
<td>SPCG-4</td>
<td>Scandinavian Prostate Cancer Group-4 RCT</td>
</tr>
<tr>
<td>SPINK</td>
<td>Serine peptidase inhibitor, Kazal type 1</td>
</tr>
<tr>
<td>SPOP</td>
<td>Speckle-type POZ protein gene</td>
</tr>
<tr>
<td>STAMPEDE</td>
<td>Systemic therapy in advancing or metastatic prostate cancer: Evaluation of drug efficacy</td>
</tr>
<tr>
<td>TMPRSS2</td>
<td>Transmembrane protease, serine 2</td>
</tr>
<tr>
<td>TP53</td>
<td>Tumour protein 53</td>
</tr>
<tr>
<td>TPM-Bx</td>
<td>Transperineal template biopsy</td>
</tr>
<tr>
<td>Tris-HCl</td>
<td>Tris-hydrochloride buffer</td>
</tr>
<tr>
<td>TRUS-PBx</td>
<td>Transrectal ultrasound-guided prostate biopsy</td>
</tr>
<tr>
<td>TSP-1</td>
<td>Thrombospondin-1</td>
</tr>
<tr>
<td>TSPY</td>
<td>Testis specific protein, Y-linked</td>
</tr>
<tr>
<td>TURP</td>
<td>Transurethral resection of the prostate</td>
</tr>
<tr>
<td>US</td>
<td>Ultrasound</td>
</tr>
<tr>
<td>VEGF</td>
<td>Vascular endothelial growth factor</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organisation</td>
</tr>
<tr>
<td>WW</td>
<td>Watchful waiting</td>
</tr>
<tr>
<td>YTHDF</td>
<td>YTH domain family member</td>
</tr>
</tbody>
</table>
Grant funding for Support of this Work

1) The Urology Foundation: One Year Research Scholarship.
   £50,000. May 2016-May 2017

2) The Urology Foundation: One Year Research Scholarship.
   £46,513. Feb 2015-Dec 2015

3) Royal College of Surgeons of England (Together with The Rosetrees Trust): One Year Research Fellowship.
   £53,500. Feb 2014-Feb 2015

4) Department of Oncology and Metabolism, University of Sheffield: Conference Fund (Learned Society). £600. UAA, Hong Kong, Aug 2017

5) Department of Oncology and Metabolism, University of Sheffield: Conference Fund (Learned Society). £600. UAA, Singapore, Jul 2016

6) Department of Oncology, University of Sheffield: Conference Fund. £750. SIU, Melbourne, Oct 2015
List of Achievements arising from the work undertaken in this thesis

Peer-reviewed publications


Abstract publications

1) Pang KH, Lesbirel S, Cumberbatch MG, Wilson SA, Catto JWF. Identification of N6-adenosine Methylated RNAs by m6A-Seq within Prostate Cancer. IJU Suppl 2016 Jul;23: p1-190 (131)

2) Pang KH, Cumberbatch MG, Wilson SA, Catto JW. N6-methyladenosine mRNA methylation in common human cancers: Identifying a potential role in androgen-resistance within prostate cancer. Eur Urol Suppl 2015 Apr; 14/2; e527
3) **Pang KH, Miah S, Drayton RM, Rehman I, Clarke R, Stoehr R, Hartmann A, Lavin M, Bryant HE, Martens-Uzunova ES, Jenster G, Hamdy FC, Gardiner RA, Catto JW. Identification and diagnostic performance of a small RNA within the PCA3 and BMCC1 gene locus that potentially targets mRNA. Eur Urol Suppl 2015 Apr; 14/2; e426**

Other publications

1) **Pang KH.** Evaluation of the diagnostic role of a small RNA within PCA3 in prostate cancer. Royal College of Surgeons of England Surgical Research Report 2017/18
   
   [https://www.rcseng.ac.uk/standards-and-research/research/surgical-research-reports/](https://www.rcseng.ac.uk/standards-and-research/research/surgical-research-reports/). Page 43

Oral presentations


that potentially targets mRNA. EAU 2015. Madrid, Spain. Mar 20-24 2015 (Best Poster of Biomarker Session Prize)


Poster presentations


Prizes

1) 1st Prize Best Oral Presentation of a Poster- PCA3 and Prostate Cancer. ‘Biomarker Session’, 30th EAU Congress Mar 2015, Madrid
List of Tables

Table 1. International Society of Urological Pathology 2014 grade groups. 43
Table 2. Tumour Node Metastasis (TNM) classification of prostate cancer. 44
Table 3. Complication rates per biopsy session. 55
Table 4. European Association of Urology risk groups for localised and locally advanced prostate cancer. 56
Table 5. MicroRNAs expression in prostate cancer. 77
Table 6. Prostate cell lines. 93
Table 7. Baseline characteristics of men with an initial negative prostate biopsy. 115
Table 8. Baseline characteristics of patients and prostate biopsy outcomes in each biopsy group. 116
Table 9. Patient and prostate biopsy characteristics in positive cohorts. 120
Table 10. Clinical and pathological characteristics of patients with and without prostate cancer at the time of analysis. 124
Table 11. Univariable and multivariable analysis of predictors for tumours detected on repeat biopsy. 125
Table 12. Published international data on repeat prostate biopsy outcomes. 127
Table 13. Predicted hairpin RNAs within PCA3 and BMCC1. 138
Table 14. Clinical and histopathological characteristics of patients and samples. 146
Table 15. Predicted mRNA targets for PCA3-shRNA2. 156
Table 16. Selected potential mRNA targets of PCA3-shRNA2. 162
Table 17. Primers and condition used to detect target mRNAs. 164
Table 18. Patients and FFPE samples analysed in this report. 175
Table 19. RNA expression stratified by eventual diagnosis. 180
Table 20. Total number of cancer and control samples within each cancer type. 192
Table 21. Summary of the 47 microarray datasets extracted from Array Express. 193
Table 22. Functional clusters common in both m6A(+) and m6A(-). 233
Table 23. LNCaP and LNCaP-LN3 MeRIP-seq read quality control. 237
Table 24. LNCaP and LNCaP-LN3 reads mapping rate. 239
Table 25. m6A peak calling. 240
List of Figures

Figure 1. The hallmarks of cancer. .................................................................26
Figure 2. Anatomy of the prostate. ..............................................................31
Figure 3. Estimated age-standardised rates (world) per 100,000. ...............36
Figure 4. Age-standardised incidence and mortality rate (per 100,000 UK men) of prostate cancer. ..............................................................37
Figure 5. Signs and symptoms of prostate cancer ......................................40
Figure 6. Gleason grading schematic diagram. ...........................................42
Figure 7. Location of PCA3. .....................................................................49
Figure 8. Summary of the treatment options for prostate cancer ...............57
Figure 9. The hypothalamus-pituitary axis. ..............................................64
Figure 10. Androgen biosynthesis .............................................................65
Figure 11. A summary of the events leading to oncogenesis ....................67
Figure 12. Illustration of RNA synthesis and regulation of mRNA ...........75
Figure 13. The N6-methyladenosine machinery. ........................................83
Figure 14. Flow chart of the study population within the ProtecT/ProMPT study. ....90
Figure 15. Percentage of prostate cancer identified on prostate biopsy ...117
Figure 16. Kaplan-Meier estimate of time from date of first prostate cancer clinic to diagnosis of a) all cancer and b) high-grade cancer .........................122
Figure 17. Identification of PCA3-shRNA2 .............................................139
Figure 18. Alignment of the transcriptomic sequences to the genome ........140
Figure 19. The genomic origin of PCA3-shRNA2 ....................................141
Figure 20. BLAST results of the PCA3-shRNA2 sequence ......................142
Figure 21. Expression of PCA3 and PCA3-shRNA2 in cell lines representing prostate cancer and other malignancies .............................143
Figure 22. The correlation between PCA3-shRNA2 and PCA3/BMCC1 ...144
Figure 23. The correlation between PCA3-shRNA2 and PCA3 in urinary RNA .........147
Figure 24. The expression of PCA3-ShRNA2, PCA3 and BMCC1 in urinary RNA ......148
Figure 25. The correlation between PCA3-shRNA2 and PCA3 in a large urinary cohort ........................................................................149
Figure 26. The expression of PCA3-shRNA2 in benign and cancerous urinary samples. ........................................................................150
Figure 27. The expression of PCA3-shRNA2 in different stages of disease ....151
Figure 28. Identification of disease by PCA3-shRNA2 ................................153
Figure 29. Androgen regulated expression of PSA, PCA3 and PCA3-shRNA2 ......154
Figure 30. Expression of PCA3-shRNA2 and primary/mature miRNAs according to nuclear and cytoplasmic localization ..................................160
Figure 31. Transfection of DU145 with PCA3-shRNA2 and controls ..........163
Figure 32. The correlation between COPS2 and PCA3-shRNA2 in urinary samples .................................................................166
Figure 33. The expression of COPS2 in urinary samples .............................167
Figure 34. Scatterplot of PCA3 and PCA3-shRNA2 RNA expression normalized to PSA mRNA in FFPE benign prostate biopsies ...............................177
Figure 35. PSA, PCA3 and PCA3-shRNA2 expression in FFPE prostate biopsies stratified by year of collection .................................................178
Figure 36. Box plot of PCA3 and PCA3-shRNA2 expression stratified for eventual diagnosis. .................................................................181
Figure 37. PSA, PCA3 and PCA3-shRNA2 expression in tissues with respect to 
Gleason grade. ...........................................................................182
Figure 38. Predictive ability of PCA3 and PCA3-shRNA2 expression for the eventual 
diagnosis of prostate cancer. .........................................................183
Figure 39. Proportion of differentially expressed RNA susceptible to N6-adenosine 
methylation. ..................................................................................204
Figure 40. Volcano plots of a) all mRNAs and b) m6A susceptible mRNAs within 
prostate cancer. .............................................................................205
Figure 41. Volcano plots of expression of all mRNAs and m6A(+) within a) all 
cancers, b) bladder, c) kidney, d) breast, e) lung, f) ovarian, g) colorectal, h) 
gastric cancer. ..................................................................................206
Figure 42. Average fold change of aberrantly expressed RNAs within each cancer 
and across all cancer types...............................................................214
Figure 43. Proportion of differentially expressed RNA susceptible to N6-adenosine 
methylation within a) top and b) bottom 10%-fold change percentile. ......216
Figure 44. The mean fold changes in a) up and b) downregulated N6- 
methyladenosine susceptible RNAs in each cancer type. ................219
Figure 45. Functional annotation of aberrantly expressed genes shared in prostate 
cancer. ............................................................................................221
Figure 46. Functional annotation of aberrantly expressed m6A(+) and m6A(-) in a) 
bladder, b) kidney, c) breast, d) lung, e) ovarian, f) colorectal, g) gastric cancer. .................................................................222
Figure 47. Annotation clusters common in each cancer type ................229
Figure 48. Functional annotation of aberrantly expressed genes shared in all cancer. ..............................231
Figure 49. Functional activities common in both m6A susceptible and non- 
susceptible transcripts. .................................................................232
Figure 50. Validation of RNA size. ......................................................234
Figure 51. Concentration of Genes whose transcripts are known to undergo N6- 
adenosine methylation ....................................................................236
Figure 52. Plot of base qualities for LNCaP-LN3-input. .......................238
Figure 53. Sequence logo. .................................................................240
Figure 54. Methylated transcripts that are differentially expressed in prostate 
cancer cell lines. .............................................................................242
Figure 55. The role of PARG and PARP in DNA repair. .........................244
Figure 56. Levels of PARP activity and mRNA in LNCaP and LNCaP-LN3 ........245
CHAPTER 1: Introduction
1.1 Cancer background
The National Cancer Institute defines cancer as an abnormal growth of cells which tend to proliferate in an uncontrolled way and in some cases, metastasize and invade other tissues. Cancer can arise from different organ structures and develops as a result of abnormal genetic and/or epigenetic events.

The International Agency for Research on Cancer (WHO, GLOBOCAN Project) estimates around 14.1 million new cancer cases, 8.2 million cancer deaths and 32.6 million people living with cancer in 2012 worldwide (GLOBOCAN, 2012).

Cancer was first described in Egypt in 1600 BC. Edwin Smith Papyrus was the ancient Egyptian textbook on trauma surgery which described eight cases of breast tumours/ulcers that were removed by cauterization with a ‘fire drill’.

The ‘Father of Medicine’, Greek physician Hippocrates (460-370 BC) used the term carcinos and carcinoma to describe tumours. The Greek word translates to ‘crab’ since tumours possess finger-like spreading projections and have the crab-like tenacity to grasp and invade tissues. In 28-50 BC, the Roman physician, Celsus translated the Greek term into cancer, the Latin word for crab. Galen (130-200 AD), another Greek physician, used oncos to describe malignant tumours, which is now used to form the name of the cancer specialty, ‘oncology’ (Hajdu, 2011).

1.1.1 Pathogenesis of cancer
Cancer arises from alterations of complex biological mechanisms. It is a disease of disruption in cell/tissue growth regulation. Under normal circumstances, cell growth is regulated by oncogenes (cell growth promoters) and tumour suppressor genes (cell growth inhibitors). Genetic changes as a result of inherited genes (5-20% germ-line mutations), or environmental factors (90-95% somatic events) such as smoking, radiation and
environmental pollutants, lead to an imbalance of oncogenes and tumour suppressor genes causing cancer development (oncogenesis or tumorigenesis). Disruption can occur at any stage from within the nucleus such as, DNA modifications and DNA transcription to RNA, to within the cytoplasm such as RNA modifications, microRNA (miRNA) regulations (epigenetic events), RNA translation to proteins, and post-translational protein modifications (Croce, 2008).

The Catalogue of Somatic Mutations in Cancer (COSMIC) (http://cancer.sanger.ac.uk) is the world’s largest and most comprehensive database for somatic mutations in human cancer. In 2004, four genes were described when COSMIC was launched (Bamford et al., 2004), and now around two million coding point mutations in over one million tumour samples across most human genes are described (Forbes et al., 2015).

Hanahan and Weinberg proposed that six hallmarks of cancer form a structured principle that allows a model for understanding the complexities of oncogenesis. In 2000, the authors identified the six hallmarks as sustaining proliferative signalling, evading growth suppressors, resisting cell death, enabling replicative immortality, inducing angiogenesis, and activating invasion and metastasis. Underlying these mechanisms are genome instability, which induces genetic diversity (Hanahan et al., 2000). Progress in the last couple of decades has added two emerging hallmarks-reprogramming of energy metabolism and evading immune destruction (Hanahan et al., 2011).

1.1.2 Hallmarks of cancer
The development of cancer and its ability to survive and spread is dependent on various important biological processes that become dysregulated. The six hallmarks of cancer described by Hanahan and Weinberg are shown in Figure 1). This section gives an overview of the features of oncogenesis. Specific mechanisms related to prostate oncogenesis are discussed in more details later on.
**Figure 1.** The hallmarks of cancer.

It has been described that cancer cells are able to alter homeostatic biological mechanisms and possess the six characteristics illustrated in this figure in order to survive, proliferate and invade (Hanahan et al., 2011)
1.1.2.1 Sustaining proliferative signalling
One of the most important features of oncogenesis, is growth. Cancer cells have the ability to sustain chronic cell growth and division (proliferation) through dysregulating growth promoting signals that instruct entry into and progression through the cell division cycle within normal tissues. Key growth signalling pathways that are interrupted by cancer includes MAP-kinase (Mitogen-activated protein kinase), mTOR-kinase (Mammalian target of rapamycin) and PI3K (Phosphoinositide 3-kinase). Cancer cells can either produce signalling factors themselves, or stimulate normal cells to release proliferative signals (Bhowmick et al., 2004).

1.1.2.2 Evading growth suppressors
In well-controlled cell cycles, growth suppressors act to inhibit overgrowth. In order to maintain continuous growth, cancer cells possess mechanisms to bypass suppression via inhibition of two main pathways, RB (Retinoblastoma) and TP53 (Tumour protein p53) pathways. The RB protein integrates signals from extracellular sources, whilst TP53 receives inputs from abnormal intracellular stress. Both regulatory circuits are gatekeepers of cell cycle progression (Sherr et al., 2002).

1.1.2.3 Activating invasion and metastasis
Cancer cells grow locally and have capabilities to invade local structures and spread to distant sites through metastasis. This process is termed the invasion-metastasis cascade. Cancer cells intravasate into nearby blood and lymphatic vessels, transit through the haematogenous and lymphatic systems and extravasate into distant tissues forming micrometastases. E-cadherin is a cell-to-cell adhesion molecule which assembles cell sheets and maintain cell quiescence. The invasion-metastasis cascade is associated with a decrease expression of E-cadherin, and the uncontrolled epithelial-mesenchymal transition (EMT) regulatory program. The EMT program can be activated transiently and enable cancer cells to invade, resist apoptosis and disseminate (Thiery et al., 2009).
1.1.2.4 Enabling replicative immortality
To form macroscopic tumours, cancer cells require unlimited replication. Within normal cells, excessive proliferation is limited by two barriers, senescence, which is the irreversible entrance into a non-proliferative, viable state; and crisis, which involves cell death. Cancer cells bypass senescence and crisis, and transit into immortalization exhibiting unlimited replicative potential. Telomeres are partly responsible for unlimited proliferation and are structures attached to the end of chromosomes protecting chromosomal DNAs from end-to-end fusions (Blasco, 2005). Cancer cells maintain telomeric DNA lengths to avoid initiating senescence or apoptosis, achieved most commonly by increasing the expression of telomerase.

1.1.2.5 Inducing angiogenesis
Like normal tissue and cells, tumours require nutrients and oxygen, and removal of metabolic waste and carbon dioxide. Following vasculogenesis and angiogenesis during embryogenesis, the normal vasculature becomes quiescent. Angiogenesis only occurs following stress such as, wound healing or female menstruation. In tumour growth, there is a persistent 'angiogenic switch' causing quiescent vasculature to expand (Baeriswyl et al., 2009). The defect partly lies upon the angiogenesis inducer, VEGF (Vascular endothelial growth factor) and inhibitor, TSP-1 (Thrombospondin-1).

1.1.2.6 Resisting cell death
Programmed cell death by apoptosis is an important activity involved in preventing cancer development. Apoptosis can be triggered by various physiological stresses and is regulated in part by pro- (Bax and Bak) and anti-apoptotic (Bcl-2) proteins. The apoptotic signals trigger a cascade of proteolysis mediated by caspases. Disassembled cells are then consumed by phagocytic cells (Adams et al., 2007). Autophagy is an intracellular degradation activity that is caspase-independent and is mediated through lysosomes (Levine et al., 2008). Cancer cells have the ability to avoid apoptotic and autophagic mechanisms and continue to proliferate in an uncontrolled manner.
1.1.3 Emerging hallmarks and characteristics
There are two additional hallmarks of cancer that have been described, including deregulating cellular metabolism and evading the immune system (Hanahan et al., 2011).

1.1.3.1 Reprogramming energy metabolism
In order to sustain the six core hallmarks described above, cancer cells require sufficient energy. Normal cells produce energy through aerobic metabolism (using oxygen) and glycolysis (metabolizing glucose). Cancer cells have the ability to reprogram energy metabolism and exhibit a metabolic switch to provide continuous glycolytic fuelling. There is a state of upregulation of glucose transporters and multiple enzymes of the glycolytic pathway (DeBerardinis et al., 2008; Hsu et al., 2008).

1.1.3.2 Evading immune destruction
The immune system serves as a protective mechanism in recognizing and eliminating infected, and possibly cancerous cells. In mice studies, tumours developed more abundantly in immunodeficient mice compared to immunocompetent controls, suggesting that the immune system has a role in oncogenesis (Kim et al., 2007). Proposed mechanisms include the ability of cancer cells to paralyze infiltrating cytotoxic T-lymphocytes and natural killer cells (Yang et al., 2010), or recruiting immunosuppressive cells such as, T-regulatory cells (Mougiakakos et al. 2010).

1.1.4 Sustaining core and emerging hallmarks
Two evolving characteristics of cancer facilitate the process of both core and emerging hallmarks: genomic instability and mutation, and tumour-promoting inflammation.

1) Genomic instability and mutation
Alterations in the genomes of cancer cells results in heritable phenotypes such as, inactivation of tumour suppressors and evading apoptosis. The mutation of genes is achieved through increased sensitivity to mutagenic
factors or through defects in the genomic maintenance machinery. The proteins that are involved in detecting DNA damage, repairing DNA, and inactivating mutagenic molecules are defective (Negrini et al., 2010). The result is abnormal and uncontrolled gene expression of factors that promote oncogenesis and factors that inhibit tumour suppression through the six core and two merging cancer hallmarks.

2) Tumour-promoting inflammation

It is thought that evading the immune system is an emerging hallmark, and this theory is consolidated by the fact that tumours are infiltrated by both innate and adaptive cells on histopathological examination of tumour specimens. As cancer cells interact with immune cells evading clearance, they enhance the release of inflammatory mediators which gives rise to the unanticipated effect of enhancing tumour progression. This is achieved by the release of inflammatory mediators that facilitate growth, angiogenesis, and invasion and metastasis (Grivennikov et al., 2010).

1.2 The prostate gland

1.2.1 Anatomy and histology of the prostate

1.2.1.1 Surface anatomy

The prostate gland is a pyramid-shaped structure with its apex directing downwards and its base directing upwards towards the bladder. The gland lies below the urinary bladder and it situated in front of the rectum. The prostate weighs ~20g and is ~3cm long, 4cm wide and 2cm thick (Figure 2) in young men. The prostate gland is anatomically split into zones, the transition zone (5-10% of prostate volume) surrounds the urethra, the peripheral zone (70% of prostate volume) includes the peripheral sections of the gland, and the central zone (25% of prostate volume) is located between the transition and peripheral zones. The ejaculatory ducts run through the central zone (Wein et al., 2015).
The prostatic urethra runs through the prostate and divides the gland into left and right lateral lobes. The posterior aspects of the lateral lobes and the median sulcus are palpable by digital rectal examination (DRE). The paired ejaculatory duct enters the prostate and opens into the middle of the prostatic urethra at the seminal colliculus (verumontanum).

The prostate is encapsulated by a dense capsule and is fixed to the pubic bone by two puboprostatic ligaments. The endopelvic fascia covers the prostate ventrally and extends to both sides and covers the levator ani muscle. The Denovilliers' fascia separates the prostate dorsally from the rectum (Wein et al., 2015).

**Figure 2. Anatomy of the prostate.**

The prostate gland is situated below the bladder and produces substances which contribute to semen. The prostatic urethra and ejaculatory duct (from seminal vesicles) run through the prostatic transition zone and transports urine and ejaculate (Based on graphics created by Cancer Research UK, http://www.cancerresearchuk.org/about-cancer/prostate-cancer/about).
1.2.1.2 Histology
The prostate is formed of tubuloalveolar glands, and have pseudostratified columnar epithelium. The columnar cells contain secretory granules. Basal cells are located between the columnar cells and are the fundamental cells for epithelial regeneration. Each gland is embedded in fibromuscular stroma (70% of the prostate mass) containing connective tissue and smooth muscle. During ejaculation, the smooth muscles contract and expulse glandular content (McNeal, 1981).

1.2.1.3 Vascular, lymphatic and nerve supply
The blood supply originates from the internal iliac vessels. The inferior vesical artery divides into urethral and capsular branches, and the urethral branches enter the basal prostate and bladder neck at 4 and 8 o’clock positions, supplying the transition zone. The capsular branches join the cavernous nerves laterally and run to the pelvic floor giving rise to smaller arteries that perforate the capsule. Additional arterial supply comes from the middle rectal artery, internal pudendal artery and obturator artery.

The venous vessels drain via the vesicoprostatic plexus (deep venous complex) to the internal iliac veins. The vesicoprostatic plexus lies under the puboprostatic ligaments and pubic bone, where the blood from the deep penile vein joins the plexus.

The lymphatic drainage of the prostate drains to the obturator and internal iliac nodes. In addition, there is lymphatic communication with the external iliac, presacral and para-aortic lymph nodes.

The autonomic innervation reaches the prostate via the lateral cavernous nerves. Parasympathetic (S2-4) signals stimulate the glandular activity and the sympathetic (L1-2) nerves controls smooth muscle contraction through alpha-receptors (Wein et al., 2015).
1.2.1.4 Structures surrounding the prostate
The paired seminal vesicles are located behind the bladder and its ducts open into the ductus deferens and form the ejaculatory duct. The seminal vesicles consist of duct-like glandular tissue, about 15cm long with a muscle-containing wall. The seminal vesicle produces an alkaline (pH 7.4) secretion of gelatinous consistency which contains fructose and forms half of the semen volume. Sperm and testosterone is produced in the testicles and matured sperms are stored in the epididymis. The vas deferens is a tube that transports semen and sperm from the epididymis to the ejaculatory duct during ejaculation (Figure 2).

1.2.2 Functions of the prostate
The epithelial cells of the prostate secrete a glycoprotein enzyme called prostate-specific antigen (PSA). It is a member of the kallikrein-related peptidase family, hence it is also termed kallikrein-3 (KLK3). PSA is produced to liquefy semen in the ejaculate to allow transportation of sperm facilitating fertilization (Balk et al., 2003). The prostate secretes fluid that forms ~20% of the semen volume. The prostatic fluid is thin, acidic (pH 6.4) and contains spermine, spermidine, prostaglandins, zinc, citric acid, immunoglobulins, phosphatases and proteases. Constituents help to liquefy semen and provide nutrition and optimal conditions for sperms to travel and fertilize the ovum (Wein et al., 2015).

The prostate functions as a valve preventing urine flow during ejaculation. Micturition is controlled by parasympathetic activity, leading to bladder neck relaxation. Ejaculation is controlled by the sympathetic nerves, and is the result of contraction of the smooth muscle stroma.

To function adequately, the prostate needs androgens. Testosterone is produced mainly by the testicles, its metabolite, dihydrotestosterone (DHT) predominantly regulates the prostate.
1.3 Prostate cancer

Prostate cancer is the most common cancer in men and occurs in the peripheral zone of the prostate gland. This chapter describes the epidemiology, clinical presentation, investigation and management of PCa.

1.3.1 Epidemiology of prostate cancer

1.3.1.1 Worldwide

Approximately 1.1 million men were diagnosed with the disease worldwide in 2012, which accounts for around 15% of cancer diagnoses in men. Around 70% of these cases occur in more developed regions (GLOBOCAN, 2012). The incidence of PCa varies worldwide, by around 25% (Figure 3). The highest rates are seen in Australia, New Zealand, Northern America and in Western/Northern Europe due to the increasing practice of PCa testing and subsequent diagnosis on prostate biopsy (PBx) in these developed regions. Incidence rates are relatively high in some less developed areas including the Caribbean, Southern Africa and South America. The lowest rates are seen in Eastern/South-central Asia (Arnold et al., 2013). Prostate cancer is the fifth leading cause of cancer-death in men with an estimated 307,000 deaths worldwide in 2012. This accounts for approximately 6.6% of total male deaths. Due to the fact that PSA testing has a greater effect on incidence than on mortality, there is less variation in mortality rates worldwide compared to incidence. The number of deaths from PCa is larger in less developed regions (Figure 3). Mortality rates are generally high in predominantly black populations, very low in Asia and intermediate in the Americas and Oceania (GLOBOCAN, 2012). There is an expected rise in the disease’s economic burden associated with the increases in life expectancy and incidence of PCa. It is estimated that the costs of PCa in Europe exceed €8.43 billion, and this amounted to €106.7-179 million for all PCa patients diagnosed in 2006 (Luengo-Fernandez et al., 2013).

1.3.1.2 United Kingdom

According to Cancer Research UK, there were around 46,690 new diagnoses of PCa in 2014, which accounts for ~13% of all new cancers in
the UK. Over the past decade, PCa incidence rates have increased by around 5% and is expected to rise by 12% between 2014 and 2035 (Figure 4a). This is largely related to increasing PSA testing. Prostate cancer is the most common cancer in UK men and is the second most common cause of cancer death in UK men after lung cancer. In 2014, around 11,300 men died from PCa (CRUK, 2016). Over the last decade, the number of deaths have decreased by around 13% and is expected to decrease by 16% between 2014 and 2035 (Figure 4b).
Figure 3. Estimated age-standardised rates (world) per 100,000.

The incidence and mortality rates vary within different parts of the world secondary to PSA testing, genetics, individual risks factors and ethnicity. The incidence of PCa in the UK is within the top 25% globally (Adapted from GLOBOCAN 2012, http://globocan.iarc.fr/old/FactSheets/cancers/prostate-new.asp).
Figure 4. Age-standardised incidence and mortality rate (per 100,000 UK men) of prostate cancer.

a) The anticipated further rise in incidence is largely related to increase awareness of PCa and PSA testing; b) The anticipated further decrease in mortality rate is largely related to increasing PSA testing, early diagnosis and subsequent treatment of PCa (Adapted from the UK national cancer intelligence network, 2013, http://www.ncin.org.uk/publications/).
1.3.2 Risk factors and aetiology of prostate cancer

The aetiology of PCa is unclear, but the three well-defined risk factors are increasing age, ethnic origin (Kheirandish et al., 2011) and heredity (Hemminki, 2012). There is evidence for a genetic component to PCa based on two factors, namely family history and ethnic background (Hemminki, 2012; Jansson et al., 2012). In addition, genome-wide association studies (GWAS), including a meta-analysis of around 87,000 individuals have identified 100 common susceptibility loci contributing to PCa (Al Olama et al., 2014).

The risk is doubled if one first-degree relative has PCa. Approximately 9% have true hereditary PCa, defined as three or more affected relatives, or at least two relatives who have been diagnosed with early-onset (before age 55) PCa (Hemminki, 2012). Patients with hereditary PCa do not differ in other ways compared with spontaneous disease, apart from that they have an onset of disease, six-seven years earlier.

Other factors found to be associated with PCa include diet, alcohol consumption, chronic inflammation and occupational exposure (Nelson et al., 2003; Leitzmann et al., 2012). However, there is currently insufficient evidence to recommend lifestyle changes in order to decrease the risk of developing PCa (Richman et al., 2011). Research into selenium, Vitamin E (SELECT trial) (Lippman et al., 2009), and lycopene (meta-analysis of eight randomised-controlled trials, RCT) (Ilic et al., 2012) all did not show any negative correlations with PCa risk.

Metabolic syndrome is weakly associated with PCa risk and among single variables of the syndrome (meta-analysis) including body mass index, dyslipidaemia/glycaemia, high triglycerides, low high-density lipoprotein-cholesterol, only hypertension and waist circumference (>102cm) were associated with a greater risk of PCa by 15% and 56% respectively (Esposito et al., 2013).
The role of medications in PCa has also been investigated. In three parallel prospective studies, the use of testosterone in 1,023 hypogonadal men did not show an increased risk of PCa (Haider et al., 2015). The use of Aspirin (meta-analysis) and non-steroidal anti-inflammatory medications reveal conflicting data (Bhindi et al., 2014; Huang et al., 2014). 5-alpha reductase inhibitors (5-ARI) such as Finasteride and Dutasteride have been studied in the Reduction by Dutasteride of Prostate Cancer (REDUCE) RCT and although it appears that this class of medications have a potential benefit in preventing or delaying PCa development, this must be weighed against side effects and the potential risk of low-grade disease progressing to high-grade PCa. However, none of the 5-ARIs have been approved for this indication (Thompson et al., 2003; Andriole et al., 2010). The investigation into statins also did not reveal any preventive effect on PCa risk (Esposito et al., 2013; Freedland et al., 2013).

1.3.3 Signs and symptoms of prostate cancer

Most PCas are situated in the peripheral zone of the prostate gland and may be palpable on DRE when the volume is >0.2ml. Prostate cancer is detected in ~18% by DRE alone irrespective of PSA level. Abnormal DRE findings warrant a PSA test and PBx. Early PCa usually give rise to no symptoms. Approximately 30% of men present with no symptoms (Miller et al., 2003). Symptoms appear when cancer invades the urethra, or obstruct the urinary flow (Figure 5). Symptoms of haematuria or lower urinary tract symptoms (LUTS), including difficulty initiating, hesitancy, poor stream and terminal dribbling may occur (Kupelian et al., 2006). In severe obstruction, the risk of urinary retention with or without kidney injury can occur. In advanced or metastatic disease, patients may complain of back pain (bony metastasis), fatigue or lethargy. When the cancer affects the spinal nerves causing cauda equina syndrome (CES), urgent assessment and management in the form of steroids, MRI (Magnetic resonance imaging) and radiotherapy/surgical decompression are needed. Symptoms/signs of CES include, back pain, lower limb weakness/numbness, perineal (saddle) anaesthesia and urinary/faecal disturbance (Dy et al., 2008).
If cancer is not obstructing urine outflow, there may be no symptoms. Obstructive symptoms include lower urinary tract symptoms (poor stream, dribbling) and in complete obstruction, urinary retention may occur. In bone metastasis, men may present with bony pain and present with signs and symptoms of cauda equina syndrome (Based on graphics created by Cancer Research UK, http://www.cancerresearchuk.org/about-cancer/prostate-cancer/symptoms).

**Figure 5.** Signs and symptoms of prostate cancer.
1.3.4 Histopathology of prostate cancer

The most common histopathological form of PCa is acinar adenocarcinoma (~95%), which is a malignant tumour formed from glandular structures in epithelial tissue. Non-acinar carcinoma variants accounts for around 5-10% of primary PCa. These histological variants include sarcomatoid carcinoma, urothelial carcinoma, squamous cell carcinoma, adenosquamous carcinoma, ductal adenocarcinoma, basal cell carcinoma and neuroendocrine tumours, including small-cell carcinoma and clear cell adenocarcinoma (Humphrey, 2012; Humphrey et al., 2016).

There are histopathological types that have malignant potential, and patients with these pathologies need to be followed-up closely to identify progression. The two main pathologies associated with PCa are high-grade prostatic intraepithelial neoplasia (HGPIN) and atypical small acinar proliferation (ASAP), the latter being less significant (Humphrey et al., 2016).

1.3.5 Prostate cancer classification

Prostate cancer is graded histopathologically and staged by means of clinical examination (DRE findings), tissue biopsy and imaging. In 2004, the World Health Organization (WHO) produced a classification system for tumours of the urinary system and male genital organs. This grading system was revised and updated in 2016 (Humphrey et al., 2016). The aim of the classification is to aid risk stratification and management protocols.

1.3.5.1 Prostate cancer grading

Gleason grading is based on the combination of two grading scores. On examination of specimens, the first (primary) and second (secondary) Gleason score is the most common and second (secondary) most common cell type/pattern seen respectively. The International Society of Urological Pathology (ISUP) 2014 Gleason grading conference of PCa introduced an updated grading system from the ISUP 2005 (Epstein et al., 2016). A score between 1-5 is given, 1, being well differentiated, and 5, being poorly differentiated (Figure 6). The total score is then further grouped into grades.
1-5 (Table 1). Gleason scores <6 is regarded as grade 1, Gleason 9-10 as grade 5, and Gleason 7 (3+4) as grade 2, and Gleason 7 (4+3) as grade 3. The ISUP 2014, has therefore further codify the clinically highly significant differences between Gleason 7, 3+4 and 4+3 PCa (4+3 being more aggressive/high-risk).

**Figure 6.** Gleason grading schematic diagram.

Gleason score range between 1 and 5 based on how differentiated cells are under histopathological examination. Two scores are added together to give a final score. Gleason scores 7-10 are considered as high-grade disease (Humphrey et al., 2016).
<table>
<thead>
<tr>
<th>Gleason Score</th>
<th>Grade Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-6</td>
<td>1</td>
</tr>
<tr>
<td>7 (3+4)</td>
<td>2</td>
</tr>
<tr>
<td>7 (4+3)</td>
<td>3</td>
</tr>
<tr>
<td>8 (4+4, 3+5, 5+3)</td>
<td>4</td>
</tr>
<tr>
<td>9-10</td>
<td>5</td>
</tr>
</tbody>
</table>

**Table 1.** International Society of Urological Pathology 2014 grade groups.

The new ISUP grading system now comprises of Gleason scores 6-10 subcategorized into 5 grade groups. Gleason 2-6 being in group 1 and Gleason 9-10 being in group 5 (Mottet et al., 2017).

1.3.5.2 Prostate cancer staging

The widely-used tumour (T), nodal (N), metastasis (M) staging system is used for staging PCa. The T-stage describes the primary tumour and the degree of invasion of the disease locally, T1-2 being confined to the prostate capsule, and T3-4 being invasion through the capsule to structures including the seminal vesicles, rectum and pelvic wall. Tumour detected on transurethral resection of prostate (TURP) or transrectal ultrasound-guided prostate biopsy (TRUS-PBx) are T1 disease and disease that are palpable on DRE are labelled as T2 disease (Table 2). The N-stage highlights the spread of PCa in relation to the regional lymph nodes drained by the prostate gland (N1). These include nodes of the pelvis, below the bifurcation of the common iliac arteries. Pelvis MRI can assess the prostate as well as identifying any enlarged/abnormal regional lymph nodes. Finally, the M-stage describes any spread to non-regional lymph nodes (M1a non-pelvic nodes) or distant structures such as the bones (M1b) or distant organs commonly the liver, brain and lungs (M1c) (Table 2). Whole body MRI can evaluate distant organ metastasis and non-regional lymph nodes. In addition, single-photon emission CT (SPECT) is useful in assessing bony metastasis. In summary staging requires a combination of DRE, PBx and imaging in the form of computerized tomography (CT) scan, MRI and isotope bone scan.
<table>
<thead>
<tr>
<th>Tumour (T) Stage</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>TX</td>
<td>Primary tumour cannot be assessed</td>
</tr>
<tr>
<td>T0</td>
<td>No evidence of tumour</td>
</tr>
<tr>
<td>T1</td>
<td>Tumour not palpable or visible by imaging</td>
</tr>
<tr>
<td>T1a</td>
<td>Tumour in ≤5% of resected tissue</td>
</tr>
<tr>
<td>T1b</td>
<td>Tumour in &gt;5% resected tissue</td>
</tr>
<tr>
<td>T1c</td>
<td>Identified by needle biopsy</td>
</tr>
<tr>
<td>T2</td>
<td>Tumour confined within the prostate (palpable on DRE)</td>
</tr>
<tr>
<td>T2a</td>
<td>Involves one half of one lobe or less</td>
</tr>
<tr>
<td>T2b</td>
<td>Involves more than half of one lobe</td>
</tr>
<tr>
<td>T2c</td>
<td>Involves both lobes</td>
</tr>
<tr>
<td>T3</td>
<td>Tumour extends through prostatic capsule</td>
</tr>
<tr>
<td>T3a</td>
<td>Extracapsular extension (uni- or bilateral)</td>
</tr>
<tr>
<td>T3b</td>
<td>Invades seminal vesicle(s)</td>
</tr>
<tr>
<td>T4</td>
<td>Fixed or invades adjacent structures other than seminal vesicles- external sphincter, rectum, levator muscles, pelvic wall</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Node (N) Regional</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>NX</td>
<td>Regional lymph nodes cannot be assessed</td>
</tr>
<tr>
<td>N0</td>
<td>No regional lymph node metastasis</td>
</tr>
<tr>
<td>N1</td>
<td>Regional lymph node metastasis</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Metastasis (M)</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>MX</td>
<td>Distant metastasis cannot be assessed</td>
</tr>
<tr>
<td>M0</td>
<td>No distant metastasis</td>
</tr>
<tr>
<td>M1</td>
<td>Distant metastasis</td>
</tr>
<tr>
<td>M1a</td>
<td>Non-regional lymph node(s)</td>
</tr>
<tr>
<td>M1b</td>
<td>Bone(s)</td>
</tr>
<tr>
<td>M1c</td>
<td>Other sites(s)</td>
</tr>
</tbody>
</table>

**Table 2.** Tumour Node Metastasis (TNM) classification of prostate cancer.

The tumour (T) stage describes the degree of prostate invasion. The node (N) and metastasis (M) stage describe the presence of nodal or distant organ involvement. The TNM staging system was developed by the American Joint Committee on Cancer (AJCC) and the Union for International Cancer Control (UICC) (Adapted from the European Association of Urology 2017 guidelines (Mottet et al., 2017).
1.3.6 Prostate cancer screening
The systematic examination and investigation of asymptomatic men is termed population screening. The aim is to reduce overall mortality with the secondary aim of maintaining quality of life (QoL). Mortality rates from PCa vary from country to country. A recent report on reduced mortality rate seen in the USA is probably related to aggressive PCa screening (through PSA testing) and aggressive treatment (Etzioni et al., 2013). Prostate cancer screening is a controversial topic with conflicting data arising from large RCTs (Andriole et al., 2009; Schröder et al., 2009; Hugosson et al., 2010). Screening for PCa in the NHS is not practiced, as overall benefits may not be superior to the risks of over diagnosis and over-treatment. A Cochrane review published in 2013 reported that screening was associated with an increased diagnosis of PCa, more localised PCa than advanced PCa. From results of five RCTs (341,000 randomised men), no PCa-specific survival benefit was observed and from results of four RCTs, no overall survival benefit was observed (Illic et al., 2013). Within the UK, at risk individuals (age over 50 years, family history, African-Americans, raised PSA) may be offered PSA testing (Mottet et al., 2017) and individual patients can ask their GPs for a PSA test.

1.3.7 Diagnosis of prostate cancer
Definitive diagnosis of PCa is based on histopathological examination of specimens obtained from PBx, TURP or prostatectomy for benign prostate hyperplasia (BPH). Tissue sampling/diagnosis is usually proceeded by suspicion detected on DRE, PSA testing and an MRI scan.

1.3.7.1 Prostate-specific antigen
Prostate-specific antigen, also known as kallikrein-3 (KLK3), is a glycoprotein enzyme encoded in humans by the KLK3 gene. PSA is present in the blood in multiple forms known as isoforms, and some of these forms are more cancer-specific. PSA testing was introduced in the late 1980’s to aid PCa diagnosis (Stamey et al., 1987). Whilst PSA is relatively organ-specific, elevations in serum PSA levels are not cancer-specific. PSA is secreted by
the prostatic glandular epithelium into the prostatic ducts. As such, no or low levels of PSA should reach the blood stream. Any physiological or pathological process that affects the prostate/blood barrier will allow PSA to reach the blood stream. Examples of these include BPH, prostatitis and malignancy (Nadler et al., 1995). The threshold for an abnormal PSA level has been 4.0ng/ml and the associated estimated sensitivity and specificity are 21-44% and 91-92% respectively. Using a cut-off of 3.0 ng/ml increases the sensitivity and decreases the specificity to 32% and 85% respectively (Wolf et al.) (Holmström et al., 2009). The positive predictive value for a PSA level >4.0 ng/ml is ~30% (Catalona et al., 1994), and the negative predictive value for a PSA value <4.0 ng/ml estimated by The Prostate Cancer Prevention Trial (PCPT) is 85% (Thompson et al., 2004). Moreover, many men may have PCa despite having a low PSA level (Dong et al., 2008).

Since there is doubt in the diagnostic accuracy of PSA levels, patients with suspected PCa undergo a PBx.

1.3.7.2 Prostate-specific antigen density and kinetics

To improve the specificity for prostate cancer, various modifications of the PSA assay have been suggested.

1) PSA density

The PSA density is the PSA serum level divided by the prostate volume determined by TRUS-PBx. The higher the density, the more likely it is that the cancer is clinically significant. Lower densities are seen in men with large glands through BPH.

2) PSA velocity and doubling time

PSA velocity (PSAV) is the increase in PSA over time (ng/ml/year) and PSA doubling time (PSA-DT) is the exponential increase in PSA measured against previous values. Both may have a role in prognosis in treated PCa, but limited diagnostic use (Mottet et al., 2017).
3) Free/total PSA ratio
Free/total (f/t) PSA ratio may be used to differentiate BPH from PCa. It stratifies the PCa risk with 4-10ng/ml total PSA and negative DRE. In a reported study, 56% of men with f/t PSA <0.10 were found to have PCa on PBx, in contrast only 8% of men with f/t PSA >0.25ng/ml were found to have PCa (Catalona et al., 1998).

4) Other forms of PSA testing, PHI, 4K and IsoPSA
A few tests measuring a range of KLK in serum and plasma are now commercially available, including the FDA-approved Prostate Health Index (PHI) test and the four kallikrein score (4K). Both the PHI and the 4K score are used with the intention to reduce the number of unnecessary PBx (Loeb et al., 2014; Bryant et al., 2015).

The PHI test combines free and total PSA and the (-2)pro-PSA isoform (p2PSA) and is calculated using the following formula, ((−2)pro-PSA/fPSA) × PSA1/2. Isoforms of PSA have similar biological roles but differ in sequence and structural form. PHI was approved by the FDA in 2012 and studies have shown that PHI can improve the detection of high-risk PCa and is associated with PCa aggressiveness (Lazzeri et al., 2013).

The 4K score measures free, intact and total PSA and kallikrein-like peptidase 2 (hK2). In addition, the 4K score also takes into account clinical information such as age and history of previous PBx results. A meta-analysis showed that the 4K score is associated with an improvement of 8-10% in predicting biopsy-confirmed PCa with an estimate of avoiding 48-56% of current PBx (Voigt et al., 2014). However, unlike PHI, the 4K score is not yet FDA-approved, and currently more prospective data is needed to compare PHI and the 4K score.

The PHI and 4K score measure only a few known isoforms of PSA that give meaningful information if they are present at a given time. IsoPSA is a structure-based (rather than concentration-based) test that incorporates the entire spectrum of PSA structural changes. IsoPSA has been shown in a
prospective study to be more superior than standard PSA at predicting PCa (versus benign) and high-grade PCa (versus low-grade and benign). Once validated, IsoPSA may be another potential assay that would help to better select at risk men for PBx (Klein et al., 2017).

1.3.7.3 Prostate Cancer Gene 3
To aid the detection of PCa, several biomarker screens have been performed using normal and malignant prostate samples. On such screen identified the Prostate cancer gene 3 PCA3 non-coding RNA (Bussemakers et al., 1999).

PCA3 ncRNA has been shown to be useful when used in conjunction with PSA (PCA3 score, PCA3/PSA mRNA ratio x 1000) to predict PCa on repeat PBx (rPBx). The PCA3 test was approved by the FDA in 2012. PCA3 is a long ncRNA that was discovered in 1999 (Bussemakers et al., 1999) and has been found to be overexpressed in >95% (53 out of 56 radical prostatectomy specimens) of primary PCa tissue. The PCA3 gene is located on chromosome 9q21-22 and was originally described as being formed by 4 exons (1, 2c, 3, 4) with alternative polyadenylation at 3 different positions in exon 4. Subsequently, 4 new transcription start sites (Exon 1), 2 new differentially spliced exons (2a and 2b), and 4 new polyadenylation sites in exon 4 have been identified (Clarke et al., 2009). The presence of large number of stop codons in all 3 reading frames and the lack of an extensive open reading frame (ORF) suggest that PCA3 is a ncRNA. The PCA3 gene is found to be embedded in intron 6 of the BCH motif-containing molecule at the carboxyl terminal region 1 gene (BMCC1, also known as PRUNE2) (Figure 7), but transcribed in the antisense orientation (Bussemakers et al., 1999; Clarke et al., 2009).
PCA3 (Blue) is embedded in intron 6 of BMCC1 (Grey) in chromosome 9 and is transcribed in the antisense direction (Red arrow). PCA3 has 6 exons and 4 transcription start sites. The ncRNA is prostate cancer specific and is overexpressed in >95% of primary prostate cancer tissue (Bussemakers et al., 1999; Clarke et al., 2009).

A quantitative reverse transcription polymerase chain reaction (qRT-PCR) analysis of PCA3 showed low levels of expression in normal prostate and BPH tissue, no expression in other tissues, and a median upregulation of PCA3 in PCa cells by ~66 folds relative to non-malignant tissue (de Kok et al., 2002; Hessels et al., 2003). Unlike PSA, this test is organ- and PCa-specific (Crawford et al.; Nadler et al., 1995), it can be detected in urinary samples (PROGENSA assays), more readily after DRE (Groskopf et al., 2006), and has a sensitivity of 67% and a specificity of 83% (Hessels et al., 2003). A European prospective multicentre study (Haese et al., 2008) and the REDUCE RCT (Aubin et al., 2010) demonstrated that the PCA3 score is higher in men with a positive biopsy than in men with a negative biopsy. A PCA3-Score threshold of 35 provides an optimal balance between sensitivity (47–58%) and specificity (72%) for detecting PCa. A score of 35 would avoid 67% of PBx, missing 21% of high-grade tumours. However, lowering the score to 20 would miss only 9% of high-grade tumours, but at the expense of avoiding less PBx (44%) (Haese et al., 2008). Overall, PCA3 is a valuable biomarker used in conjunction with PSA to guide the need for rPBx in the detection of PCa (Tinzl et al., 2004; van Gils et al., 2007; Roobol et al., 2010).

1.3.7.3.1 A function of PCA3

The function of PCA3 is not entirely clear. However, PCGEM1 (Prostate-specific transcript 1, chromosome 2q32) is also a ncRNA and like PCA3, it is
prostate-specific and is overexpressed in PCa (Srikantan et al., 2000). PCGEM1 has been found to promote cell proliferation and inhibit apoptosis (Petrovics et al., 2004; Fu et al., 2006), these functional roles may also apply to PCA3. BMCC1/PRUNE2 has also been found to be overexpressed in PCa tissues and thought to have a role in cellular proliferation, apoptosis, and cellular transformation (Clarke et al., 2009). Knowing that the PCA3 gene is embedded in the BMCC1 gene, it may have roles similar to BMCC1. In addition, it has been shown that PCA3 has a dominant-negative oncogenic role in regulating tumour suppressor gene BMCC1, through RNA editing mediated by a complex containing adenosine deaminase acting on RNA (ADAR) family members (Salameh et al., 2015).

Although PCA3 is overexpressed in PCa tissue, its role in predicting the presence of PCa on rPBx in men with elevated PSA level and a previous negative biopsy is not entirely clear. Although studies have shown that PCA3 scores can predict the probability of a positive rPBx result, the optimum cut-off value is unknown (Haese et al., 2008; Aubin et al., 2010; Kirby et al., 2012).

1.3.7.3.2 Clinical barriers to use of the PCA3 assay
Although PCA3 helps in the diagnosis of PCa, there are logistic and implementation problems associated with this test. Firstly, the target for detection is a long RNA (gene size is 25kb, PCR target assay size is 380bp) (Bussemakers et al., 1999). These molecules are unstable and prone to digestion by endogenous RNases. As such, handling the sample prior to the assay requires laborious stringency, which limits its adoption and ensures a high cost. Secondly, the function of PCA3 is unclear, producing a biological gap in knowledge that hampers the scientific community’s acceptance of this assay. Finally, the test has a lower sensitivity than PSA for PCa, including potentially aggressive disease (Roobol et al., 2010). Although, PCA3 is able to detect both low- and high-grade disease, like PSA, its sensitivity and specificity vary according to different cut-off scores. In addition, the relationship between PCA3-scores and parameters (i.e. tumour volume and
Gleason Score) of PCa aggressiveness is unclear. Although some studies have established a positive correlation between PCA3-scores and more serious disease (>T2, Gleason >6), others have not (Roobol et al., 2010).

Little is known about the function of many long ncRNAs, and until recently their importance biologically was unclear. The recent Gencode human genome sequence assembly identified 10,000 long ncRNAs with features similar to PCA3 (Derrien et al., 2012). Many are located within or adjacent to other RNAs (PCA3 is located within BMCC1) (Clarke et al., 2009) and many are processed to smaller more active species (Röther et al., 2011).

1.3.7.4 Other biomarkers

Biomarkers are measurable markers detected in urine, blood/plasma or tissue specimens, and function to help diagnose disease, predict progression/prognosis and to monitor for recurrence following treatment. The evolution of PSA led to earlier diagnosis and treatment of PCa, however, as discussed, PSA is often associated with false-negative results. This has focused research on markers to detect PCa, and to differentiate indolent disease from aggressive disease. RNA profiling using microarray-based techniques is used to trace changes in gene expression during oncogenesis.

Two current promising RNA biomarkers are the PCA3 ncRNA and transmembrane protease, serine 2 (TMPRSS2) fusion transcripts.

In men with an initial negative PBx who have a persistent risk, such as elevated PSA or abnormal DRE, there are additional tests available to aid rPBx decisions. The urinary PCA3 test discussed earlier is one additional test that can be used. Other tests that are available are the serum 4K, PHI and IsoPSA tests (Chapter 1.3.7.2), or a tissue-based epigenetic test (ConfirmMDx). The ConfirmMDx test quantifies the methylation level of promoter regions of three genes (RASSF1, GSTP1 and APC) in benign prostatic tissue. A multicentre study showed a negative predictive value of 88% when methylation was absent in all three markers, suggesting that rPBx could be avoided in these men (Stewart et al., 2013; Partin et al., 2014).
Given the limited currently available data, no recommendation can be made regarding its routine application.

Gene rearrangements have been described in multiple cancers. The TMPRSS2-ERG fusion gene comprises the androgen-responsive genes TMPRSS2 and erythroblast transformation-specific (ETS)-related genes (ERG), and was observed in ~40-80% of PCa in 2005 (Tomlins et al., 2005). Both genes are located on chromosome 21, and the TMPRSS2-ERG score is calculated using the formula, (TMPRSS2-ERG mRNA/PSA RNA copies) x 100,000. A recent meta-analysis showed that TMPRSS2-ERG overexpression is associated with tumour stage, but not with disease recurrence or mortality in men treated with radical prostatectomy (Pettersson et al., 2012).

The Mi-Prostate Score (MiPS) combines the prognostic significance of urine TMPRSS2-ERG and urine PCA3 with serum PSA to generate a PCa risk assessment score. Although not yet FDA-approved, the MiPS has been shown to be more superior to PSA alone in predicting biopsy-confirmed PCa and high-grade disease (Tomlins et al., 2016).

Biological markers, including urine PCA3, transmembrane protease, serine 2- ETS-related gene (TMPRSS2-ERG) fusion, or PSA isoforms such as the PHI index appear promising as does genomics on the tissue sample itself (Jerónimo et al., 2011; Berg et al., 2014; Cantiello et al., 2016). However, further study data will be needed before such markers can be used in standard clinical practice.

1.3.7.5 Prostate biopsy
A PBx is considered if the PSA level is raised and/or abnormal findings on DRE is detected. Limited PSA elevation alone does not necessitate immediate PBx, and warrants repeating. This is because as discussed before, PSA is not cancer-specific and may be raised under numerous
conditions. There are two main routes for biopsy—either through the rectum or through the perineum.

1.3.7.5.1 Transrectal Ultrasound-guided biopsy

A transrectal ultrasound-guided prostate biopsy (TRUS-PBx) is currently a standard biopsy modality of choice (NICE, 2014b; Mottet et al., 2017). It is performed under local anaesthetics with prophylactic antibiotics cover. The ultrasound probe is inserted into the rectum and 6 or more cores of tissue are taken with a needle from each lobe of the prostate.

Over 100,000 TRUS-PBxs are performed annually in the UK and approximately 70% of UK men are found not to have PCa (Lane et al., 2010). The low detection rate maybe due to small tumours, tumours located peripherally or anteriorly that are difficult to sample, or no tumours in the first place (PBx following a raised PSA that is not cancer related). Men with an initial negative biopsy and persistently raised PSA are subject to rPBx. The PCa detection rates on rPBx is around 10 to 30% (Keetch et al., 1994; Djavan et al., 2001). Unnecessary PBx increases healthcare costs, patients’ anxiety and puts patients at risk of PBx complications including infection and bleeding (Rosario et al., 2012; Loeb et al., 2013b). The percentage of men hospitalized for complications after biopsy is 4.1%, and this figure seems to be increasing (Nam et al., 2010). In addition, there is a risk of over-diagnosing insignificant tumours and delaying the detection of significant disease. Each rPBx could cost ~£310 if they are transrectal or £650 if they are transperineal (NICE, 2014a). European guidelines suggest that one set of rPBx is warranted where there is an abnormal DRE, persistently elevated PSA value and a histopathological finding suggestive of malignancy (HGPIN) on initial biopsy. There are no further recommendations on subsequent repeat biopsies (Mottet et al., 2017). The decision as to how best to proceed is based predominantly on retrospective international data and the man’s original experience of biopsy (Wade et al., 2013). Although there are new biomarkers found such as PCA3 to help better select men for initial and rPBx,
these are not used globally secondary to its high cost, and its unclear biological role in PCa (Kirby et al., 2012).

1.3.7.5.2 Transperineal biopsy
Other approaches include the transperineal approach, which is performed under general anaesthetics and requires accessing the prostate gland through the perineum. Cancer detection rates are comparable to TRUS-PBx (Takenaka et al., 2007).

Saturation biopsy involves taken >20 cores and the PCa detection rate is between 30 and 43% (Walz et al., 2006). The increased number of cores taken is to maximise detection rate on rPBx following a negative initial PBx. Saturation biopsy can be performed with the transperineal approach (Transperineal template biopsy, TPM-Bx), which detects an additional 38% of cancer, however, there is a ~10% risk of urinary retention (Moran et al., 2006).

Multiparametric MRI (mpMRI) are increasingly performed prior to PBx to improve the detection of clinically significant PCa and to allow MRI-targeted biopsy in case of positive mpMRI (Schoots et al., 2015). MRI-targeted PBx can be performed through USS/mpMRI fusion software. However, there are contradictory data as to whether there is a difference in PCa detection rate between MRI-targeted and systematic PBx and systematic PBx alone (Panebianco et al., 2015; Tonttila et al., 2016).

1.3.7.5.3 Prostate biopsy complications
Prostate biopsy complications are listed in Table 3. Although antibiotic cover reduces the risk of severe infection, this is on the incline as a result of antibiotic resistance (Loeb et al., 2013b).
### Table 3. Complication rates per biopsy session.

Adapted from the European Association of Urology 2016 guidelines. Complication rates are reported irrespective of the number of cores taken. The most common complication is haematospermia (37.4%) and haematuria lasting over 1 day (14.5%). Other complication rates are ≤1% (Mottet et al., 2017).

<table>
<thead>
<tr>
<th>Complications</th>
<th>Percentage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Haematospermia</td>
<td>37.4</td>
</tr>
<tr>
<td>Haematuria &gt;1day</td>
<td>14.5</td>
</tr>
<tr>
<td>Rectal bleeding &lt;2days</td>
<td>2.2</td>
</tr>
<tr>
<td>Prostatitis</td>
<td>1.0</td>
</tr>
<tr>
<td>Fever &gt;38.5°C</td>
<td>0.8</td>
</tr>
<tr>
<td>Epididymitis</td>
<td>0.7</td>
</tr>
<tr>
<td>Rectal bleeding &gt;2days</td>
<td>0.7</td>
</tr>
<tr>
<td>Urinary retention</td>
<td>0.2</td>
</tr>
<tr>
<td>Other complications requiring hospitalization</td>
<td>0.3</td>
</tr>
</tbody>
</table>

1.3.7.6 Imaging

For staging (Section 1.3.5.2), following tissue diagnosis of PCa, cross-sectional abdomino-pelvic imaging using MRI and a bone scan is used to detect distant organ/lymph node and bone involvement respectively.

Multiparametric MRI of the pelvis is a dynamic contrast-enhanced MRI technique for local detection/staging of PCa. Correlation with radical prostatectomy shows that mpMRI has sensitivity rates of 80% for detecting Gleason ≥7 tumours of <0.5ml volume and 100% for detecting >2ml tumours (Turkbey et al., 2011; Bratan et al., 2013).

1.3.8 Management of prostate cancer

The treatment of PCa depends on the patient’s fitness and performance status (including co-morbidities) and their cancer (stage, grade and risk of the disease) (Table 4). Localised disease can be managed by watchful waiting, active monitoring/surveillance, radical radiotherapy or radical surgery (NICE, 2014b; Mottet et al., 2017). The optimal treatment option for localised
disease is not known and the UK Prostate Testing for Cancer and Treatment (ProtecT) RCT aimed to evaluate this (Lane et al., 2010, 2014). Between 1999 and 2009, ProtecT recruited 82,429 men for PSA testing within nine UK cities. A total of 2664 men received a diagnosis of localised PCa and 1643 were randomised to active monitoring (n=545), radical prostatectomy (n=553) or radical radiotherapy (n=545). At a median of 10 years follow-up, disease-specific mortality was not significantly different amongst treatments, however, there were higher rates of disease progression in the active monitoring group (Hamdy et al., 2016). Metastatic disease is managed with hormonal therapy or chemotherapy. However, a proportion of men who receive hormonal therapy become resistant to the treatment and deteriorate rapidly. The cancer progresses in an androgen-independent (Castration-resistant PCa, CRPC) manner and the mechanisms underlying this are not entirely clear. The treatment options for localised and locally advanced/metastatic PCa are summarized in Figure 8.

<table>
<thead>
<tr>
<th>Low-risk</th>
<th>Intermediate-risk</th>
<th>High-risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>PSA &lt;10ng/ml and Gleason &lt;7 and cT1-2a</td>
<td>PSA 10-20ng/ml or Gleason 7 or cT2b</td>
<td>PSA &gt;20ng/ml or Gleason &gt;7 or cT2c</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Any PSA, any Gleason, but cT3-4 or cN+</td>
</tr>
</tbody>
</table>

| Localised | Locally advanced |

**Table 4.** European Association of Urology risk groups for localised and locally advanced prostate cancer.

Risk stratification (low, intermediate and high) is based on PSA, Gleason score and staging (imaging). Treatment decisions are based on PCa risk, patient comorbidities and patient’s wishes (Mottet et al., 2017).
Figure 8. Summary of the treatment options for prostate cancer.

Treatment depends on the risk of disease and whether the disease is localised, locally advanced or metastatic. Options include watchful waiting, active surveillance, radical surgery or radiotherapy and hormonal therapy.

WW, watchful waiting; AS, active surveillance; RP, radical prostatectomy; RT, radiotherapy; HIFU, high-intensity focused ultrasound; ADT, androgen deprivation therapy; MAB, maximal androgen blockade (for castration-resistant disease)
1.3.8.1 Localised prostate cancer
Localised PCa is disease confined to the prostate, tumour stage T1-2. Insignificant or indolent disease may not need immediate radical treatment. Epstein defined indolent PCa as, Gleason score <6; <3cores positive; <50% positive per core (Epstein et al., 1994).

1.3.8.1.1 Watchful waiting
Many men with incidental finding or screened-detected PCa will not necessarily need definitive treatment. These men could be managed conservatively to reduce over-treatment of insignificant low-risk disease. Watchful waiting (WW) is also termed deferred or symptom-guided treatment, which refers to monitoring until the patient develops symptoms. This approach is more of a palliative approach that is preferred in men with a life expectancy of less than 10 years. The aim is to minimize treatment-related adverse effects (Albertsen, 2015). Follow-up is patient-specific and assessment is not pre-defined. Many small, localised and low-grade disease do not progress, and radical treatment may result in over-treatment. This was confirmed by a recent analysis in 19,659 men with 10 years follow-up. Men with low co-morbidity index scores had a low-risk of death at 10 years or died from competing causes. In addition, men with higher co-morbidity scores, tumour aggressiveness had little impact on overall survival (Albertsen et al., 2011).

1.3.8.1.2 Active surveillance
For those who do not need or want immediate treatment, active surveillance (AS) may be offered. Treatment is deferred until there are clinical features of disease progression. Follow-up is pre-defined and assessments include DRE, PSA, rPBx and mpMRI. The aim is to detect progressing disease that would initiate prompt curative treatment. The AS approach is preferred in men with life expectancy of over 10 years (Thomsen et al., 2014; Welty et al., 2014). The current selection criteria for AS is based on two systematic reviews and include, Gleason 6, <2-3 positive cores on PBx with <50% cancer involvement or each core, a clinical T1c or T2a, PSA <10ng/ml (Thomsen et
al., 2014; Loeb et al., 2015). The initiation of active treatment should be based on a change in PBx results (Gleason score, number of positive cores), T-stage progression, or upon a patient’s request.

1.3.8.1.3 Radical prostatectomy
Radical prostatectomy (RP) should be offered to a patient with low- and intermediate-risk PCa with at least 10 years of life expectancy. The procedure involves eradication of disease by removal of the entire prostate, both seminal vesicles and pelvic lymph nodes. Due to the abundant nerve supply around the prostate, there is a risk of incontinence and impotency. The prostate could be removed via the traditional open approach (retropubic prostatectomy, RRP), laparoscopic prostatectomy (LP) or robotic-assisted LP (RALP). Prostatectomy was traditionally indicated in organ-confined disease (T1-T2), however, in recent years, there has been an interest in performing RP for locally advanced T3 disease (Ward et al., 2005; Hsu et al., 2007). Minimally-invasive approaches may have the benefit of better operating view for the surgeon, reduced peri-operative morbidity and hospital stay, however, the oncological outcomes and survival may not differ significantly (Novara et al., 2012; Ramsay et al., 2012).

1.3.8.1.4 Radiotherapy
External beam radiotherapy (EBRT) can be used as primary or adjuvant treatment, and can be used in conjunction with hormonal therapy to improve outcomes (D’Amico et al., 2008; Bolla et al., 2010). EBRT can be offered to all risk groups of non-metastatic PCa. Combination therapy with hormonal therapy is recommended short-term in intermediate-risk PCa and long-term in high-risk localised disease. Men with low-risk disease, without a previous TURP and with a prostate volume <50ml may be offered brachytherapy as a monotherapy. Low-dose-rate (LDR) brachytherapy uses radioactive seeds permanently implanted into the prostate (Davis et al., 2012). Radiotherapy complications include gastrointestinal and genitourinary toxicity (Zelefsky et al., 2008), erectile dysfunction and an increased risk of being diagnosed with secondary malignancies (Abdel-Wahab et al., 2008). However, radiotherapy
has a lesser effect on erectile function compared with surgery (Robinson et al., 2002). The risk of developing rectal cancer and bladder cancer following radiotherapy can be a 1.7-fold (Baxter et al., 2005) and 2.34-fold (Liauw et al., 2006) increase retrospectively compared to surgery.

1.3.8.1.5 Options other than surgery or radiotherapy

Other modalities for managing localised PCa include cryosurgery (CSAP), high-intensity focused US (HIFU), photodynamic therapy, radiofrequency ablation and electroporation. A lot of these treatment modalities are in the early phases of evaluation, and there are sufficient data only on CSAP and HIFU.

Cryosurgery involves freezing (-40°C) of the prostate by the placement of a cryoneedle under TRUS guidance. Men who are eligible for CSAP are ones who have a life expectancy of >10 years, prostate gland <40ml, PSA <20ng/ml and Gleason <7 (Rees et al., 2004; Ramsay et al., 2015).

High-intensity focused ultrasound uses focused US waves to induce tissue damage by thermal (65°C) effects. This procedure is performed under general or spinal anaesthesia and is offered to patients with low- to intermediate-risk disease. In a recent systematic review comparing CSAP (n=3995) and HIFU (n=4000) with AS, RP and EBRT, there was no evidence that mortality at 4 years, or other cancer-specific outcomes differed between treatments (Ramsay et al., 2015).

Focal therapy refers to treatment of low-volume (disease occupying 5-10% of the prostate volume) unifocal or unilateral disease, usually in the form of ablative therapy such as CSAP or HIFU. The main aim of focal therapy is to ablate tumours selectively and limiting toxicity by sparing the neurovascular bundles, sphincter and urethra (Eggener et al., 2007). Although in the same systematic review published by Ramsay et al, there were no significant differences in oncological outcome at 3 years amongst a subgroup of focal therapy against RP and EBRT, focal therapy remains investigational and
should not be offered as a therapeutic alternative outside clinical trials (Ramsay et al., 2015).

A recent RCT comparing AS (n=207) and TOOKAD® Soluble-Vascular Photodynamic Therapy (VTP, n=206) showed that at a median (IQR) follow-up of 24 (24-25) months the disease progression rate was lower in the VTP group (VTP, 28% and AS, 58%, p<0.0001). Therefore men with localised, low-risk PCa can be managed with tissue-preserving focal therapy through VTP with good disease-free progression outcomes (compared with AS) and reduced need for whole-gland radical treatment (Azzouzi et al., 2017).

1.3.8.2 Locally advanced and metastatic prostate cancer

1.3.8.2.1 Androgen deprivation therapy

The pathogenesis of PCa is associated with androgens, hence targeting the hormonal pathway suppresses disease progression. Androgen deprivation can be achieved by either inhibiting the action on androgens at the receptors, or suppressing the secretion of testicular androgens.

Androgen deprivation therapy (ADT) is achieved surgically by bilateral orchidectomy or pharmacologically with anti-androgens, LHRH antagonists or agonists. Surgery results in castration (testosterone <20ng/dl) within 12 hours, and is a straight-forward procedure that can be performed under local anaesthetics (Desmond et al., 1988). Drugs that suppress the secretion of testicular androgens include Luteinising-hormone-releasing hormone (LHRH) agonists (Leuprolide, Goserelin) and LHRH antagonists (Abarelix, Degarelix), and drugs that act on androgen receptors (AR) include anti-androgens (steroidal- Cyproterone acetate; non-steroidal- Flutamide, Bicalutamide). LHRH agonists are delivered as depot injections on a 1-, 2-, 3- or 6-monthly basis, and castration is usually obtained within 2-4 weeks. In contrast LHRH antagonists are given on a 1-monthly basis and castration is achieved within days (Pagliarulo et al., 2012). In the first instance of managing locally advanced disease, either LHRH agonists/antagonists or anti-androgens are used as monotherapy. A systematic review of the side effects of ADT has
recently been published and common/severe effects include erectile dysfunction, hot flushes, non-metastatic bone fractures, metabolic effects (lipids, glucose), cardiovascular morbidity and fatigue (Ahmadi et al., 2013).

1.3.8.2.2 Castration-resistant prostate cancer
During castration, the occurrence of castration-resistant (CRPC) is systematic. When CRPC develops (rising PSA or radiological progression despite testosterone <50ng/ml), maximal androgen blockade (MAB) is achieved by combining LHRH agonist/antagonist and anti-androgens together (Pagliarulo et al., 2012). Newer medications that manage CRPC include Abiraterone acetate (CYP17 inhibitor- suppresses testosterone synthesis), and Enzalutamide (a novel anti-androgen).

1.3.8.2.3 Metastatic prostate cancer
Primary ADT is the standard of care for patients with metastatic PCa. In metastatic CRPC (mCRPC), MAB should be initiated (Pagliarulo et al., 2012). There is evidence from RCTs that ADT combined with chemotherapy (Docetaxel) provides a more superior overall survival rate. The STAMPEDE (Systemic Therapy in Advancing or Metastatic Prostate Cancer: Evaluation of Drug Efficacy) RCT randomised 2962 men to four groups of treatment and showed that the median overall survival in the ADT alone and ADT + Docetaxel group to be 71 and 81 months respectively (James et al., 2016). Sweeney et al who randomised 790 men also showed a significant higher survival rate in the combination group (57.6 months versus 44 months) compared to ADT alone (Sweeney et al., 2015). There is evolving evidence that ADT combined with radical prostatectomy or radiotherapy provides an overall survival benefit in newly diagnosed metastatic PCa (Culp et al., 2014; Gratzke et al., 2014).

1.4 Prostate cancer and androgen regulation
Androgens play a crucial role in male sexual development and prostate physiology. Testosterone is produced by testicular Leydig cells, and dihydrotestosterone (DHT) is produced from testosterone in peripheral
tissues by 5-alpha reductase (Michaud et al., 2015). Testosterone is bound to sex hormone-binding globulin (SHBG), however, it is the unbound, free testosterone that is the active form (Figure 9). It is known that androgens have an important role in the pathogenesis and progression of PCa (androgen hypothesis). This was first brought to attention when Huggins and Hodges observed the benefits of castration in patients with PCa (Huggins et al., 1941). Several cell lines studies have shown that PCa cells grow in an androgen-dependent manner. Hormonal therapy remains the first line treatment for advanced or metastatic disease (discussed earlier 1.3.8.2), which targets the hypothalamus-pituitary-adrenal/testicular-prostate axis (Heidenreich et al., 2008). However, ADT is not curative and in some men, PCa reactivates AR-signalling and resumes proliferation despite low levels (<20ng/dl) of testosterone (CRPC). AR-signalling is maintained through multiple mechanism including AR mutation, AR amplication, altered co-regulator profiles, and extra-gonadal androgen production (Karantanos et al., 2015).
Figure 9. The hypothalamus-pituitary axis.

The hypothalamus releases luteinizing hormone-releasing hormone (LHRH) to trigger the pituitary gland to release luteinizing hormone (LH) and adrenocorticotropic hormone (ACTH). LH facilitates the production of testosterone in testicular Leydig cells. Testosterone is converted into dihydrotestosterone (DHT) via 5-alpha reductase. ACTH acts on adrenal glands to stimulate the release of adrenal androgens (Dehydroepiandrosterone, DHEA), and together with DHT, prostate growth is enhanced.

1.4.1 Paracrine/autocrine synthesis

In normal homeostasis, testosterone and DHT are produced to provide survival and proliferation of prostatic epithelium by a paracrine loop. During PCa progression, an autocrine loop is established and PCa cells produce numerous factors/androgens to support their own growth (Logothetis et al., 2013). The levels of DHT and key enzymes in androgen synthesis, such as CYP17A1 (Cytochrome P450 hydroxylase type enzyme) and HSD3B1 are higher in CRPC. This suggests increased androgen synthesis in advanced PCa, partly due to upregulation of enzymes through autocrine androgen synthesis (Chang et al., 2011). There are multiple steps in the synthesis of adrenal and testicular androgens, and some hormonal treatments are based on inhibiting androgen synthesis through inhibiting enzymes. An example is the inhibition of CYP17A1 by Abiraterone acetate (Figure 10).
Figure 10. Androgen biosynthesis.

The biochemical pathways in androgen synthesis is summarized in this figure. Current and new medications aim to block receptors or inhibit key enzymes involved in androgen synthesis, such as the inhibition of CYP17A1 by Abiraterone acetate (Karantanos et al., 2015).

1.4.2 Androgen receptor mutation
The AR gene is located on the X chromosome at Xq11-Xq12. Androgen receptor mutations occur in >10% of patients with CRPC. The mutant broadens the ligand-binding specificity of AR, sensitizing it to other steroid hormones such as progesterone and oestrogens (Grasso et al., 2012a). H874Y mutation enhances the binding of AR co-regulators and increases AR transcriptional activity through AR protein conformational change. In addition, mutation of W435L promotes androgen-independent AR activation (Schröder, 2008). Results suggest that AR mutations provide survival benefits to PCa cells and promote resistance to anti-androgens (Karantanos et al., 2015).
Enzalutamide is a novel antagonist of AR, inhibiting nuclear translocation, chromatin binding and interactions with AR co-regulators (Tran et al., 2009).

1.5 Genetic mechanisms of prostate cancer
Prostate cancer is a heterogeneous disease, although it’s slow growing, a proportion of men will have aggressive disease with metastasis and die from the disease. The majority of men will have indolent (Epstein criteria) disease that will not necessarily progress, and over-treatment of such low-risk disease leads to greater risks of morbidity (Epstein et al., 1994). An appreciation of the molecular basis of PCa allows us to understand the varying behaviour of the disease.

Normal cells divide, turnover and produce proteins in a controlled manner. When these processes become unregulated, oncogenesis occurs (Hanahan et al., 2011). Regulations may occur at the genetic or epigenetic level. Genetic processes are irreversible and include DNA mutation or chromosomal translocation, deletion and amplification (Croce, 2009). Epigenetic changes do not alter gene sequence or chromosomal structure and are commonly reversible (Figure 11).
Figure 11. A summary of the events leading to oncogenesis.

The development of tumours is the result of genetic and/or epigenetic events. Genetic and epigenetic alterations/mutations lead to uncontrolled cell proliferation, avoidance of apoptosis, increased angiogenesis and tissue invasion.
Different types of genetic variations can affect tumorigenesis, including germline variations (Choudhury et al., 2012) and somatic alterations. Germline changes are transmittable to offspring and are present in every cell, in comparison, somatic alterations arise in prostate cells. Genetic alterations causing amplification, deletion or translocation of segment of chromosomes lead to an imbalance of amplification of oncogenes and deletion of tumour suppressor genes. Together with gene fusions, these alterations promote oncogenesis. There is a diverse list of genes contributing to key pathways in oncogenesis across a wide-spectrum of cancers as well as PCa alone (Barbieri et al., 2013, 2014; Boström et al., 2015; Mitchell et al., 2015).

1.5.1 Cancer pathways
1.5.1.1 Phosphoinositide 3-kinase pathway
The phosphoinositide-3-kinase (PI3K) pathway is a commonly altered signalling pathway in human cancer and is altered in around 25-70% of PCa. The alteration of this pathway contributes to cell proliferation, survival and invasion. The phosphatase and tensin homolog (PTEN) gene (chromosome 10q23) is a frequently mutated TS gene and acts to deactivate PI2K-dependent signalling (Barbieri et al., 2013; Boström et al., 2015). Deletion of PTEN occurs in 40% of PCa and studies support the role of PTEN as an important tumour suppressor in PCa (Carver et al., 2009). Deletion of PTEN is associated with higher grade disease, progression, metastasis and higher risk of recurrence in treated disease (McMenamin et al., 1999), (Krohn et al., 2012), (Choucair et al., 2012).

1.5.1.2 Ras/Raf/mitogen-activated protein kinase pathway
The mitogen-activated protein kinase (MAPK) pathway has a less established role in PCa compared to other common cancers such as lung, ovarian and gastrointestinal cancer. However, activation of MAPK signalling by Ras and Raf intermediates may enhance transcriptional activity of AR and appears to be more enriched in metastatic PCa (Bakin et al., 2003). Other rare fusion genes that are associated with the MAPK pathway in PCa include KRAS, RAF1 and BRAF (Palanisamy et al., 2010).
1.5.1.3 Tumour Protein 53
Tumour protein p53 acts as a transcription factor activating the transcription of genes involved in cell cycle arrest, DNA repair and apoptosis. It is the most commonly mutated gene in cancer and deletions at the p53 locus are seen in around 25-40% of PCa (Barbieri et al., 2012) (Taylor et al., 2010). Interestingly, TP53 defects may be an early onset in PCa, as ~25-30% of clinically localised PCa harbour lesions in TP53.

1.5.1.4 Retinoblastoma protein
Tumour suppressor Retinoblastoma protein (Rb) regulates cell cycling and is deleted or mutated in a number of human cancers. The Rb1 gene (chromosome 13q14) is commonly inactivated in CRPC (~45%), and has been shown to modulate AR-signalling and inhibit progression to castration resistance (Aparicio et al., 2011).

1.5.1.5 v-myc myelocytomatosis viral oncogene homolog
The v-myc myelocytomatosis viral oncogene homolog (MYC) gene located at 8q24 encodes c-Myc, which is a transcription factor involved in cell cycle progression, cell survival and oncogenesis. It is a common oncogene in human cancers and is commonly amplified in PCa. The amplification usually involves the entire chromosome 8 arm, which may result in amplifying other genes within this region (Barbieri et al., 2012; Grasso et al., 2012b).

1.5.2 Prostate cancer specific pathways
There are generic pathways that are common to a range of human cancers, in addition, there are genetic mechanisms that are highly PCa specific.

1.5.2.1 Androgen-signalling
Focusing on PCa, the discovery of disease regression through suppressing androgens (castration) in men with advanced PCa emphasized the importance of the androgen-signalling pathway in PCa. Androgen receptor (AR) is a ligand-dependent nuclear transcription factor. The presence of amplification and mutation of AR in treated metastatic disease and the
absence of these AR lesions in primary localised PCa imply that the AR gene
do not have a role in the pathogenesis of PCa, but have a role in androgen
resistance during treatment (Taylor et al., 2010; Barbieri et al., 2012). PCa
that continue to progress despite castration become androgen-independent
and show features of active AR-signalling (Waltering et al., 2012).

There is also a list of genes encoding proteins that interact and modulate AR
activity such as transcriptional coactivators (the nuclear receptor coactivator
2 gene (NCOA2), E1A binding protein p300 gene (EP300), nuclear receptor
interacting protein 1 gene (NRIP1)), transcriptional corepressors (nuclear
receptor corepressors 2 gene (NCOR2)) and the forkhead-box family of
transcription factors (FOXA1). (Taylor et al., 2010; Barbieri et al., 2012;
Grasso et al., 2012).

Interactions between AR-signalling and other oncogenic pathways have
been established. For example, PI3K/Akt signalling pathway inhibits AR-
signalling and by negative feedback, AR inhibition activates Akt signalling.
This may be a key component of CRPC progression (Carver et al., 2011).

1.5.2.2 E26 transformation-specific (ETS) gene fusions
The presence of gene fusions between androgen-regulated genes and
members of the ETS family of transcription factors has an important role in
prostate oncogenesis. The most common fusion is the transmembrane
protease, serine 2 (TMPRSS2) gene to the v-ets erythroblastosis virus E26
oncogene homolog (ERG) gene (Tomlins et al., 2009; Barbieri et al., 2013;
Boström et al., 2015). Other members of the ETS family that serves as
partners include ets variant 1, 4, 5 (ETV-1, 4, 5) and Friend leukaemia virus
integration 1 (FLI1) (Paulo et al., 2012). ETS rearrangements occur in 27-79%
of radical prostatectomy and PBx samples, and it has been shown that
transgenic TMPRSS2-ERG mice develop prostatic intraepithelial neoplasia
(PIN). Tumours consisting of TMPRSS2-ERG also show PTEN loss,
suggesting cooperation in prostate oncogenesis (King et al., 2009). ETS
fusions are associated with both aggressive and indolent disease. Some
studies report increased Gleason grade and BCR, while some report lower Gleason score and increased recurrence-free survival (Tomlins et al., 2009). Tumours with ERG rearrangement have increased lesions in PTEN and TP53, the high prevalence of ETS fusions and its association with other oncogenic proteins led to its evaluation as a therapeutic target. A popular interacting enzyme involved in DNA repair, poly (ADP-ribose) polymerase 1 (PARP1) is under clinical investigation in numerous cancers (Fathers et al., 2012). Studies have shown inhibiting PARP1 leads to decreased growth of ETS fusion-positive lesions.

1.5.2.3 Speckle-type POZ protein (SPOP) mutations
Mutations in the speckle-type POZ protein (SPOP) gene are the most common point mutations seen in primary PCa, representing recurrent mutations in 6-13% (Barbieri et al., 2012). SPOP mutations are mutually exclusive with TMPRSS2-ERG fusion and TP53 mutations (Barbieri et al., 2012). Another well-studied protein is the serine peptidase inhibitor, Kazal type 1 (SPINK1), when overexpressed, is associated with decreased biochemical recurrence-free survival (Tomlins et al., 2008).

1.5.2.4 Somatic Copy Number Aberrations and gene expression
Somatic copy number aberrations (SCNAs) are the gain or loss of segments of DNA, leading to oncogenes amplification or TS genes deletion. SCNAs are associated with high-grade disease, advanced tumour stage and other factors associated with poor prognosis reflecting the importance of genomic instability in prostate tumorigenesis (Robbins et al., 2011; Tapia-Laliena et al., 2014).

1.5.2.5 Cytokine signalling
As discussed earlier, inflammation and immunobiology is becoming an ‘emerging’ cancer hallmark (Hanahan et al., 2011). Cytokine-signalling pathways have been shown to be implicated in PCa. The inflammatory cytokine Interleukin-6 (IL-6) is overexpressed in PCa and regulates cell proliferation, apoptosis and angiogenesis through activation of multiple
downstream pathways, such as MAPK (Nguyen et al., 2014). Clinical trials testing anti-IL6 antibody have been initiated (Karkera et al., 2011). In addition, Inhibitors of cytokine-signalling 3 inhibits apoptosis in AR-negative models (Puhr et al., 2009).

1.6 Epigenetics of prostate cancer
Epigenetic mechanisms are inheritable changes that alter expression without changing gene sequence or chromosomal structure. The three main components of epigenetics include DNA methylation, chromatin remodelling and microRNA (miRNA) regulation (Feinberg et al., 2004; Catto et al., 2011; Jerónimo et al., 2011). These mechanisms are reversible unlike genetic events.

1.6.1 DNA methylation
DNA methylation is the best-studied epigenetic change and occurs mostly at cytosines followed by a guanine nucleotide (within CpG dinucleotides). An addition of a methyl group to the fifth carbon of the cytosine residue ring produces 5-methylcytosine (m5C). This process is mediated by DNA methyltransferase (DNMT) and uses S-adenosylmethionine (SAM) as the methyl donor (Lopez-Serra et al., 2008). DNA hypomethylation is associated with oncogenes activation and genetic instability, and DNA hypermethylation is thought to promote gene silencing (Sharma et al., 2010).

Cytosine hypomethylation has been shown to be present in metastatic PCa affecting chromosome instability and disease progression (Bedford et al., 1987). Repetitive DNA regions such as LINE1, are hypomethylated in approximately 50% of PCa, and more abundant in lymphatic metastases (Santourlidis et al., 1999). Genes found to be upregulated following promoter hypomethylation in PCa include IGF2, CAGE, CYP1B1, HPSE, PLAU, CRIP1, S100P, WNT5A (Jerónimo et al., 2011). Interestingly, PLAU gene encodes urokinase plasminogen activator and is associated with castration-resistance.
DNA hypermethylation is the best-known epigenetic mechanism in PCa and over 50 hypermethylated genes have been described. The genes that are hypermethylated play key roles in cell cycle control (CCND2, CDKN2A), apoptosis (ASC, BCL2), hormone response (AR, ESR2), DNA repair (GSTP1, GSTM1), signal induction (EDNRB, RASSF1A, DKK3) and tumour invasion (APC, CAV1, CDH1) (Jerónimo et al., 2011). Identification of hypermethylation (APC, CCND2, GSTP1, PTGS2, RARB2, RASSF1A) in HGPIN and normal prostate tissue, suggests that epigenetic events may occur early in the onset of prostate oncogenesis.

1.6.2 Histone and chromatin modifications
DNA is wrapped around histones to form chromatins and further packed into units of nucleosomes forming chromosomes. Histones regulate DNA transcription, repair and replication, and are prone to post-translational modification (histone 'tail') for example: acetylation, methylation and phosphorylation (Lennartsson et al., 2009). Histone modifications are thought to form a code (histone code) and are associated with transcription activation or repression, and DNA methylation (Jenuwein et al., 2001; Esteller, 2008).

Mutated genes involved in histone modifications include the lysine (K)-specific demethylase 6A gene (KDM6A) and the myeloid/lymphoid or mixed-lineage leukaemia 2 and 3 gene (MLL2, MLL3). These genes encode proteins that alter methylation of the histone variant H3, which is involved in the regulation of chromatin states and transcriptional control. The chromodomain helicase DNA binding protein 1 gene (CHD1) encodes a binding protein that remodels chromatin states, and is recurrently deleted in PCa at around 10-25%. Primary and metastatic PCa with CHD1 deletion have an increase in genomic rearrangements (Barbieri et al., 2013).

Several histone-modifying enzymes including histone deacetylase (HDAC), histone methyltransferase (HMT) and histone demethylase (HDM) are altered in PCa. The best-studied HMT enzyme is Histone-lysine N-methyltransferase (EZH2) which catalyses the trimethylation of histone
H3K27 and dimethylates H3K9 (Cao et al., 2002). EZH2 upregulation is associated with promoter hypermethylation and repression of some genes resulting in increased proliferation rate and aggressiveness of PCa (Viré et al., 2006). Overexpression of HDM and HDAC are associated with hormone-refractory disease. Interestingly, HDAC1 is also overexpressed in PCa containing TMPRSS2-ERG fusion (Halkidou et al., 2004). The HDM, lysine-specific demethylase 1 (LSD1) removes methyl groups from H3K4 and H3K9, and overexpression is associated with aggressiveness and CRPC (Metzger et al., 2005).

1.6.3 Non-coding and MicroRNAs

RNAs are transcribed from DNA in the nucleus and are formed of two types, ones that are translated into protein (coding-RNAs) and ones that are not translated (ncRNAs). Short ncRNAs, named microRNA (miRNA) are around 19-22 bases in length, and were first documented in 1993 (Miah et al., 2014). miRNAs are synthesized (pri-miR) and processed in the nucleus (pre-miR) before exportation to the cytoplasm (mature miR). Nearly 2000 miRNAs are now reported (see http://www.mirbase.org). The prefix ‘miR’ is followed by a number, the latter indicates order of naming/discovery. A capitalized ‘miR’-refers to the mature form of miRNA and the uncapitalized ‘mir’ refers to the pre/pri-miRNA forms. The 5’ end of a miRNA contains the ‘seed’ region that binds mRNAs with complementary sequences at the 3’ untranslated region (3’ UTR) (Bartel, 2009). The miRNA-mRNA pairing recruits a RNA-induced silencing complex (RISC) to modulate mRNA expression (Figure 12). Each miRNA can interact with multiple mRNAs and each mRNA can be targeted by multiple miRNAs. Up to 30% of human genes are regulated by miRNAs. MicroRNAs (~40%) are located within coding mRNAs (intronic or exonic) or on their own close to CpG islands (~20-40%), and exist solitarily or in clusters (~30%). Around 30% of miRNAs are clustered, meaning that a single miRNA triggered event can affect several miRNAs within the same cluster altering thousands of mRNA/protein targets. (Catto et al., 2011).
**Figure 12.** Illustration of RNA synthesis and regulation of mRNA.

RNA is transcribed from DNA in the nucleus. mRNAs are exported and translated into proteins in the cytoplasm. miRNAs contain the ‘seed’ region at the 5’ end and bind complementary sequences at the 3’ UTR end of mRNAs. The pair recruits a RNA-induced silencing complex (RISC) which binds target mRNAs resulting in degradation (perfect complementation) or altered translation (imperfect complementation) of the target mRNAs.
1.6.4 MicroRNA and prostate cancer

Porkka et al. reported the first systematic profiling of miRNA expression in PCa comparing the aberrant expressions of miRNA in PCa and benign prostate cells (Porkka et al., 2007). To date, over 100 reports have investigated miRNA expression in PCa using both cell lines and prostate specimens.

MicroRNA expression are altered in cancer acting as oncogenes when overexpressed, or tumour suppressors when underexpressed (Garzon et al., 2009). Interestingly miRNAs are also involved with DNA methylation and chromatin modification, suggesting interlinks between the three major epigenetic events (Guil et al., 2009). Approximately 50 miRNAs (Table 5) have been reported in PCa, however, not all have been proven to contribute to the disease. In PCa, miRNAs are associated with AR-signalling or androgen-independent growth (MiR- 125b, 146a, 221, 222, 331, 488) and avoidance of apoptosis (MiR- 21 and 34a and c) (Catto et al., 2011).
A range of miRNAs has been identified and investigated in prostate cancer. Common miRNA upregulated and downregulated in prostate cancer, along with their target mRNAs and associated downstream pathways are summarized in this table. All pathways are associated with oncogenesis (Catto et al., 2011; Fabris et al., 2016).

A1, androgen-independence.

Table 5. MicroRNAs expression in prostate cancer.

<table>
<thead>
<tr>
<th>MiRNA</th>
<th>mRNA target</th>
<th>Pathway</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upregulation miR-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20a</td>
<td>E2F1-3</td>
<td>Apoptosis</td>
</tr>
<tr>
<td>21</td>
<td>PTEN, AKT, androgen pathway</td>
<td>Apoptosis, mTOR, AI</td>
</tr>
<tr>
<td>24</td>
<td>FAF1</td>
<td>Apoptosis</td>
</tr>
<tr>
<td>25</td>
<td>PTEN</td>
<td>Proliferation</td>
</tr>
<tr>
<td>32</td>
<td>BCL2</td>
<td>Apoptosis</td>
</tr>
<tr>
<td>93</td>
<td>LATS2</td>
<td>Metastasis</td>
</tr>
<tr>
<td>106b</td>
<td>P21, E2F1</td>
<td>Apoptosis, cell cycle</td>
</tr>
<tr>
<td>125b</td>
<td>P53, BAK1</td>
<td>Apoptosis</td>
</tr>
<tr>
<td>148a</td>
<td>CAND1</td>
<td>Cell cycle</td>
</tr>
<tr>
<td>221</td>
<td>P27</td>
<td>Cell cycle, AI</td>
</tr>
<tr>
<td>222</td>
<td>P27</td>
<td>Cell cycle, AI</td>
</tr>
<tr>
<td>521</td>
<td>Cockayne syndrome protein A</td>
<td>DNA repair</td>
</tr>
<tr>
<td>Downregulation miR-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Exportin 6, tyrosine kinase 9</td>
<td>Gene expression</td>
</tr>
<tr>
<td>7</td>
<td>ERBB-2 (EGFR, HER2)</td>
<td>Signal transduction</td>
</tr>
<tr>
<td>15a-16 cluster</td>
<td>CCND1, Wnt3a</td>
<td>Cell cycle, apoptosis</td>
</tr>
<tr>
<td>34a</td>
<td>HuR/Bcl2/SIRT1</td>
<td>Apoptosis</td>
</tr>
<tr>
<td>34c</td>
<td>E2F3, Bcl2</td>
<td>Apoptosis, proliferation</td>
</tr>
<tr>
<td>101</td>
<td>EZH2</td>
<td>Gene expression</td>
</tr>
<tr>
<td>107</td>
<td>Granulin</td>
<td>Proliferation</td>
</tr>
<tr>
<td>125b</td>
<td>BAK1</td>
<td>Apoptosis, AI</td>
</tr>
<tr>
<td>143</td>
<td>MYO6, ERK5</td>
<td>Cell migration, proliferation</td>
</tr>
<tr>
<td>145</td>
<td>MYC, MYO6</td>
<td>Cell migration, apoptosis</td>
</tr>
<tr>
<td>146a</td>
<td>ROCK1</td>
<td>Proliferation, invasion</td>
</tr>
<tr>
<td>148a</td>
<td>MSK1</td>
<td>Proliferation, stress response</td>
</tr>
<tr>
<td>205</td>
<td>IL-24, -32</td>
<td>Cell growth, invasion, EMT</td>
</tr>
<tr>
<td>331-3p</td>
<td>ERBB-2, CDCA5</td>
<td>Signal transduction, cell cycle</td>
</tr>
<tr>
<td>449a</td>
<td>HDAC-1</td>
<td>Gene expression</td>
</tr>
<tr>
<td>1206</td>
<td>MCM family</td>
<td>DNA replication</td>
</tr>
<tr>
<td>let-7a</td>
<td>E2F2, CCND2</td>
<td>Cell cycle, proliferation</td>
</tr>
<tr>
<td>let-7b</td>
<td>Ras, CycD1</td>
<td>Cell cycle, apoptosis</td>
</tr>
</tbody>
</table>
1.6.4.1 MicroRNA and cellular pathways

Avoidance of apoptosis can be driven by many miRNAs in PCa. For example, the overexpression of miR-20a (also miR25/205) leads to the inhibition of transcription factor E2F1, which in turn, results in cell proliferation and reduced p53 and caspase-mediated apoptosis (Sylvestre et al., 2007; Gandellini et al., 2009). miR-21 targets programmed cell death 4 (PDCD4) and PTEN mRNAs to suppress apoptosis (Lu et al., 2008). Reduced levels of miR-34 has also been shown to inhibit silent information regulator 1 (SIRT1), which inhibits p53 mediated apoptosis. Reduced levels of p53 further inhibits miR-34 resulting in an auto-regulatory loop (Yamakuchi et al., 2008).

miR-15 and -16 are downregulated in ~80% of prostate tumours. Loss of these two miRNAs induces upregulation of cyclin D1 and increased cell proliferation. In addition, reduced miR-15/16 facilitates pro-carcinogenic (cell proliferation and migration) Wnt pathway activation (Bonci et al., 2008).

DNA repair is an important activity in suppressing oncogenesis. Stress, radiation or any form of mechanism that causes DNA damage leads to a complex DNA repair cascade. Upregulation of miR-521 in PCa cells reduces cell response to cell damage by targeting Cockayne syndrome protein A (CSA). miR-34 exert a similar effect through p53 regulation (Josson et al., 2008).

1.6.4.2 MicroRNA and androgen-signalling

Androgen-signalling is one of the most important PCa specific activity. There is a complex link between miRNAs and the androgen pathway. miR-125b is regulated by androgens via an androgen-responsive element (ARE), and upregulation results in androgen-independent growth in LNCaP cells and suppression of apoptosis by targeting BAK1 and p53 (Shi et al., 2007, 2011). Through miRNA profiling studies, miR-146a has been found to be downregulated in hormone-resistant cell lines and transfection with miR-146a results in suppression of ROCK1 and subsequently reduced cell proliferation, invasion and metastasis (Lin et al., 2008). miR-141 is upregulated in
androgen-regulated cells, and its overexpression results in increased PCa growth (Waltering et al., 2011). miR-221 and -222 are both upregulated in CRPC cells and exhibit androgen-independent growth of prostate cell lines (Sun et al., 2009). ERBB-2 tyrosine kinase receptor is overexpressed in PCa and is associated with disease progression and androgen-signalling. miR-331-3p expression is decreased in PCa and transfection of miR-331-3p results in reduced ERBB-2 mRNA and downstream PI3KAKT signalling, and blockade of the AR-signalling pathway (Epis et al., 2009).

1.6.4.3 MicroRNA as biomarkers
Evolving information on the roles of microRNAs has established them as potential biomarkers and therapeutic targets. They are utilized as urinary markers for diagnosing urological cancers and appear to show great potential in managing cancers (Miah et al., 2012). Short or microRNAs appear to be promising biomarkers as their small size protects them from endogenous RNase degradation. They are stable, active and resistant to freeze-thawing (Miah et al., 2012). Its ability to individually, or as clusters to interact with multiple mRNAs involved in PCa pathways makes them attractive as therapeutic targets. In addition, short/miRNAs are detectable in urine, blood/serum, ejaculate and prostate tissue, which are ideal patient specimens to use for diagnosing or monitoring disease progression/recurrence in a clinical setting (Fabris et al., 2016). Current techniques available for miRNA detection include quantitative real-time polymerase chain reaction (qRT-PCR), microarray and small RNA sequencing.

Many studies report numerous upregulated and downregulated miRNAs, with conflicting data. This reflects different profiling strategies, differences in analytical thresholds, study design (samples and methods) and disease heterogeneity. miRNA expression is variable and differs according to the phases of development (initiation, progression or metastasis) or treatment exposure (ADT, radiotherapy, chemotherapy). This diversity allows in-depth evaluation and search for miRNAs as diagnostic and prognostic markers.
Many studies have looked into miRNA as diagnostic and prognostic biomarkers. Fabris et al conducted a systematic review and highlighted plasma miRNAs that were consistently altered in PCa with diagnostic properties. These miRNAs include upregulatory, miR-141, 375, 221, 21; and downregulatory, miR-181a (Fabris et al., 2016).

Profiling experiments report some consistency in results for miRNAs as diagnostic markers. Once PCa is diagnosed and management initiated in the form of radical surgery or radiotherapy, or hormonal therapy, disease recurrence needs to monitored. This usually requires clinical examination, imaging, and some form of biomarker measurement. PSA is expected to fall following surgery/hormones, thus patients with detectable PSA after surgery are thought to have biochemical recurrence (BCR) and detectable PSA following hormonal treatment are thought to have progressed to CRPC (if testosterone is <20ng/dl).

miRNAs have been investigated as biomarkers for BCR and progression of PCa to CRPC and metastatic PCa. Studies focus on analysing RP and PBx specimens and bodily fluids. Some of the miRNAs implicated in BCR include miR-96 (Schaefer et al., 2010; Haflidadóttir et al., 2013; Ilic et al., 2016), -21 (Melbø-Jørgensen et al., 2014; Leite et al., 2015), -221 (Martens-Uzunova et al., 2012) and -1193, -4516, -626 (Bell et al., 2015). Larne et al established a miR index quote (miQ) that could predict PCa aggressiveness and metastatic status, and BCR following RP. The miQ consists of two upregulatory (miR-96-5p and miR-183-5p) and two downregulatory (miR-145-5p and miR-221-5p) miRNAs (Larne et al., 2013).

As with BCR, numerous miRNAs have been identified to be associated with the onset of CRPC or mCRPC. These include, miR-21 (Ribas et al., 2009), -141 (Agaoglu et al., 2011), -221, -222 (Sun et al., 2012), -375, 1290 (Huang et al., 2015). Interestingly most of these miRNA markers of CRPC are also linked to the AR-signalling pathway, such as miR-21, -221 and -222. So far, miR-221 appears to be the most promising diagnostic and prognostic
biomarker, and is associated with clinico-pathologic factors including the Gleason score and clinical recurrence (Spahn et al., 2010; Agaoglu et al., 2011; Larne et al., 2013).

1.6.4.4 RNA methylation

DNA methylation is known to be a common and important epigenetic modification. The final production of proteins give rise to individual phenotypes. Since RNA translation is the final step of gene expression, RNA modifications have also been evaluated extensively. More than hundred types of post-transcriptional modifications have been identified since the 1950s. RNA methylation is the most common modification and has been known since the 1970s. Methylation can occur at the adenosine (m6A) or cytosine (m5C) residues, and at nucleotides (Nm) (N. Liu et al., 2014). The interest in N6-methyladenosine (m6A) revived when Jia et al found that m6A was reversible through demethylator (eraser) FTO (Fat mass and obesity-associated protein). The methyl group is donated from SAM (also implicated in DNA methylation) and is mediated by methylators (writers) such as N6-adenosine-methyltransferases METTL3 and METTL14. The m6A machinery (Figure 13) is diverse and consists of ‘writers’ (methylators, METTL3), ‘erasers’ (demethylators, FTO) and ‘readers’ which facilitates and potentiates methylation (YTH domain family member, YTHDF3) (Niu et al., 2013; Fu et al., 2014).

The first writer identified is METTL3 encoded by the METTL3 gene, and knock down of METTL3 leads to apoptosis in human cell lines. METTL14 also catalyses m6A methylation and form complexes with METTL3 (J. Liu et al., 2014). The METTL3-METTL14 complex interacts with Wilm’s tumour 1-associating protein (WTAP), which is a mRNA splicing regulator involved in controlling cell cycles through stabilization of cyclin A2 mRNA (Horiuchi et al., 2006). Knock down of WTAP results in the largest decrease in m6A in Hela cell lines, indicating that WTAP has an important role in methylation, possibly through enhancing recruitment of METTL3-METTL14 complex to target RNAs (J. Liu et al., 2014). Silencing of the METTL3-METTL14 complex led
to an increase in levels of their target RNAs, suggesting that m6A acts as a negative regulator of gene expression (J. Liu et al., 2014).

The first m6A eraser described was the Fat mass and obesity associated enzyme (FTO), which removes the methyl group through oxidation (Jia et al., 2011). The functional consequences of such demethylation are unclear, however, the discovery of this first demethylator indicates that m6A is subject to sophisticated control. Another m6A eraser is the protein encoded by the alkB, alkylation repair homolog 5 (ALKBH5) gene which belongs to the same protein family as FTO (Zheng et al., 2013). ALKBH5 knockdown in human cell lines results in increased m6A levels and accelerated export of these RNAs from the nucleus to the cytoplasm (Zheng et al., 2013).

Embryonic lethal, abnormal vision (ELAVL1) binds UA-rich regions located in 3'UTR sites of mRNAs (Kundu et al., 2012). RNA-ELALV1 interactions regulate the stability of many mRNAs in embryonic stem cells in a m6A-dependent manner (Y. Wang et al., 2014). YTH-domain family protein, YTHDF2 binds more than 3000 cellular RNAs and competes with ribosomes for translatable mRNAs. Successful binding results in mRNA localisation to mRNA decay sites such as processing bodies (X. Wang et al., 2014). Several cytoplasmic mRNA degradation pathways have been established (Schoenberg et al., 2012), however, the YTHDF2-mediated mRNA degradation is dependent on the methylated state of the target mRNA. Therefore, the m6A state of mRNA could regulate its rate of degradation. YTHDF1 has been reported to interact with initiation factors to promote translation (X. Wang et al., 2014). Recently, eukaryotic initiation factor 3 (eIF3) has been shown to bind to m6A 5'UTR and promote translation under stress (Meyer et al., 2015). HNRNPA2B1 is a nuclear reader of m6A and on binding, regulates pre-mRNA processing and splicing (Alarcón et al., 2015).

Knowing that m6A is reversible and is associated with inflammatory and malignant (Leukaemia, prostate, breast, colorectal, gastric) processes, it has become a target for mass investigation into biological significance and clinical relevance (Maity et al., 2015). The m6A distribution has been...
mapped in liver cell lines, however, its distribution and significance in PCa has not been evaluated to date (Dominissini et al., 2012). The methylation of N6-adenosine in PCa will be explored in this project.

Figure 13. The N6-methyladenosine machinery.

Over one hundred post-transcriptional modifications have been described. RNA methylation is the most common RNA modification and the most common methylated site is at the adenosine base (m6A). N6-adenosine methylation is induced by ‘writers’ including METTL3/14 and WTAP. This process is reversible and demethylation is mediated by ‘erasers’ such as FTO and ALKBH5. RNAs that are exported into the cytoplasm are modulated by ‘readers’ including ELAVL1, YTHDF1-3, eIF3 and HNRNPA2B1, which can affect mRNA processing, exportation, storage, translation and RNA degradation.
1.7 Aims

The evolving knowledge on the biological basis of PCa has allowed the development of markers to aid PCa diagnosis, stratify men for PBx and monitor PCa recurrence (BCR) post-treatment (Post- surgery, radiotherapy or hormonal therapy).

Prostate cancer is a heterogeneous disease and some of the current problems include, 1) a lack of rPBx protocol for patients with initial negative PBx. Automatic rPBx possess procedural complications, and the risk of over-diagnosing and over-treating indolent PCa. Likewise, under-performing rPBx may result in mis-diagnosing aggressive disease; 2) there are no biomarkers that could reliably stratify localised disease from advanced (metastatic) disease on (r)PBx, or biomarkers that could predict progression of PCa to CRPC (when on ADT treatment) or metastatic disease. Hence, urgent investigations into new biomarkers with greater sensitivity and specificity are needed. This is the ultimate aim of the current project.

International data suggests that the PCa diagnostic rate on rPBx is ~30% (Keetch et al., 1994; Djavan et al., 2001). However, rPBx data is not readily available from a UK cohort of patient, therefore a retrospective analysis was performed to identify the rPBx rate in the Sheffield cohort of patients within the national ProtecT RCT and to identify men/specimens for subsequent laboratory analyses.

Knowing that ncRNA PCA3 is currently the most prominent FDA approved urinary marker to date, we aimed to evaluate this RNA further. PCA3 is a long ncRNA and is vulnerable to digestion by urinary RNase, hence prior to analysis in the laboratory, urinary samples need to be treated with RNA inhibitors which ensures a high cost. In addition, the biological role of PCA3 is unknown, therefore it’s not widely adopted worldwide, including the UK NHS. As discussed earlier, short or miRNAs are more stable and active species than long RNAs, in addition, we know that long RNAs are processed into shorter forms (Röther et al., 2011). With this in mind, we aimed to explore this marker to overcome some of the issues related to it, by
identifying a shorter segment within PCA3 and exploring potential biological roles of short-PCA3.

As the project progressed, an increasing interest in RNA methylation occurred. DNA methylation is a key component of epigenetic pathways in cancer. Although RNA methylation has been investigated extensively in the past, a specific methylated site, N6-methyladenosine (m6A) has been shown to be reversible and implicated in cancer. These recent findings initiated further investigation amongst groups interested in cancer epigenetics. Since there are no current data on PCa and m6A, this thesis set out to identify the distribution of m6A in PCa through an in-silico analysis, followed by laboratory evaluation.

In summary, the aim of this thesis was to develop molecular strategies to rationalise rPBx in men with elevated PSA. In order to achieve this, the following was undertaken:

1) A clinical analysis of rPBx outcomes of men who had an initial negative PBx but persistent suspicion of PCa.

2) Identification of a short segment of PCA3 that could potentially replace the current PCA3 assay, and evaluation of its diagnostic and biological role.

3) Identification of the distribution of m6A in PCa.
1.8 Significance

Prostate cancer is the most common cancer in men and is the second most common cause of cancer-related death in the UK. Early diagnosis can improve the care of men with this disease, if cases are identified in a prompt manner and then managed with the most suitable care pathway (such as surveillance for low-risk and radical treatment for high-risk disease). This project aimed to improve the detection of PCa and disease progression using a small RNA within PCA3 or methylated (specifically, m6A) transcripts that are known to be implicated in PCa. With regards to the future, therapeutic RNA molecules that target small ncRNAs are in phase 2 and phase 3 clinical trials (see http://www.santaris.com/product-pipeline) of hepatitis and sold tumours. As such, short RNAs could represent a therapeutic target within PCa.
CHAPTER 2: Materials and Methods
2.1 ProtecT and ProMPT studies

2.1.1 Patient recruitment, consent and ethical approval

All samples including urine and PBx tissue used in the current study were obtained from patients within the ProtecT (Prostate testing for cancer and treatment) and ProMPT (Prostate Cancer: Mechanisms of Progression and Treatment) studies (Donovan et al., 2003; Lane et al., 2010, 2014; Hamdy et al., 2016). Men recruited into the ProtecT study were identified from community medical practices across nine cities within the UK. Letters of invitation detailing the rationale for the ProtecT study were sent to medical practices, and participating practices in turn sent letters inviting men aged between 50 and 69 years with no prostatic symptoms and prior PSA testing to attend for PSA counselling. Nurse led clinics were held in a primary care setting, where participants with an estimated life expectancy of a minimum of ten years and without significant cardio-respiratory co-morbidity were given detailed information about the implications of PSA testing, treatment uncertainties and the need for a RCT. Consent was obtained thereafter for PSA testing.

Men with an initial, single PSA value between 3.0 and 19.9ng/ml were offered a TRUS-PBx. Men with a positive biopsy and a diagnosis of localised PCa were randomised to one of the three treatment arms, 1) active surveillance, 2) radical radiotherapy or 3) radical prostatectomy. Shortly after the commencement of recruitment, a new study, ProMPT was initiated. This study focused on in-vitro and in-vivo analysis of specimens collected from ProtecT patients. A detailed rationale for the study was discussed with ProtecT patients and a new consent was sought for permission to collect and analyse their urine, blood and tissue specimens.

Between 2001 and 2009, 227,000 men were identified at 352 practices and invited for PSA counselling within the ProtecT study at a nurse led clinic. A total of 111,148 men attended and 10,297 were offered TRUS-PBx. Within the Sheffield cohort, 16,656 men were invited for PSA testing, 10,412
underwent testing, and 920 had a TRUS-PBx. Prostate cancer was diagnosed in 321 men within Sheffield and were subject to randomization (Figure 14). The 599 men who had a negative PBx were included in the rPBx analysis (Chapter 3).

The ProMPT study began recruitment in 2002 and all centres have stopped recruitment except for Oxford at the time of writing. To date, 6414 men have consented to participate in ProMPT. Sheffield recruited 1578 men between 2002 and 2015. ProMPT is a collaborative translational research group studying advanced and progressing PCa, focusing on molecular pathology.

Baseline clinical and histopathological data including age at initial PBx, age at diagnosis of cancer, number of PSA tests and PBx undergone per patient, PSA value/PSA kinetics, PBx characteristics (laterality, number cores, taken, number cores positive and length), Gleason score, stage were obtained from the ProtecT database, or from laboratory (ICE)/radiological (PACS) computer resources at the Royal Hallamshire Hospital, Sheffield Teaching Hospitals NHS Trust.

The ProtecT study is registered with clinicaltrials.gov (NCT02044172), and as an International Standard Randomised Controlled Trial (ISRCTN20141297). The ProtecT trial and the later ProMPT study received approval from the Trent Multicentre Research Ethics Committee, MREC (HTA 96/20/06; HTA 96/20/99) and the National Research Ethics Service (NRES 01/04/061).
Figure 14. Flow chart of the study population within the ProtecT/ProMPT study.

Patients were recruited from nine UK cities (Blue map) between 2001 and 2009. Within Sheffield, a total of 321 men were diagnosed with PCa on first biopsy, 599 men had a negative biopsy and were subject to a rPBx.

ProtecT, Prostate testing for cancer and treatment; PSA, prostate-specific antigen; TRUS-PBx, transrectal ultrasound-guided prostate biopsy; PCa, prostate cancer.
2.2. General laboratory equipment and reagents

2.2.1 Laboratory equipment

Glassware
Glassware used for this project was washed with RBS detergent (Chemical concentrates), tap water and de-ionised water. Washed glassware was dried in a hot oven and items that required sterilization were autoclaved for 15 minutes at 15p.s.i.

-80°C freezer
4°C, -20°C fridge/freezer
Ice machine AF100
Heraeus megafuge 40
Geneflow microcentrifuge
Vortex genie 2
Sub-aqua dual water bath
Barnsread nanopure water system
Forma CO₂ incubator
GeneAmp PCR 2700 (Thermal cycler)
HT7900 PCR system
Qiagen rotor-gene PCR
Biological safety cabinet
Nanodrop bioanalyzer
P2, 10, 20, 200, 1000 pipettes
Immunoprecipitation magnetic rack

Sanyo
Liebherr
Scotsman
ThermoFisher
Sigma
Scientific Industries
Grant
ThermoFisher
Applied Biosystems
Qiagen
ThermoFisher
Agilent
Gilson
ThermoFisher
2.2.2 Plastic and disposable equipment

Plastic pasteur pipettes
5, 10, 25ml Plastic pipettes
Filter tips for Gilson P10, 20, 200, 1000
0.5, 1.5, 2ml eppendorf tubes
1.5, 2ml microcentrifuge tubes
15ml, 50ml centrifuge tubes
BD Microlance 21G, 23G needles
Plastipak 1ml, 3ml syringes
T75, T175 NUNC cell culture flasks with filter cap
Nalgene cryogenic 2ml vials
Microamp 0.1, 0.2ml PCR tubes
96 well PCR plates
8 strip PCR caps
384 well PCR plates
384 well PCR plate seals
Semperguard nitrile powder-free gloves

2.2.3 General laboratory chemicals and reagents

Ethanol
Precept
FCS (Fetal calf serum)
Fetal calf serum
NEAA (Non-essential amino acid)
Penicillin-Streptomycin
Trypsin
DMSO (Dimethyl-sulfoxide)
Molecular biology water
β-Mercaptoethanol (β-ME)
2.3 General materials and methods

2.3.1 Cell Lines and cell cultures

a) Cells (Table 6) were purchased from ATCC and grown in appropriate media according to standard methods (https://www.lgcstandards-atcc.org/Products/Cells_and_Microorganisms/Cell_Lines.aspx?geo_country=gb) (Leiblich et al., 2006).

<table>
<thead>
<tr>
<th>Cell Line</th>
<th>Androgen-sensitive</th>
<th>Derivation</th>
<th>Metastasis</th>
</tr>
</thead>
<tbody>
<tr>
<td>LNCaP</td>
<td>Yes</td>
<td>Lymph nodes</td>
<td>Lymph nodes</td>
</tr>
<tr>
<td>LNCaP-LN3</td>
<td>Less than LNCaP</td>
<td>Lymph nodes</td>
<td>Lymph nodes</td>
</tr>
<tr>
<td>DU145</td>
<td>No</td>
<td>Brain</td>
<td>Liver, lung</td>
</tr>
<tr>
<td>PC3</td>
<td>No</td>
<td>Vertebræ</td>
<td>Liver, lung, bone</td>
</tr>
</tbody>
</table>

Table 6. Prostate cell lines.

The common prostate cell lines used in the current studies are demonstrated. Its sensitivity to androgen, origin and sites of metastasis are detailed in this table. LNCaP-LN3 is the metastatic sibling of LNCaP (Sobel et al., 2005a).

b) Phosphate Buffered Saline (PBS)

Oxoid PBS tablets were purchased from Unipath Ltd. One tablet of PBS was dissolved in 100ml de-ionised water and autoclaved at $115^\circ$C for 10 mins to sterilize.

All cell lines were cultured in T75 and T175 flasks with filtered caps in 30-60ml RPMI L-Glutamine (Lonza) media supplemented with: 10% (50ml) FCS, 10ml Penicillin (10,000U)-streptomycin (10mg), 10ml NEAA (10mM). LNCaP cells, LNCaP-LN3, DU145 or PC3 cells (Sobel et al., 2005a, 2005b) were split or harvested when they reached 80-90% confluence. All cell lines were incubated at $37^\circ$C and 5% CO$_2$ and grown in monolayer.

On harvesting, RPMI medium was removed and cells washed twice with 10-25ml PBS depending on the size of the flask. A volume of 1.5-3ml of Trypsin
was added and flasks incubated for 5mins at 37°C to detach adherent cells. Cells were lifted of the base and either redistributed into two new flasks containing media (splitting) or washed again with PBS (to inactivate trypsin) for cell counting.

For counting, cells were re-suspended in PBS and 10µl was used for the microscope-counting chamber (hemocytometer). Cells in two 4x4 (1mm) quadrants were counted and the total divided by two, giving X x10^4 number of cells per ml. Cells were washed with PBS after counting and either stored at -80°C or used immediately for RNA extraction.

2.3.2 Urinary sample preparation
All urinary samples were obtained from patients within the ProtecT and ProMPT studies. Samples were collected in out-patient urology consultant or nurse led clinics at Royal Hallamshire Hospital, Sheffield. In order to optimize samples for PCA3 analysis, urine was obtained following prostatic massage (Hessels et al., 2003). The first 10-20ml of freshly voided urine was collected and immediate centrifugation at ~3,300g for 10 mins was undertaken. The supernatant was removed and the cell pellet washed twice in PBS before freezing at -80°C until use.

2.3.3 Prostate tissue collection
All prostate tissue samples were obtained from men within the ProtecT and ProMPT studies. Out-patient TRUS-PBx was carried out in the left lateral decubitus position under peri-prostatic infiltration of local anaesthetic (1-2% Lignocaine) and antibiotic cover by using a 10-core lateral biopsy template. Saturation biopsies (20-30 cores) +/- transurethral resection biopsies of the transitional zones of the prostate (TURP) were performed in selected men with ≥1 negative rPBx at the discretion of the attending clinician.

2.3.3.1 Prostate tissue fixation and embedding
Formalin-fixed Paraffin embedded (FFPE) tissues were prepared as follow. Fresh PBx cores of tissue were fixed in freshly prepared 4%
paraformaldehyde for 3 hours at room temperature. The specimens were dehydrated by passing through ascending alcohol concentrations (50% for 30 mins; 70% 30 mins; 90% 1 hour; absolute for 1 hour), then cleared in xylene for 30 mins at 60°C in a glass pot. The tissue was then embedded in a 1:1 xylene/molten paraplast mixture at 60°C for 30 mins, then in fresh paraplast at 60°C and allowed to solidify by cooling to room temperature. Embedded blocks were stored at room temperature in the histopathology department, Royal Hallamshire hospital until use.

2.3.3.2 FFPE sectioning
In order to obtain enriched tissue for analysis, the best block with the most abundant tissue visually from each patient was subject to sectioning. A microtome was used to cut 4x 10µm section per block, placed in 1.5ml eppendorf tubes and stored at -20°C (Performed by Maggie Glover).

2.4 Specific materials and methods

2.4.1 RNA Extraction
Total and/or miRNA were extracted from cell lines, urinary pellets and FFPE prostate biopsies from three different commercially purchased RNA extraction kits.
2.4.1.1 Cell lines and urinary pellets

a) MirVana RNA Isolation Kit (Ambion)
This kit was used to isolate RNA from cell lines and urinary pellets for the PCA3 study. Extraction was performed according to the manufacturer’s protocol.

Materials
- PBS
- Acid-Phenol:Chloroform (APC)
- Lysis buffer
- miRNA homogenate additive
- Wash solution
- Elution solution

Approximately \(10^2\)-\(10^7\) cells were washed following re-suspension in PBS and placed on ice. PBS was removed and 300-600\(\mu l\) lysis solution for 100-\(10^7\) cells was added. Solution was vortexed and pipetted vigorously to completely lyse the cells and obtain a homogenous lysate. One tenth volume of miRNA homogenate additive was added, mixed well and placed on ice for 10 mins. A volume (300-600\(\mu l\)) of APC that was equal to the lysate volume before the addition of homogenate additive was added, and vortexed for 30-60 sec. The upper aqueous phase was removed (without disturbing the lower phase) and transferred to a fresh tube. For small RNA extraction, 1/3 (1.25 volume for total RNA) volume of room temperature 100% ethanol was added to the previously collected aqueous phase. A filter cartridge was placed into a collection tube, and the lysate/ethanol mix was pipetted onto the filter cartridge. A max. of 700\(\mu l\) was applied to the cartridge at a time and centrifuged at RCF 10,000xg (10,000rpm) for 15 sec to pass the mixture through the filter. The mixture was applied in successive applications to the same filter. The flow-through was discarded, and repeated until all the mixture was through the filter. For total RNAs, the filtrate was subject to
wash and RNA elution. Extra steps were required to extract small RNAs (described below).

*Enrichment for small RNAs*

These steps were only applicable to small RNAs. The filtrate was collected and if the initial lysate/ethanol mix was >700µl, the flow-through was transferred to a fresh tube, and steps repeated until all of the lysate mixture was through the filter. The collected filtrate was pooled 2/3 vol. room temperature 100% ethanol was added to the filtrate and mixed. The filtrate/ethanol was pipetted onto a second filter cartridge. Up to 700µl can be applied at a time. For greater volumes, the mixture was applied in successive applications to the same filter. The mix was centrifuged for 15 sec at RCF 10,000xg (10,000rpm). The flow-through was discarded and repeated until all of the filtrate/ethanol was through the filter.

*Washing and elution*

The steps in washing and eluting was the same for small and total RNAs. 700µl of wash 1 solution was applied to the cartridge and centrifuged for 5-10 sec. The flow-through was discarded and the cartridge was placed into the same collection tube. The filter was washed twice with 500µl wash solution 2/3. The flow-through was discarded each time, and the filter was spun for 1 min after the 2nd wash. The cartridge was transferred into a fresh collection tube and 100µl of pre-heated (95°C) elution solution (or nuclease-free water) was applied to the center of the filter. The tube was spun for 20-30 sec at max. speed to recover the RNA. The eluate (containing RNA) was collected and stored at -80°C or placed on ice and used immediately for reverse transcription PCR to cDNA. RNA was analysed using 2100 bioanalyzer (Agilent) after purification. The concentration was determined by measuring the absorbance at 260nm (A_{260}). The ratio of A260 to A280 provides an indication of RNA purity. For highly pure RNA a ratio of 1.8-2.1 is expected.
b) Prime 5 PerfectPure RNA cultured cell kit (Scientific Laboratory Supplies)

This kit was used to isolate RNA for the N6-methyladenosine analyses as suggested by Dominissini et al (Dominissini et al., 2013). Extraction was performed according to the manufacturer’s protocol.

Materials
- Lysis solution + 143mM β-Mercaptoethanol (β-ME)
- Wash 1 solution
- DNase solution
- DNase wash solution
- Wash 2 solution
- Elution solution

Lysis solution (800µl for 20-50x 10^6 cells) was added to the frozen pellets to release nucleic acid from the cells. The lysate was passed through 21 and 23 gauge needles 8-10 times and vortexed vigorously for 2 mins until pellet was resuspended. 400µl lysate was pipetted onto a purification column and centrifuged for 1 min at 14,500xg. This was repeated until all the lysate was passed through. The column was placed into a new tube and 400µl Wash 1 solution was added and centrifuged for 1 min at 14,500xg. The column was placed into a new tube and 50µl DNase solution was added and incubated at room temperature for 15 mins to remove genomic DNA. 200µl DNase wash solution was added at centrifuged at 14,500xg for 1min and a further 200µl DNase wash solution was added and centrifuged at 14,500xg for 2 mins. The column was transferred to a new collection tube and 200µl Wash 2 solution was added and centrifuged at 14,500xg for 1 min. A further 200µl Wash 2 solution was added and centrifuged at 14,500xg for 2 mins. The column was transferred to a new collection tube and 50µl Elution solution was added and centrifuged at 14,500xg for 1 min. RNA was analysed using 2100 bioanalyzer (Agilent) after purification.
2.4.1.2 Prostate biopsy FFPE

miRNeasy FFPE Kit (Qiagen)
This kit was used to extract total and miRNA for the PCA3 rPBx analysis.

Materials
- Deparaffinization solution (Qiagen)
- Protein kinase digestion (PKD) buffer
- Proteinase K
- DNase booster buffer
- DNase I stock solution
- RBC buffer
- 100% Ethanol
- RPE buffer
- RNase-free water

Approximately 300µl Deparaffinization solution was added to the FFPE microdissected sections and vortexed for 10 sec to remove the paraffin. Samples were incubated at 56°C for 3 mins and allowed to cool to room temperature. 150µl Buffer PKD was added, tubes vortexed and centrifuged at 11,000xg (10,000rpm) for 1 min. 10µl proteinase K was added to the lower clear phase and mixed gently to release RNA from the sections. Samples were than incubated at 56°C for 15 mins and then at 80°C for another 15 mins. The lower, clear phases were transferred into a new 2ml microcentrifuge tube and incubated on ice for 3 mins, then centrifuged for 20 mins at 20,000xg (13,500rpm). The supernatants were transferred into a new microcentrifuge tube and 25µl DNase Booster Buffer and 10µl DNase I stock solution were added to eliminate all genomic DNA. Samples were mixed gently and incubated at room temperature for 15 mins. 320µl Buffer RBC was added to adjust binding condition, and mixed. 1120µl ethanol (100%) was added and mixed well by pipetting up and down. 700µl of the sample was transferred to an RNeasy MinElute spin column placed in a 2ml collection tube and centrifuged for 15 sec at 8,000xg (10,000rpm). The flow-
through was discarded. This step was repeated until the entire sample has passed through the spin column. 500µl Buffer RPE was added to the spin column and centrifuged for 15 sec at 8,000xg (10,000rpm). The flow-through was discarded. 500µl Buffer RPE was added to the spin column and centrifuged for 2 mins at 8,000xg (10,000rpm). The flow-through was discarded. The spin column was placed in a new collection tube and centrifuged at full speed for 5 mins. The collection tube with the flow-through was discarded. The spin column was placed in a new 1.5ml collection tube, 25µl RNase-free water was added and centrifuged for 1 min at full speed to elute the RNA. RNA was analysed using 2100 bioanalyzer (Agilent) after purification.

2.4.2 Reverse transcription
A number of different kits were used for reverse transcription (RT) of RNA to cDNA. In general, the high-capacity kit RT cDNA kit (Applied Biosystems) was used for mRNA, and TaqMan RT microRNA kit (Applied Biosystems) was used for miRNA studies.

a) High capacity cDNA reverse transcription kit (Applied Biosystems)

Materials
- 10x RT buffer
- 25x 100mM deoxyribonucleotide triphosphates (dNTP)
- 10x RT random primers
- MultiScribe reverse transcriptase
- Molecular biology water (Sigma)

High-capacity cDNA reverse transcription kit (Applied Biosystems) was used to reverse transcribe RNA to cDNA. The total input amount of total RNA used was up to 2µg of total RNA per 20µl reaction. This input varied depending on the type of samples used, for urinary RNA, up to 200ng was used.
For one sample, 2x RT master mix was prepared (10µl), mixed and placed on ice. The mix contained 2µl 10x RT buffer, 0.8 µl 25x dNTP (100mM), 2µl 10x RT random primers, 1µl MultiScribe reverse transcriptase, 1µl RNase inhibitor, 3.2µl Nuclease-free (NF) water. 10µl 2x RT master mix was pipetted into an individual 0.5ml eppendorf or into each well of a 96-well reaction plate. 10µl RNA (up to 2µg) was added to the 10µl of master mix and mixed gently by pipetting up and down (total 20µl volume). Plates were sealed with 8-cap strips. The plates/tubes were briefly centrifuged to spin down the contents. The plates/tubes were placed in a PCR thermal cycler set to run under the following conditions: 25°C 10 min, 37°C 120 min, 85°C 5 min, 4°C finish.

cDNA was stored at -20°C or placed on ice and used immediately for quantitative PCR (qPCR).

b) TaqMan MicroRNA reverse transcription kit (Applied Biosystems)

Materials
- 100mM dNTP (with dTTP)
- MultiScribe Reverse transcriptase, 50U/µl
- 10x RT buffer
- RNase inhibitor, 20U/µl
- NF water

The total RNA input ranged between 1-10ng as recommended by the manufacturer. For each reaction, 7µl of master mix (0.15µl dNTP, 1µl reverse transcriptase, 1.5µl RT buffer, 0.19µl RNase inhibitor, 4.16µl NF water), 3µl 5x RT primer and 5µl RNA (containing 1-10ng) was combined in a 0.2ml RT PCR tube or in a well of a 96-well PCR plate. The mix was incubated on ice for 5 mins and then incubated in the thermal cycler at 16°C for 30 mins, 42°C 40 mins, 85°C 5 mins and 4°C finish.
cDNA was stored at -20°C or placed on ice and used immediately for quantitative PCR (qPCR).

2.4.3 Real-time PCR
Quantitative real-time PCR was performed using two kits. SYBR select (Applied Biosystems) was used for cDNA synthesized from mRNA, and TaqMan assays/master mix (Applied Biosystems) was used for cDNA synthesized from miRNA.

a) SYBR select master mix (Applied Biosystems)

Materials
- SYBR Select Master mix (2x)
- NF water
- Custom designed primers 3mM

A 10µl reaction was made up for each sample containing 5µl SYBR select master mix, 1µl forward primer, 1µl reverse primer, 2µl NF water and 1µl cDNA. The mixture was made up in a 0.2ml PCR tube or in wells of a 384-well PCR plate and centrifuged briefly. The plate was sealed with a PCR adhesive seal and qPCR was performed on Qiagen rotor-gene (small sample tubes) or HT7900 (Applied Biosystems) (384-well plates) PCR machines. The cycling mode was as follow:

- Polymerase Activation: 50°C 2 mins then 95°C 2 mins
- Denature: 95°C 15 sec
- Anneal: 55-60°C 15 sec 45 cycles
- Extend: 72°C 1 min
b) TaqMan gene expression master mix (Applied Biosystems)
The custom designed TaqMan assays (primers/probes) were used with the TaqMan master mix.

Materials
- TaqMan gene expression Master mix (2x)
- NF water
- Custom designed TaqMan assay (20x)

A 20µl reaction was made up for each sample containing 10µl TaqMan master mix, 1µl TaqMan assay and 9µl cDNA/NF water (containing 10-100ng cDNA). The mixture was made up in wells of a 384-well PCR plate and centrifuged briefly. The plate was sealed with a PCR adhesive seal and qPCR was performed on HT7900 (Applied Biosystems) (384-well plates) PCR machine. The cycling mode was as follow:

- Polymerase Activation: 50°C 2 mins then 95°C 10 mins
- Denature: 95°C 15 sec
- Anneal/extent: 60°C 1 min 45 cycles

2.5 Immunoprecipitation and sequencing of m6A
RNA extraction, immunoprecipitation (IP), sequencing and bioinformatic analyses were performed as per Dominissini et al’s protocol (Dominissini et al., 2013), this protocol has been slightly modified by Professor Chuan He’s group, Department of Chemistry, University of Chicago, USA (Hsu et al., 2018).

Total RNA was extracted from each cell line LNCaP and LNCaP-LN3 using Prime 5 PerfectPure RNA cultured cell kit (Scientific Laboratory Supplies).
2.5.1 Immunoprecipitation of m6A

Materials
- Tris-hydrochloride pH 7.0 and 7.4, 1M (Tris-HCl, Sigma-Aldrich)
- RNasin Plus RNase inhibitor (Promega)
- NaCl (Sodium Chloride, Sigma-Aldrich)
- Igepal CA-630 (Sigma-Aldrich)
- Anti-m6A rabbit polyclonal antibody (Synaptic Systems)
- m6A 5’-monophosphate sodium salt (Sigma-Aldrich)
- Ethanol (Sigma-Aldrich)
- Dynabeads Protein A for immunoprecipitation (ThermoFisher)

Reagents setup
- IP buffer (5x). 10ml was made with 0.5ml Tris-HCl (1M, pH7.4), 1.5ml NaCl (5M), 0.5ml Igepal CA-630 (10% vol/vol stock) and RNase-free water.
- Elution buffer (1x). 90µl 5x IP buffer, 150µl m6A salt (20mM), 7µl RNasin Plus and 203µl RNase-free water.

2.5.1.1 mRNA purification and RNA fragmentation

*mRNA purification*
mRNA purification was performed using the Dynabeads mRNA Purification Kit (Ambion)

Materials
- Dynabeads oligo (dT)25
- Binding buffer
- Washing Buffer B
- 10mM Tris-HCl (pH7.5)
A total of 100µg of total RNA from each LNCaP and LNCaP-LN3 cell line was used. The RNA was adjusted to 100µl with RNase-free water and heated to 65°C for 2 mins to disrupt secondary structures. Dynabeads were resuspended and 250µl was transferred to a micro-centrifuge tube and placed on a magnetic rack for 30 sec. The supernatant was discarded and the Dynabeads were washed in 100µl Binding Buffer. Supernatant was discarded and the beads were re-suspended in 100µl Binding Buffer.

Total RNA was added to the Dynabeads/Binding Buffer suspension and mixed on a rotator for 10 mins at room temperature to allow mRNA to anneal to the oligo (dT)25 on the beads. After 10 mins, the tube was placed on the magnet, and the supernatant was removed. The beads were washed twice with 100µl Washing Buffer B and then 20µl 10mM Tris-HCl (pH7.5) was added to the beads and incubated at 65°C for 2 mins. The tube was placed immediately on the magnet and the eluted mRNA was transferred to a new RNase-free tube.

**mRNA fragmentation**

A total of 1µg of mRNA (polyA+) was made up to 100µl with RNase-free water in 0.65ml sonication tubes (diagenode). RNA was fragmented to ~200nt using a sonication machine (diagenode) as follow:

- 4°C 30 sec on
- 4°C 30 sec off
- 4°C Hold

### 2.5.1.2 Immunoprecipitation

Approximately 5ng (5µl) of untreated fragmented RNA was saved to serve as input control in RNA-seq. The remaining 95µl of fragmented RNA (per cell line) was made up to 500µl in a IP mix containing 100µl 5x IP buffer, 6µl m6A-antibody (0.5mg/ml), 5µl RNasin (100U) and 294µl RNase-free water. The RNA/IP mix was incubated with head-over-tail rotation for 2 hours at 4°C.
During incubation, 40µl (per reaction) protein A Dynabeads were washed twice in 1ml 1x IP buffer on the IP magnetic rack. The IP mix (following 2h of incubation) was transferred to the bead-containing tubes and incubated further on a rotating wheel for 2 hours at 4°C. After 2 hours, the beads were spun down and washed 3x in 1ml 1x IP buffer.

2.5.1.3 Elution
A total of 50µl elution buffer was added to each sample and incubated for 30 min on a rotator at 4°C. Beads were spun down and supernatant removed at retained (now containing RNA). This elution step was repeated once more. All eluates from the same sample were combined (total 100µl) and beads were discarded. Each tube containing eluates was supplemented with 1/10th 3M sodium acetate (pH5.2) and 2.5 volumes of 100% ethanol. The tubes were mixed and incubated overnight at 80°C.

2.5.1.4 RNA purification
RNA purification was performed using the RNA Clean & Concentrator kit (ZYMO RESEARCH).

Materials
- RNA Binding Buffer
- RNA Prep Buffer
- RNA Wash Buffer

Each 100µl eluted RNA was mixed with 200µl RNA Binding Buffer (1:2) and equal volume (300µl) of 100% ethanol was added. The sample was transferred to a Zymo-spin column in a collection tube and centrifuged for 30 sec at max. speed. The flow-through was discarded and 400µl RNA Prep Buffer was added to the column and centrifuged for 30 sec. The flow-through was discarded and 700µl RNA Wash Buffer was added and centrifuged for 30 sec. The flow-through was discarded and a further 400µl RNA Wash buffer was added and centrifuged for 2 mins to ensure complete removal of
the Wash Buffer. The column was transferred to a new RNase-free tube and 14µl RNase-free water was added and centrifuged for 1 min. The eluted RNA was stored at -80°C overnight.

2.5.1.5 Immunoprecipitation quality control
The success of IP was validated by using qRT-PCR to assess the presence of methylated transcripts in the m6A-antibody IP samples relative to bead-only IP (control) samples. Methylated transcripts were chosen from the list provided by Dominissini et al (Dominissini et al., 2012).

2.5.2 Library preparation and sequencing
TruSeq Stranded mRNA library preparation kit Set A (Illumina) was used to produce cDNA from input (untreated fragmented) and IP RNA. Libraries were subject to next generation sequencing on the Illumina Hiseq 2500 machine.

Materials (Illumina kit)
- Fragment, Prime, Finish Mix (FPF)
- SuperScript II enzyme
- Act D
- Second Strand Marking Master mix (SMM)
- Resuspension buffer
- A-tailing mix (ATL)
- Index adaptors
- Ligation mix (LIG)
- Stop Ligase Buffer (STL)
- RNase-free water
- Ethanol 80%
- AMPure XP beads (Beckman)
- PCR Master mix
- PCR Primers
Approximately 7µl IP mRNA and the previously saved 5µl fragmented untreated input mRNA was made up to 18µl with FPF mix. The input and IP samples were incubated at 94°C for 20 sec to fragment and prime the RNA. First strand synthesis was performed by adding 8µl mixture containing 0.8µl SuperScript III and 7.2µl Act D to each sample and running the samples under the following conditions:

- 25°C 10 min
- 50°C 15 min
- 70°C 15 min
- 4°C Hold

Second strand synthesis (20µl product) was performed using the SX-8G IP-Star Compact Automated System (diagenode). In brief, second strand synthesis involves incubating and washing (80% Ethanol) the samples under various temperatures (automated) with,

1) 25µl second strand synthesis mixture (20µl SMM and 5µl Resuspension buffer)
2) 15µl A-tailing mix (12.5µl ATL and 2.5µl Resuspension buffer)
3) 7.5µl ligation mix (2.5µl index adaptor, 2.5µl LIG and 2.5µl Resuspension buffer)
4) 5µl STL

Following ligation, the samples were purified using AMPure beads (Beckman) and amplified using qPCR as per the Illumina protocol with 25µl PCR master mix, 5µl PCR primers, 1µl RNase-free water and 19µl cDNA library. The PCR mixture (50µl) was incubated under the following conditions:
- 98°C 20 sec
- 98°C 10 sec
- 60°C 30 sec
- 72°C 30 sec
- 72°C 30 sec
- 5°C Hold

The PCR product was then purified using AMPure beads (Beckman). The 50µl PCR product was incubated with 50µl AMPure beads (1:1) at room temperature for 15 mins. The mixture was put on the magnet for 5 mins and the supernatant was discarded. The beads were washed twice on the magnet with 200µl 80% ethanol and left to air dry (~15 mins). The beads were resuspended with 22µl RNase-free water and incubated at room temperature for 2 mins. The mixture was put on the magnet for 5 mins and 20µl of the eluate was removed and placed on ice.

Bioinformatic analysis was performed by colleagues from the University of Sheffield Bioinformatic hub. The analysis design was performed in 4 steps, 1) quality control; 2) mapping with splice aware mapper; 3) peak calling with MACS2; 4) peak annotation. Data were analysed as ChIP-seq data instead of RNA-seq data as per Dominiissini et al’s protocol (Dominiissini et al., 2013).
2.6 Statistical Analysis

Statistical analyses were performed using Statistical Package for Social Sciences (Version 23.0, 2016, SPSS, Inc., Chicago, IL, USA) software. In general, continuous data between groups were compared using Student’s t-test (two independent groups) or 1-way analysis of variance (ANOVA, greater than two independent groups) for parametric data, and Mann-U-Whitney test for non-parametric data. Categorical data were analysed using Chi-Square test or Fisher’s exact test (for small sample size). A p-value <0.05 was considered statistically significant. For average calculations, mean and standard deviation (SD) were used for parametric data, and median and interquartile range (IQR) were used for non-parametric data. Specific statistical analysis (for example Kaplan-Meier plots, logistic regression, multivariable analysis) are described in relevant result chapters. All graphs were plotted using SPSS, GraphPad Prism (Version 6.0, 2014) or Excel (Microsoft, version 15, 2016).
CHAPTER 3: A Retrospective Analysis of Repeat Prostate Biopsy Outcomes
4.1 Background

Transrectal ultrasound-guided prostate biopsy (TRUS-PBx) has historically been the gold-standard investigation and tissue diagnosis of PCa. However, as discussed in Chapter 1.3.7.5 around 30% or patients who undergo a PBx are diagnosed with PCa and around the same percentage, ~30% are diagnosed with PCa on repeat PBx (rPBx) (Keetch et al., 1994; Djavan et al., 2001).

There is limited data regarding the yield and predictors for PCa on initial and rPBx in a UK population, where the rate of PSA testing is low compared to other comparable countries (Melia et al., 2004). As such, the risks of PCa detection on initial and rPBx are difficult to quote when counselling men with a negative initial biopsy. There is no validated agreed protocol defining the need for rPBx and the European Association of Urology (EAU) guidelines only provide indications for the first rPBx (Mottet et al., 2017).

Within the ProtecT RCT, those diagnosed with PCa on PBx were randomised to active surveillance, radical radiotherapy or radical prostatectomy. Those with initial negative PBx underwent routine care (Lane et al., 2010; Hamdy et al., 2016).

Biomarkers such as the FDA approved PCA3 are used to help stratify men for rPBx, with the aim of reducing under- or over-treatment of disease. The primary aim of the current analysis was to report a single centre’s experience on the PCa detection rate on pre-MRI rPBx in a previously unscreened population of community-dwelling men in the UK with a negative initial PBx within the ProtecT study. Urinary and prostate biopsy specimens (FFPE) from men (Sheffield) within the ProtecT/ProMPT studies were retrieved for PCA3-shRNA2 evaluation (Chapter 4 and 5).

The secondary aim was to identify clinical features that may predict PCa and aid decision to rPBx.
4.2 Methods

4.2.1 Design and patient population
The current study was a prospective observational cohort study within the ProtecT study. Within the Sheffield cohort, 16656 men were invited for PSA testing, 10412 men underwent testing and 920 had a TRUS-PBx. Of these, 599 (65%, 95% CI 62-68) men had a negative PBx and were included in the current analysis. The interval of repeat PSA testing was left at the discretion of the attending Urologist. Men with inadequate or HGPIN findings on previous PBx and %free PSA <12% were offered a rPBx. Prostate tissue was obtained via TRUS-PBx, as described in Chapter 2.3.3.

4.2.2 Data collection and analysis
Data regarding age, PSA kinetics (PSA value, %free PSA, PSA velocity) and PBx/prostatectomy characteristics including pathology, Gleason scores, biopsy cores, tumour length, tumour volume and staging were gathered. Tumours with a Gleason 7-10 score were classified as high-grade tumours. Indolent tumours were classified using the Epstein criteria (Gleason score ≤6; <3 cores positive; ≤50% positive per core) (Epstein et al., 1994). Characteristics between biopsy groups, and patients with and without PCa were compared.

4.2.3 Statistical analysis
General statistical tests used were described in Chapter 2.6. PCa diagnostic rates were analysed using Kaplan-Meier survival estimates and log-rank test. Time to PCa diagnosis was calculated from the date of first PCa clinic. Patients who died, who were lost to follow-up or who have not reached endpoint were censored. Uni- and multivariable logistic regression analysis was used to identify predictors of PCa on rPBx.
4.3 Results

4.3.1 Patient population within the Sheffield cohort
In total, 321/920 (34.9%) men (mean age (standard deviation) 62.1 (+4.9) years) undergoing TRUS-PBx between 30 November 2001 and 28 November 2008 were found to have PCa on initial biopsy. Of the remaining 599 men (mean age 61.8 (+5.0) years, median (interquartile range (IQR)) baseline PSA of 4.1ng/ml (3.5-4.5) and a median follow-up of 17 (10-45) months), 248 (41.4%) had >1 rPBx (Table 7). A total of 66/248 (26.6%) men who underwent a rPBx were found to have PCa.

4.3.2 Outcomes of repeat prostate biopsy
A total of 337 rPBx were performed, a single rPBx was performed on 248 men; 2nd rPBx on 71; a 3rd on 16; and a 4th on 2 men. Clinical parameters and histopathology results of rPBx are demonstrated in Table 8. PCa was detected in 66/337 (biopsy yield, 19.6%) rPBx with 41/248 (16.5%) and 25/89 (28.1%) PCa detected on 1st and 2nd – 4th rPBx respectively (Table 9 and Figure 15).
<table>
<thead>
<tr>
<th>Characteristics</th>
<th>n= 599</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline Age (Years)</td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>61.8 (5.0)</td>
</tr>
<tr>
<td>Baseline PSA (ng/ml)</td>
<td></td>
</tr>
<tr>
<td>Median (IQR)</td>
<td>4.1(3.5-5.5)</td>
</tr>
<tr>
<td>Number of PSA tests (n)</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td>2</td>
<td>118</td>
</tr>
<tr>
<td>3</td>
<td>193</td>
</tr>
<tr>
<td>&gt;4</td>
<td>280</td>
</tr>
<tr>
<td>Number of Biopsy (n)</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>351</td>
</tr>
<tr>
<td>2*</td>
<td>177</td>
</tr>
<tr>
<td>3*</td>
<td>55</td>
</tr>
<tr>
<td>4*</td>
<td>14</td>
</tr>
<tr>
<td>5*</td>
<td>2</td>
</tr>
<tr>
<td>Follow-up (Months)</td>
<td></td>
</tr>
<tr>
<td>Median (IQR)</td>
<td>17.0 (10.0-45.0)</td>
</tr>
</tbody>
</table>

**Table 7.** Baseline characteristics of men with an initial negative prostate biopsy.

In total, 599 of the 920 men biopsied within Sheffield had an initial negative biopsy. At a median follow-up was 17 months, 248 men underwent at least one rPBx, of which two men received a 4\textsuperscript{th} rPBx.

2-5*, Repeat prostate biopsies.
Table 8. Baseline characteristics of patients and prostate biopsy outcomes in each biopsy group.

Age, PSA details and histopathological results from prostate biopsies are detailed here. The baseline %free PSA (p=0.009), PSA (p<0.001) and PSA velocity (0.05) at biopsy were significantly different in the 2nd-4th rPBx group compared to the initial PBx/1st rPBx group.
Figure 15. Percentage of prostate cancer identified on prostate biopsy.

A total of 321/920 (34.9%) PCa was detected on initial PBx. Prostate cancer was detected in 66/337 (biopsy yield, 19.6%) rPBx with 41/248 (16.5%) and 25/89 (28.1%) PCa detected on 1st and 2nd – 4th rPBx respectively.
4.3.3 Prostate cancer identified by prostate biopsy

Baseline clinical features (age, PSA, time to diagnosis) and biopsy characteristics of positive (PCa) biopsies are shown in Table 9. The median (IQR) time to diagnosis on rPBx in the 2nd-4th rPBx group was 31 (11.5-57.5) months. The median (IQR) aggregate tumour length (mm) was 7.5 (2-26), 4 (1.5-10), 4 (1.5-11) mm (p=0.001) in the initial PBx, 1st rPBx and 2nd-4th rPBx groups respectively. Out of all tumours detected on initial and rPBx, 93/387 (24.0%) high-grade (Gleason 7-10) tumours were detected. In the first set of biopsies 78/321 (24.3%) were high-grade tumours compared to 15/66 (22.7%) of tumours detected on rPBx (p=0.88). Of all high-grade tumours 83.9% (78/93), 8.6% (8/93) and 7.5% (7/93) were detected on initial PBx, 1st rPBx and 2nd-4th rPBx respectively. A total of 114/321 (35.5%) tumours found on initial PBx and 33/66 (50%) tumours found on rPBx were classed as indolent (p=0.04).

4.3.4 The risks of being diagnosed with cancer on repeat biopsy

The Kaplan-Meier estimated risk of being diagnosed with any grade PCa on rPBx was 8.8% at 2 years, 14.3% at 3 years, 15.4% at 4 years and 17.4% at 5 years (Figure 16a). The estimated risk of being diagnosed with high-grade Gleason 7-10 PCa was, 1.4% at 2 years, 3.2% at 3 years, 4.6% at 4 years and 5.5% at 5 years (Figure 16b).
### Characteristics PCa

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Initial PBx (n= 321/920, 34.9%)</th>
<th>1&lt;sup&gt;st&lt;/sup&gt; rPBx (n= 41/248, 16.5%)</th>
<th>2&lt;sup&gt;nd&lt;/sup&gt; to 4&lt;sup&gt;th&lt;/sup&gt; rPBx (n=25/89, 28.1%)</th>
<th>p-value (Initial Vs all rPBx)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Baseline Age (years)</strong> Mean (SD)</td>
<td>61.8 (4.9)</td>
<td>60.2 (4.6)</td>
<td>58.7 (4.9)</td>
<td>0.57</td>
</tr>
<tr>
<td><strong>Age at diagnosis (years)</strong> Mean (SD)</td>
<td>62.0 (4.9)</td>
<td>61.3 (4.6)</td>
<td>61.9 (5.4)</td>
<td>0.71</td>
</tr>
<tr>
<td><strong>Baseline PSA (ng/ml)</strong> Median (IQR)</td>
<td>4.4 (3.5-6.2)</td>
<td>5.1 (4.2-7.5)</td>
<td>5.4 (3.9-6.7)</td>
<td>0.007</td>
</tr>
<tr>
<td><strong>Baseline %free PSA</strong> Median (IQR)</td>
<td>13.0 (10.2-15.9)</td>
<td>8.3 (5.6-10.7)</td>
<td>0.006&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td><strong>PSA at diagnosis (ng/ml)</strong> Median (IQR)</td>
<td>4.1 (3.2-6.1)</td>
<td>5.8 (4.3-7.8)</td>
<td>6.2 (4.4-11.2)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>PSA velocity at diagnosis (ng/ml/y)</strong> Median (IQR)</td>
<td>0.8 (-0.5-2.4)</td>
<td>0.4 (0.05-2.7)</td>
<td>0.91&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td><strong>Time to diagnosis from 1&lt;sup&gt;st&lt;/sup&gt; Bx (months)</strong> Median (IQR)</td>
<td>4 (2.5-21)</td>
<td>31 (11.5-57.5)</td>
<td>&lt;0.001&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
</tr>
</tbody>
</table>

#### Type of Biopsy, n (%)

<table>
<thead>
<tr>
<th>Type of Biopsy</th>
<th>Initial PBx</th>
<th>1&lt;sup&gt;st&lt;/sup&gt; rPBx</th>
<th>2&lt;sup&gt;nd&lt;/sup&gt; to 4&lt;sup&gt;th&lt;/sup&gt; rPBx</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRUS</td>
<td>321 (100)</td>
<td>41 (100)</td>
<td>10 (40)</td>
</tr>
<tr>
<td>Saturation</td>
<td>0</td>
<td>0</td>
<td>5 (20)</td>
</tr>
<tr>
<td>Saturation + TURP</td>
<td>0</td>
<td>0</td>
<td>10 (40)</td>
</tr>
</tbody>
</table>

#### Laterality, n (%)

<table>
<thead>
<tr>
<th>Laterality</th>
<th>Initial PBx</th>
<th>1&lt;sup&gt;st&lt;/sup&gt; rPBx</th>
<th>2&lt;sup&gt;nd&lt;/sup&gt; to 4&lt;sup&gt;th&lt;/sup&gt; rPBx</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unilateral</td>
<td>187 (58.3)</td>
<td>29 (70.7)</td>
<td>15 (60)</td>
</tr>
<tr>
<td>Bilateral</td>
<td>131 (40.8)</td>
<td>10 (24.4)</td>
<td>8 (32)</td>
</tr>
<tr>
<td>Unknown</td>
<td>3 (0.9)</td>
<td>2 (4.9)</td>
<td>2 (8)</td>
</tr>
</tbody>
</table>
Clinical parameters and detailed histopathological data of men who underwent initial and rPBx are shown. PSA is higher in men who underwent rPBx. Tumours found on rPBx appear to be smaller and indolent compared with tumours detected on initial PBx.

Table 9. Patient and prostate biopsy characteristics in positive cohorts.

Comparison between 1st rPBx and 2nd-4th rPBx; *indolent PCa (Epstein): Gleason score ≤6; <3cores positive; ≤50% positive per core.
### Table

<table>
<thead>
<tr>
<th>Time (months)</th>
<th>0</th>
<th>12</th>
<th>24</th>
<th>36</th>
<th>48</th>
<th>60</th>
<th>72</th>
<th>84</th>
<th>96</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number at risk</td>
<td>599</td>
<td>443</td>
<td>267</td>
<td>190</td>
<td>148</td>
<td>103</td>
<td>82</td>
<td>35</td>
<td>16</td>
</tr>
</tbody>
</table>

| % PCa-free    | 100 | 93.8| 91.2| 85.7| 84.6| 82.6| 78.8| 75.3| 70.8|
Figure 16. Kaplan-Meier estimate of time from date of first prostate cancer clinic to diagnosis of a) all cancer and b) high-grade cancer.

a) The Kaplan-Meier estimated risk of being diagnosed with any grade PCa on rPBx was 8.8% at 2 years and 17.4% at 5 years; b) The estimated risk of being diagnosed with high-grade Gleason 7-10 PCa was 1.4% at 2 years and 5.5% at 5 years.
4.3.5 Comparison between men with and without cancer
The median (IQR) time to diagnosis was 14 (5.8-33.3) months. The mean (SD) baseline age in the cancer group was 60.7 (4.9) years and 61.9 (5.0) years in the non-PCa group (p=0.06). The median (IQR) baseline PSA and median (IQR) PSA velocity was 5.1 (3.9-6.4) ng/ml and 4.1 (3.4-5.4) ng/ml (p<0.001), and 0.76 (-0.2-2.3) and -0.1 (-1.1-0.3) ng/ml/year (p<0.001) in the PCa and non-PCa group respectively. The percentage of ASAP on previous negative PBx in the PCa and non-PCa group was 13.6% and 4.1% (p=0.001) respectively (Table 10).

4.3.6 Uni- and Multivariable analysis of predictors
Uni- and multivariable analyses were used to analyse dependent variables. Both demonstrated that baseline PSA and %free PSA (<12%), PSA velocity and ASAP detected on previous PBx are positive predictors of PCa (Table 11).

4.3.7 Review of the international data
A literature review of the international data on rPBx was performed. Both retrospective and prospective studies reporting data on standard TRUS, extended and saturation PBx were included. A total of 20 studies were identified between 1994 and 2013. The diagnostic rate of PCa on rPBx is between 7.45 and 26.2% (Table 12).
Table 10. Clinical and pathological characteristics of patients with and without prostate cancer at the time of analysis.

Of the 599 who had an initial negative PBx, 66 men were found to have PCa on rPBx. Baseline PSA and PSA velocity were higher, and %free PSA was lower in the PCa group. There were more ASAP detected in the no PCa group.
Table 11. Univariable and multivariable analysis of predictors for tumours detected on repeat biopsy.

Univariable and multivariable analyses were performed to identify predictors of PCa on rPBx. Analyses revealed that PSA, PSA velocity, %free PSA, ASAP were associated with PCa detected on rPBx.

HR, Hazard ratio; CI, Confidence interval
<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Country</th>
<th>Study type</th>
<th>Specimen</th>
<th>Protocol</th>
<th>rPBx (n)</th>
<th>PCA on rPBx (n)</th>
<th>Initial PBx</th>
<th>1st rPBx</th>
<th>2nd rPBx</th>
<th>3rd rPBx</th>
<th>&gt;4 rPBx</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>Current</td>
<td>2014</td>
<td>England</td>
<td>Pros</td>
<td>PBx</td>
<td>10-core (5-33)</td>
<td>337</td>
<td>66 (19.6%)</td>
<td>321/920 (34.9%)</td>
<td>41/248 (16.5%)</td>
<td>17/71 (23.9%)</td>
<td>7/16 (43.8%)</td>
<td>1/2 (50%)</td>
<td>(Elshafei et al., 2013)</td>
</tr>
<tr>
<td>Elshafei</td>
<td>2013</td>
<td>USA</td>
<td>Retro</td>
<td>PBx</td>
<td></td>
<td>682</td>
<td>179 (26.2%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(Elshafei et al., 2013)</td>
</tr>
<tr>
<td>Ploussard</td>
<td>2013</td>
<td>France</td>
<td>Pros</td>
<td>PBx</td>
<td>Extended 21-core</td>
<td>847</td>
<td>139 (16.4%)</td>
<td>103/617 (16.7%)</td>
<td>28/166 (16.9%)</td>
<td>6/48 (12.5%)</td>
<td>2/16 (12.5%)</td>
<td>(Ploussard et al., 2013)</td>
<td></td>
</tr>
<tr>
<td>Cussenot</td>
<td>2013</td>
<td>France</td>
<td>Retro</td>
<td>PBx</td>
<td></td>
<td>176</td>
<td>176 (26%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(Cussenot et al., 2013)</td>
</tr>
<tr>
<td>Park B</td>
<td>2013</td>
<td>Korea</td>
<td>Retro</td>
<td>PBx</td>
<td>10-12-core</td>
<td>1180</td>
<td>190 (16.1%)</td>
<td>1956/7191 (27.2%)</td>
<td>142/976 (14.5%)</td>
<td>38/174 (21.8%)</td>
<td>9/27 (33.3%)</td>
<td>1/3 (33.3%)</td>
<td>(Park et al., 2013)</td>
</tr>
<tr>
<td>Bakardzhiev</td>
<td>2012</td>
<td>Bulgaria, Germany</td>
<td>Retro</td>
<td>PBx</td>
<td></td>
<td>166</td>
<td>22 (13.3%)</td>
<td>18/113 (15.9%)</td>
<td>3/35 (8.6%)</td>
<td>1/18 (5.5%)</td>
<td></td>
<td></td>
<td>(Bakardzhiev et al., 2012)</td>
</tr>
<tr>
<td>Najari</td>
<td>2012</td>
<td>Germany, USA</td>
<td>Retro</td>
<td>PBx</td>
<td>&gt;10-core</td>
<td>764</td>
<td>199 (26.1%)</td>
<td>3671/8729 (54.6%)</td>
<td>199/764 (26.1%)</td>
<td></td>
<td></td>
<td></td>
<td>(Najari et al., 2012)</td>
</tr>
<tr>
<td>Resnick</td>
<td>2011</td>
<td>USA</td>
<td>Retro</td>
<td>RP</td>
<td></td>
<td>456</td>
<td>1887</td>
<td>1867</td>
<td>281</td>
<td>175</td>
<td></td>
<td></td>
<td>(Resnick et al., 2011)</td>
</tr>
<tr>
<td>Quinlan</td>
<td>2009</td>
<td>Republic of Ireland</td>
<td>Retro</td>
<td>PBx</td>
<td>10-core</td>
<td>175</td>
<td>27 (15.4%)</td>
<td></td>
<td>16</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>(Quinlan et al., 2009)</td>
</tr>
<tr>
<td>Leite</td>
<td>2008</td>
<td>Brazil</td>
<td>Retro</td>
<td>PBx</td>
<td>Sextant extended 9-32</td>
<td>99</td>
<td>8 (8.1%)</td>
<td>524/1177 (44.5%)</td>
<td>6/76 (7.9%)</td>
<td>1/7 (5.9%)</td>
<td>1/5 (20%)</td>
<td>0/1 (0%)</td>
<td>(Leite et al., 2008)</td>
</tr>
<tr>
<td>Tan</td>
<td>2008</td>
<td>USA</td>
<td>Retro</td>
<td>PBx</td>
<td>Standard &lt;20 and saturation &gt;20</td>
<td>966</td>
<td>215 (22.3%)</td>
<td>690/1212 (56.9%)</td>
<td>142/621 (22.9%)</td>
<td>73/345 (21.2%)</td>
<td></td>
<td></td>
<td>(Tan et al., 2008)</td>
</tr>
</tbody>
</table>
Table 12. Published international data on repeat prostate biopsy outcomes.

A literature review was performed on prostate biopsies. A total of 20 studies were identified between 1994 and 2013. The diagnostic rate of PCa on rPBx is between 7.45 and 26.2%.

Pros, Prospective; Retro, Retrospective; (r)PBx, (repeat) prostate biopsy; RP, Retropubic Prostatectomy; PCa, prostate cancer
4.4 Discussion

4.4.1 Diagnostic rate of prostate cancer on repeat biopsies
The rPBx outcomes in UK men within the ProtecT study in a region without a screening programme was analysed. A total of 66 of 599 (11.0% risk) men with an elevated PSA between 3 and 19.9 ng/mL and an initial negative PBx were found to have PCa during the study period (median 17; IQR 10-45 months). The yield of PCa on subsequent rPBx (66/337, 19.6%) was found to be similar to previously reported data, 7.45-26.2% (Table 12) from the international literature. Population movement within South Yorkshire is limited and men tend to stay with the same practice. Although not all men were biopsied, giving a risk of verification bias, no clinical cases of PCa presented in the unbiopsied cohort during the study period.

Over-performing rPBx increases healthcare costs, places patients at risk of rPBx complications, and increases the risk of diagnosing and over-treating insignificant disease. It is known that the majority of PCa is detected on initial PBx, this was also evident in the current study (321/387, 83.0%). Therefore, optimisation of the first PBx would be highly valuable, including the use of extended PBx (Roehl et al., 2002; Hong et al., 2004; Eskicorapci et al., 2007), taking additional anterior apical cores (Wright et al., 2006), and saturation PBx (sPBx). The optimal rPBx protocol is unclear, several series have shown that sPBx and TURP enhances PCa detection on subsequent rPBx (Scattoni et al., 2007, 2010; Ploussard et al., 2009; Zaytoun, Moussa, et al., 2011). In the current study, 60% of PCa detected on 2nd-4th rPBx were diagnosed on saturation rPBx +/- TURP. Transperineal template-guided biopsy (TPM-Bx) may detect additional PCa, but this procedure requires general/regional anaesthesia and can result in additional complications, such as urinary retention (Symons et al., 2013).

A more contemporaneous approach is the use of mpMRI scanning prior to a decision for biopsy (Moore et al., 2013). Whereas several centres have reported promising results with such an approach in terms of reducing the
number of negative biopsies (Lawrentschuk et al., 2009; Hoeks et al., 2012), prospective data on the standard diagnostic criteria, risks and benefits, as well as the cost-effectiveness are limited (Hamoen et al., 2014). In the setting of an initial negative biopsy, mpMRI-guided rPBx (mpMRI-Bx) requires fewer cores and the initial PCa detection rates are similar to TRUS-PBx (Moore et al., 2013). A recent meta-regression analysis comparing PCa detection using TRUS-PBx (30.0%), TPM-Bx (36.8%) and mpMRI-Bx (37.6%) in a re-biopsy setting could not define which strategy offers the highest cancer detection rate, however MRI-PBx may potentially detect more PCa and the authors of the meta-analysis concluded that more well-designed prospective comparative studies with standardised outcome measures are required to define an optimum rPBx strategy (Nelson et al., 2013). The national UK PROMIS study (PROstate MRI Imaging Study) was a multicentre, paired-cohort study that tested diagnostic accuracy of mpMRI and TRUS-PBx against TPM-Bx as the reference standard. A total of 576 men underwent all three diagnostic tests and concluded that 27% of men may avoid a primary PBx and diagnosis of 5% fewer clinically insignificant PCa (reducing over-diagnosis of clinically insignificant PCa). If subsequent TRUS-PBx were indicated by mpMRI findings, up to 18% more cases of clinically significant PCa may be detected compared with the standard pathways of TRUS-PBx (PROMIS, 2012; El-Shater Bosaily et al., 2015; Ahmed et al., 2017).

4.4.2 Grades of prostate cancer
In keeping with other reports (Djavan et al., 2001), the majority of PCa identified on rPBx in this study were graded Gleason 6 (77.3% of all PCa). Results showed that tumours identified on rPBx tend to be smaller, similarly differentiated and less significant (Epstein criteria) (Epstein et al., 1994) when compared to tumours found on initial PBx. PSA also appeared to be higher in the rPBx group indicating its use in stratifying men for rPBx. Although the risk of having high-grade disease on rPBx is low (15/337, 4.5%), these poorly-differentiated tumours could result in poor outcomes if not detected and acted upon. On comparing histology from RP specimens, a
delay in diagnosis does not appear to alter tumour stage, but a larger proportion are high-grade, which contrasts with the findings on biopsy. The reason for this paradox is not immediately apparent.

Our Kaplan-Meier analyses provide evidence for the estimated risk of PCa being diagnosed on rPBx on a man presenting with an initial negative rPBx and elevated PSA- such data will be useful in counselling men considering biopsy for an elevated PSA, particularly in the setting of an initial negative biopsy.

4.4.3 Predictors of prostate cancer
In the current observation, it appears that rPBx were performed on patients with a higher PSA and PSA velocity, and a lower %free PSA. Patients appeared to be younger (baseline age) in the rPBx group, probably indicating an age-bias, however, this did not reach statistical significance.

Positive predictors of PCa on rPBx identified on multivariable analysis include baseline PSA and %free PSA, PSA velocity and ASAP detected on previous PBx. These predictors are amongst the list reported in the literature (Catalona et al., 1997; Djavan et al., 2000; Gann et al., 2010). Other predictors described include age, family history, DRE, PSA slope, HGPIN, prostate volume and prostate cancer antigen-3 (PCA3) (Zaytoun and Jones, 2011). Despite all the data available on PCa detection rates and predictors of PCa on subsequent rPBx, unfortunately, there is currently still no agreed protocol on rPBx. However, evolving data suggest that performing MRI prior to biopsy may avoid potentially unnecessary biopsies (Nelson et al., 2013; Ahmed et al., 2017). Knowing around 70% of men who undergo initial or rPBx have a negative result, other molecular strategies in addition to diagnostic imaging are needed. The clinical data collected in the current analysis assisted in identify men and specimens for subsequent laboratory analyses in Chapters 4 and 5.
4.4.4 Limitations
Men with features set out in the inclusion criteria were recruited for PSA testing via written invitation sent out by GPs, therefore non-responders were not included. Data in the current study may only apply to such selected men and not include men seen in routine UK clinical practice with clinically suspected PCa (urinary symptoms, suspicious DRE). The decision for rPBx was based on PSA and clinicians’ interpretation of the whole scenario as guidelines on >1 rPBx are not available. This represents a degree of verification bias, and as demonstrated patients with high PSA and high PSA velocity were selected for rPBx. Although no further cases of clinical PCa were reported in the region in the unbiopsied cohort, it is unknown whether they may have received further rPBx or were diagnosed with PCa in a different trust or region. There could therefore be an underestimate of the rates of PCa diagnosis.

4.4.5 Generalisability
The current study included large number of UK men recruited for the ProtecT study. The results are likely to be generalizable to UK men aged between 50 and 69 years undergoing TRUS-PBx for the first time as a result of a first-time PSA level between 3.0 and 19.9ng/ml, rather than men with clinically suspected PCa.

4.5 Conclusions
Around 1 in 10 men with an initial negative PBx are diagnosed with PCa on rPBx, with 2.5% found to have high-grade disease. Biopsy characteristics of cancers identified on rPBx suggested smaller, similarly differentiated tumours compared to tumours detected on initial PBx. These data are useful to quote when counselling patients with a persistently elevated PSA level and contemplating rPBx. Recent data on mpMRI and PBx appear promising and this modality may soon be adopted wide-spread in the UK to better select men for rPBx.
CHAPTER 4: The Identification and Role of PCA3-shRNA2 In Prostate Cancer
3.1 Background
As discussed in Chapter 1.3.7.3, PCA3 is a FDA approved long ncRNA biomarker used in conjunction with PSA (PCA3 score- PCA3/PSA mRNA ratio) to guide rPBx decisions (Hessels et al., 2003; Clarke et al., 2009). However, the limitations discussed prevent its global use including in the UK NHS. It is known that long ncRNAs have few exons, can be processed into short active RNAs (Röther et al., 2011), and are not conserved from primitive species. As no functional role for PCA3 has been investigated at the time of our study, we questioned whether PCA3 may encode a shorter active RNA that targets mRNA. To test this hypothesis, we searched for probable short ncRNAs derived from sequences from the BMCC1 transcript, which spans PCA3, and investigated their translational role.

3.2 Methods

3.2.1 Identification of hairpin RNA structures
Together with Dr Ross Drayton (post doc in the Catto lab) I searched BMCC1 gene for predicted RNA hairpins using Probabilistic miRNA prediction (ProMir II) (Nam et al., 2006) and miRNA Predictor (MiPred) (Jiang et al., 2007) bioinformatic algorithms. These bioinformatic online programs compare random sequences within the target hairpin. A small RNA transcriptome generated from malignant prostatic tissue using deep sequencing (Martens-Uzunova et al., 2012) was searched for RNA sequences derived from these predicted hairpins. The expression of any identified short RNA hairpins was measured using custom stem loop primers (TaqMan small RNA assays, Applied Biosystems) with qRT-PCR in cell lines, human urinary and prostate tissue samples.
3.2.2 Cell lines and androgen regulation of RNA
The following cell lines (purchased from ATCC) were used:

- Prostate- DU145, LNCaP, LNCaP-LN3, LNCaP-pro5, PC2, PC3M, PC3M-LN4
- Lung- A549, NCI-H460
- Endometrial- AN3CA
- Bladder- EJ/T24, RT112, RT4
- Colorectal- HCT-116
- Human embryonic kidney- HEK293
- Cervical- HeLa
- T-cell lymphoma- Jurkat
- Breast- MCF-7, T47D
- Lung fibroblasts- MRC5
- Ovarian- SKOV-3
- Melanoma- WM793

*Performed by Dr Ross Drayton*

For the androgen regulation experiment, LNCaP cells (androgen-dependent cells) were grown in androgen-depleted media (phenol red free RPMI-1640 + 10% charcoal stripped fetal bovine serum (FBS) (Sigma)) with no (0nmol/L), 1nmol/L and 10nmol/L of testosterone (Sigma). Total RNA was extracted and PSA, PCA3, and PCA3-shRNA2 were measured using qRT-PCR with appropriate primers (Clarke et al., 2009).

3.2.3 Expression of PCA3-shRNA in urinary samples

The expression of short RNAs in urinary samples was measured by qRT-PCR. RNA from exfoliated prostatic urinary cells was collected following prostatic massage in men with PCa and matching controls within a pilot and a validation cohort from the University of Sheffield. Controls were matched for age and PSA, and selected if they had undergone two or more negative PBx. Following prostatic massage, 10-20ml of urine was collected and centrifuged. The cell pellet was then washed in PBS before storage at -80°C.
3.2.4 RNA extraction and quantification (cell lines and urine)
Total RNA was extracted from cell lines and patient urinary samples and were measured using qRT-PCR (Clarke et al., 2009; Dudziec et al., 2011). Primers for BMCC1, PCA3 and PSA were purchased from Applied Biosystems, Life Technologies. PCA3-shRNA primers (Sigma) were designed using sequences obtained from previous bioinformatics analysis.

- Sequence A: ACTGCACTCCAGCCTGGGCA
- Sequence B: CACTGCACTCCAGCCTGGGCA

(Ambion: assay IDs, SCSGJ090 and CSHSNF8 respectively).

Expression of PCA3, PCA3-shRNA and BMCC1 was normalized to PSA and fold changes calculated using Delta cycle threshold (Dct) values (Catto et al., 2009). In brief, the Dct value was calculated from subtracting the control Ct value (PSA) from the test Ct value (i.e PCA3-shRNA). Delta Dct (DDct) was the difference between the test Dct (cancer) and control Dct (benign). Fold change was calculated using –log2 of DDct (Livak et al., 2001). For RNA localization studies, nuclear and cytoplasmic RNA fractions were extracted separately (Stuart et al., 2004). The expressions of the potential mRNA targets were measured in cell lines and patient urinary samples using qRT-PCR (Dudziec et al., 2011).

3.2.5 Cloning primary transcripts using 3'RACE
*Performed by Dr Ross Drayton*
To determine the sequence and genomic origin of the primary RNA transcript producing short-PCA3, 3’ Rapid Amplification of cDNA ends (3’ RACE) in PC3 cells using the GeneRacer Kit according to the manufacturer’s guidelines (Life Technologies) was performed. Total RNA was extracted from PC3 cells and cDNA synthesized by PCR using an adapter primer (AP) that targets the poly(A) tail of mRNA. Amplification of cDNA by PCR was performed using PCA3-shRNA primers.
The target sequences were cloned into E. coli (Top10; Left Technologies) and then extracted, purified and sent for Sanger sequencing. Sequences were aligned (Sequencher 5.1, Gene Codes), and genomic matches were identified using Basic Local Alignment Search Tool (BLAST, NCBI) (Altschul et al., 1990).

3.2.6 mRNA target analysis and knock-up
TargetScan (v4.2) was used to identify putative target mRNAs with complementary sequences to the short-PCA3. Potential targets were analysed for PCa relevance by cross-referencing with aberrantly expressed genes obtained from publicly available microarray datasets (Arrayexpress ID: E-GEOD-8218 (Wang et al., 2010)) Cellular functions and pathway enrichment for these mRNAs were analysed using Database for Annotation, Visualization and Integrated Discovery (DAVID) v6.7 (Huang et al., 2009).

Primers were designed for selected targets, and expression measured using qRT-PCR in LNCaP cells following PCA3-shRNA knock-up. Knockup was performed by transfecting LNCaP cells with a custom made hairpin precursor designed to generate PCA3-shRNA using Lipofectamine RNAiMAX (Catto et al., 2009). Transfection with scrambled RNA was used as a control. Success of transfection was obtained by measuring PCA3-shRNA expression using qRT-PCR.

3.3 Results

3.3.1 Identification of PCA3-shRNA
A total of 13 potential RNA hairpins (Table 13) were identified following an *in silico* analysis of the BMCC1 locus. Each hairpin was derived from sequence within a BMCC1 intron and most were located around the PCA3 locus. A search of the prostate transcriptome (Martens-Uzunova et al., 2012) identified five of these RNAs, including RNA2 (second in the list), which we termed PCA3-shRNA2 (short RNA number 2), Table 13. This RNA accounted for 72 of 79 (91%) of hits. PCA3-shRNA2 was identified to be
located within PCA3 gene intron 1, adjacent to a region of high species conservation (Figure 17).

Alignment of the transcriptomic sequences to the genome revealed two potential 5' start sequences for PCA3-shRNA2; ACUG and a minority member starting with CACUG (Figure 18). We designed TaqMan assays to each (given that the 5' end of short RNA is vital for mRNA targeting) and named these assays PCA3-shRNA2a and PCA3-shRNA2b, respectively (Bioinformatic identification of PCA3-shRNA was performed together with Dr Ross Drayton).

3.3.2 Determination of the genomic origin of PCA3-shRNA
These data support our transcriptomic analysis, however, do not prove that our short RNA is derived from sequence within the PCA3 intron. To analyse this, 3' RACE to clone the primary sequence from the PCA3-shRNA2 primer was performed (performed by Dr Ross Drayton). Sequenced RACE products aligned to the PCA3 intronic locus and supported our in-silico prediction of a 98-bp hairpin (Figure 19). A BLAST search of the 98-bp sequence revealed strong homology (97%) for one locus in the genome- within PCA3 intron 1 (Figure 20).
ProMir and MiPred were used to search for predicted RNA hairpins from BMCC1. A total of 13 hairpins were found, 10 were predicted to be real, whilst three were pseudo- hairpin sequences with similar stem-loop features. Short RNA number 2 was found within intron 6 of the BMCC1 gene. The short RNA accounted for 72/79 (91.1%) of hits when the sequence was searched in a prostate cancer transcriptome. (Drayton et al., 2015).

<table>
<thead>
<tr>
<th>End (bp)</th>
<th>BMCC1 Region</th>
<th>Note</th>
<th>Free energy</th>
<th>GC ratio</th>
<th>Entropy</th>
<th>Promir value</th>
<th>MiPred Result</th>
<th>MiPred Confidence</th>
<th>Freq in Transcriptomic data</th>
<th>% of hits</th>
</tr>
</thead>
<tbody>
<tr>
<td>31,365</td>
<td>Intron 1</td>
<td></td>
<td>-25.5</td>
<td>0.38</td>
<td>1.9508</td>
<td>0.0489</td>
<td>Pseudo</td>
<td>68.70%</td>
<td>1</td>
<td>1.3%</td>
</tr>
<tr>
<td>93,278</td>
<td>Intron 6 (downstream PCA3)</td>
<td>Repeat, first 75 bp is AluSx (SINE/Alu)</td>
<td>-30.8</td>
<td>0.45</td>
<td>1.97326</td>
<td>0.0451</td>
<td>Real</td>
<td>54.00%</td>
<td>72</td>
<td>91.1%</td>
</tr>
<tr>
<td>100,952</td>
<td>Intron 6 (downstream PCA3)</td>
<td>Repeat, MER5A (DNA/MER1_type)</td>
<td>-36.44</td>
<td>0.45</td>
<td>1.97809</td>
<td>0.0617</td>
<td>Real</td>
<td>66.90%</td>
<td>0</td>
<td>0.0%</td>
</tr>
<tr>
<td>105,132</td>
<td>Intron 6 (downstream PCA3)</td>
<td>Conserved in 4 mammals</td>
<td>-27</td>
<td>0.33</td>
<td>1.8797</td>
<td>0.0635</td>
<td>Real</td>
<td>74.20%</td>
<td>0</td>
<td>0.0%</td>
</tr>
<tr>
<td>115,526</td>
<td>Intron 6 (downstream PCA3)</td>
<td></td>
<td>-49.7</td>
<td>0.3</td>
<td>1.85017</td>
<td>5.0776</td>
<td>Real</td>
<td>74.20%</td>
<td>0</td>
<td>0.0%</td>
</tr>
<tr>
<td>163,666</td>
<td>Intron 6 (upstream PCA3)</td>
<td></td>
<td>-46.3</td>
<td>0.46</td>
<td>1.98194</td>
<td>1.5115</td>
<td>Real</td>
<td>68.60%</td>
<td>0</td>
<td>0.0%</td>
</tr>
<tr>
<td>176,673</td>
<td>Intron 6 (upstream PCA3)</td>
<td>Repeat, MER5A (DNA/MER1_type)</td>
<td>-37</td>
<td>0.46</td>
<td>1.99233</td>
<td>0.0559</td>
<td>Real</td>
<td>62.90%</td>
<td>0</td>
<td>0.0%</td>
</tr>
<tr>
<td>181,993</td>
<td>Intron 6 (upstream PCA3)</td>
<td>Conserved in human dog (see the alignment in the end of the file)</td>
<td>-28.1</td>
<td>0.33</td>
<td>1.89539</td>
<td>0.0742</td>
<td>Real</td>
<td>61.10%</td>
<td>3</td>
<td>3.8%</td>
</tr>
<tr>
<td>215,235</td>
<td>Intron 9</td>
<td></td>
<td>-30.9</td>
<td>0.51</td>
<td>1.99248</td>
<td>0.1165</td>
<td>Pseudo</td>
<td>50.80%</td>
<td>1</td>
<td>1.3%</td>
</tr>
<tr>
<td>236,277</td>
<td>Intron 9</td>
<td>Repeat, LIHS (LINE/L1)</td>
<td>-34.8</td>
<td>0.47</td>
<td>1.91696</td>
<td>0.0531</td>
<td>Real</td>
<td>62.30%</td>
<td>2</td>
<td>2.5%</td>
</tr>
<tr>
<td>247,401</td>
<td>Intron 9</td>
<td>Repeat, Tigger4a (DNA/MER2_type)</td>
<td>-39.2</td>
<td>0.38</td>
<td>1.95071</td>
<td>85.884</td>
<td>Real</td>
<td>75.80%</td>
<td>0</td>
<td>0.0%</td>
</tr>
<tr>
<td>254,347</td>
<td>Intron 10</td>
<td>Repeat, Tigger4a (DNA/MER2_type)</td>
<td>-29.03</td>
<td>0.36</td>
<td>1.94008</td>
<td>0.2239</td>
<td>Real</td>
<td>80.00%</td>
<td>0</td>
<td>0.0%</td>
</tr>
<tr>
<td>268,339</td>
<td>Intron 12</td>
<td></td>
<td>-27.85</td>
<td>0.39</td>
<td>1.9288</td>
<td>0.6003</td>
<td>Pseudo</td>
<td>51.20%</td>
<td>0</td>
<td>0.0%</td>
</tr>
</tbody>
</table>

**Table 13.** Predicted hairpin RNAs within PCA3 and BMCC1.
Figure 17. Identification of PCA3-shRNA2.

Our potential short-PCA3 is located in exon 6 of the BMCC1 gene. PCA3-shRNA2 is situated adjacent to a region of high species conservation within intron 1 of PCA3 (Drayton et al., 2015).
Figure 18. Alignment of the transcriptomic sequences to the genome.

The predicted short PCA3 RNA hairpin identified is shown. Prostate cancer RNA transcriptomic data identified a relative abundance of our predicted RNA. The bases in red are those identified within the prostate transcriptome. Alignment of the transcriptomic sequences to the genome revealed two potential 5’ start sequences. (Drayton et al., 2015).
Figure 19. The genomic origin of PCA3-shRNA2.

To determine the origin of our short RNA, 3' RACE was used to clone the primary sequence from the PCA3-shRNA2 primer. 3' RACE identified the longer hairpin structure in LNCaP cells. The predicted 98-bp PCA3-shRNA2 is highlighted in red (Drayton et al., 2015).
Figure 20. BLAST results of the PCA3-shRNA2 sequence.

BLAST results of the 98bp fragment derived from PCA3-shRNA2 using 3’RACE indicated that the sequence is found within PCA3 intron 1 (Drayton et al., 2015).
3.3.3 Expression of PCA3-shRNA

3.3.3.1 Cell lines

The expression of PCA3-shRNA2 was measured in 22 cell lines and expression was detectable in all 7 prostate and 15 other cancer cell lines (Figure 21). PCA3 and PCA3-shRNA2 expression did not vary significantly with organ of origin for these cell lines. RNA expression was normalized to PSA mRNA since the commercial PCA3 test uses PSA as a reference gene. A correlation between the expression of PCA3 and PCA3-shRNA2 (PCA3-shRNA2a assay: r=0.92; P<0.001 and PCA3-shRNA2b assay: r=0.92; P<0.001) was identified, which was closer than for BMCC1 (r=0.56, P=0.01) (Figure 22).

![Figure 21. Expression of PCA3 and PCA3-shRNA2 in cell lines representing prostate cancer and other malignancies](chart.png)

The expression of PCA3 and PCA3-shRNA2 is shown (normalized to PSA mRNA expression) for the 22 cell lines. Prostate cell lines are coloured in blue, and other malignancies in red. The non-prostate cancer cell lines are not labelled for clarity. In order of PCA3-shRNA2 expression these are (from HCT-116 (PCA3-shRNA2, DCT=-15.85), HEK 293, A549, NCI-H460, WM793, RT112, T47D, MRC5, AN3CA, RT4, SKOV-3, EJ, MCF-7, Jurkat and HeLa (PCA3-shRNA2, DCT=5.08)) (Drayton et al., 2015).
RNA was extracted from PCa cell lines and qRT-PCR was used to measure the expression of BMCC1, PCA3 and PCA3-shRNA2. Expression of PCA3-shRNA2 ($r=0.92$, $p<0.001$) is closely correlated with PCA3 and less so to BMCC1 (DCT values normalized to PSA expression shown) (Drayton et al., 2015).

**Figure 22.** The correlation between PCA3-shRNA2 and PCA3/BMCC1.
3.3.3.2 Urinary samples
The clinical use for PCA3 is to test for PCa using exfoliated prostatic urinary cells. To explore this function for PCA3-shRNA2, its expression in 179 post-DRE urinary samples (Table 14) from men with (n=129) and without PCa (n=50) was examined. Once again, there was a close correlation between PCA3 and PCA3-shRNA2 expression (Figure 23; r=0.84; P<0.001).

Overall, there was upregulation of PCA3 (86.2±53.1 fold change (mean±SD)), PCA3-shRNA2 (273±0.1) and BMCC1 (2.7±0.1) in specimens from men with cancer, when compared with controls (all t-test P<0.003, Figure 24).
Table 14. Clinical and histopathological characteristics of patients and samples.

A pilot (n=179) and validation (n=471) urinary cohort was used in the current analysis. The age and PSA are shown in both the benign and malignant groups. Gleason score and T-stage are shown in the malignant group (Drayton et al., 2015).

<table>
<thead>
<tr>
<th>Material</th>
<th>Urinary pilot cohort</th>
<th>Disaggregated urinary cells</th>
<th>Urinary validation cohort</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>179</td>
<td>471</td>
<td></td>
</tr>
<tr>
<td>Benign tissue/controls</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>50</td>
<td>116</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>68.3</td>
<td>66.2</td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>8.9</td>
<td>7.2</td>
<td></td>
</tr>
<tr>
<td>PSA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>7.8</td>
<td>6.2</td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>5.9</td>
<td>6.2</td>
<td></td>
</tr>
<tr>
<td>Prostate cancer</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>129</td>
<td>355</td>
<td></td>
</tr>
<tr>
<td>Stage</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pT1-2</td>
<td>34</td>
<td>279</td>
<td></td>
</tr>
<tr>
<td>pT3</td>
<td>30</td>
<td>61</td>
<td></td>
</tr>
<tr>
<td>Metastatic</td>
<td>49</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>7</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Gleason sum</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>54</td>
<td>191</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>29</td>
<td>125</td>
<td></td>
</tr>
<tr>
<td>8–10</td>
<td>21</td>
<td>31</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>14</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Age, years</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td></td>
<td>65.1</td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td></td>
<td>7.0</td>
<td></td>
</tr>
<tr>
<td>PSA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>7.5</td>
<td>7.4</td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>175.1</td>
<td>418.9</td>
<td></td>
</tr>
<tr>
<td>0–10</td>
<td>79</td>
<td>248</td>
<td></td>
</tr>
<tr>
<td>11–20</td>
<td>27</td>
<td>64</td>
<td></td>
</tr>
<tr>
<td>&gt;20</td>
<td>16</td>
<td>37</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>0</td>
<td>9</td>
<td></td>
</tr>
</tbody>
</table>
Figure 23. The correlation between PCA3-shRNA2 and PCA3 in urinary RNA.

RNA was extracted from the pilot (n=179) cohort of urinary samples obtained from patients with (n=129) and without (n=50) PCa. qRT-PCR was used to measure the expression of PCA3 and PCA3-shRNA2. The expression of PCA3-shRNA2 was closely correlated with PCA3 expression (r=0.84, p<0.001) (Drayton et al., 2015).
We showed that PCA3-shRNA2 expression correlated to that of PCA3. We compared the expression (qRT-PCR) of BMCC1, PCA3 and PCA3-shRNA2 in benign (n=50) and malignant (n=129) samples and found that the expression was higher in urinary pellets from men with PCa than in benign controls for PCA3, PCA3-shRNA2 and BMCC1 (P<0.003) (Drayton et al., 2015).

To explore the robustness of these findings, a separate larger validation cohort of 471 urinary samples was examined. Once again, PCA3-shRNA2 expression was correlated with PCA3 (r=0.60, p<0.01) (Figure 25). qRT-PCR revealed that PCA3-shRNA2 expression was higher in samples from men with prostate cancer than controls (13.0±2.8-fold upregulation (mean±SD) in malignant samples; t-test P<0.001; Figure 26). However, expression of PCA3-shRNA2 did not vary with tumour stage (Figure 27).
**Figure 25.** The correlation between PCA3-shRNA2 and PCA3 in a large urinary cohort.

The expression (qRT-PCR) of PCA3 and PCA3-shRNA2 was measured in a larger validation (n=471) cohort of urinary samples. Expression of PCA3-shRNA2 was correlated with PCA3 expression (r=0.60, P<0.01), supporting our pilot exploration findings (Drayton et al., 2015).
Figure 26. The expression of PCA3-shRNA2 in benign and cancerous urinary samples.

The expression (qRT-PCR) of PCA3 and PCA3-shRNA2 in benign (n=116) and malignant (n=355) samples from the larger validation (n=471) urinary cohort was compared. PCA3-shRNA2 is overexpressed in urinary samples obtained from men with PCa (Drayton et al., 2015).
Figure 27. The expression of PCA3-shRNA2 in different stages of disease.

The expression (qRT-PCR) of PCA3-shRNA2 was compared in different stages of PCa. Expression was higher in urinary pellets from men with PCa than in BPH controls, but did not vary with cancer stage (ANOVA, P=0.46 between stages) (Drayton et al., 2015).
3.3.4 The ability of PCA3-shRNA2 to identify disease (urine)

PCA3-shRNA2 allowed the identification of malignancy in most men. Concordance indices show that PCA3 (C-index 0.78) and PCA3-shRNA2 (C-index 0.75) had similar accuracy for PCa, and both were superior to BMCC1 (C-index 0.66) when analysing results from our pilot study (Figure 28a). On analysing the ability of PCA3-shRNA2 to identify disease within a larger validation cohort of urinary samples, results supported our pilot exploration outcomes (C-index 0.81, Figure 28b).

![Graph showing sensitivity and 1-specificity for PCA3, PCA3-shRNA2, and BMCC1](image)
b) Receiver Operating Characteristic (ROC) curve analysis was used to explore the diagnostic performance of PCA3-shRNA. a) In comparison, PCA3 and PCA3-shRNA2 expressions were more reliable (C-indices 0.78, 0.75, respectively) than BMCC1 (C-index 0.66) at identify the presence of the cancer; b) PCA3-shRNA2 expression could identify the presence of prostate cancer in most men (C-index 0.81) in a large validation (n=471) urinary cohort (Drayton et al., 2015).

**Figure 28.** Identification of disease by PCA3-shRNA2.
3.3.5 Androgen regulation of PCA3-shRNA2

It is known that many RNAs important in prostate oncogenesis are regulated by androgens. In LNCaP cells (chosen for their androgen dependency), both PCA3 (2.1±0.31 fold change (mean±SD)) and PCA3-shRNA2 (2.75±0.23 fold change (mean±SD)) were upregulated in a dose-dependent manner (Figure 29) by testosterone. The changes were minimal for PSA (219.0±25.2 fold upregulation (mean±SD)).

Figure 29. Androgen regulated expression of PSA, PCA3 and PCA3-shRNA2.

LNCaP cells were cultured in androgen-depleted media with 0, 1 and 10nM of testosterone. PSA, PCA3 and PCA3-shRNA2 expression was measured using qRT-PCR. PCA3 and PCA3-shRNA2 expression was directly related to the concentration of testosterone (Drayton et al., 2015).
3.3.6 Functional role of PCA3-shRNA2
It is known that miRNA/short RNA function by binding to target mRNA in the cytoplasm. To explore this role, the localization of PCA3-shRNA2 was examined followed by identifying target mRNAs. The expression of target mRNAs was measured in PCa cell lines and patient urinary samples.

3.3.6.1 Localization of PCA3-shRNA2
qRT-PCR of total and nuclear fractions revealed a cytoplasmic enrichment (nuclear:cytoplasmic ratio=0.6) for PCA3-shRNA2, close to that seen for established miRNAs (Figure 30), and very different from PCA3 (with its mostly nuclear localization). This suggests a potential mRNA targeting capacity within the cytoplasm.

3.3.6.2 Identifying potential mRNA targets
The genome for complementary sequences was searched. Using TargetScan, 178 mRNAs with complementary seed sequences (Table 15) were identified. Gene enrichment analysis (DAVID) revealed significant associations (P<0.05) with pathways important in cell regulation (i.e. cell adhesion and growth, and cell signalling) and prostate biology (i.e. response to steroids, TGF-β signalling, and uro-genital development).
Table 15. Predicted mRNA targets for PCA3-shRNA2.

TargetScan revealed 178 mRNA targets with complementary sequence with PCA3-shRNA2. Gene ID and Gene name are displayed in this table.

<table>
<thead>
<tr>
<th>TargetScan ranking</th>
<th>Gene ID</th>
<th>Gene name</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>KIAA0515</td>
<td>KIAA0515</td>
</tr>
<tr>
<td>2</td>
<td>COPS2</td>
<td>COP9 constitutive photomorphogenic homolog subunit 2</td>
</tr>
<tr>
<td>3</td>
<td>AFF3</td>
<td>AF4/FMR2 family, member 3</td>
</tr>
<tr>
<td>4</td>
<td>SFRS2</td>
<td>splicing factor, arginine/serine-rich 2</td>
</tr>
<tr>
<td>5</td>
<td>TEAD1</td>
<td>TEA domain family member 1 (SV40 transcriptional enhancer factor)</td>
</tr>
<tr>
<td>6</td>
<td>ANKHD1-EIF4EBP3</td>
<td>ANKHD1-EIF4EBP3</td>
</tr>
<tr>
<td>7</td>
<td>ANKR5D7</td>
<td>ankyrin repeat domain 57</td>
</tr>
<tr>
<td>8</td>
<td>BTBD3</td>
<td>BTB (POZ) domain containing 3</td>
</tr>
<tr>
<td>9</td>
<td>C13orf36</td>
<td>chromosome 13 open reading frame 36</td>
</tr>
<tr>
<td>10</td>
<td>C8orf33</td>
<td>chromosome 8 open reading frame 33</td>
</tr>
<tr>
<td>11</td>
<td>CDH2</td>
<td>cadherin 2, type 1, N-cadherin (neuronal)</td>
</tr>
<tr>
<td>12</td>
<td>CLDN22</td>
<td>claudin 22</td>
</tr>
<tr>
<td>13</td>
<td>DCAKD</td>
<td>dephospho-CoA kinase domain containing</td>
</tr>
<tr>
<td>14</td>
<td>EIF4EBP3</td>
<td>eukaryotic translation initiation factor 4E binding protein 3</td>
</tr>
<tr>
<td>15</td>
<td>ETV1</td>
<td>ets variant gene 1</td>
</tr>
<tr>
<td>16</td>
<td>ETV5</td>
<td>ets variant gene 5 (ets-related molecule)</td>
</tr>
<tr>
<td>17</td>
<td>FAM123A</td>
<td>family with sequence similarity 123A</td>
</tr>
<tr>
<td>18</td>
<td>FAM123B</td>
<td>family with sequence similarity 123B</td>
</tr>
<tr>
<td>19</td>
<td>FAM40B</td>
<td>family with sequence similarity 40, member B</td>
</tr>
<tr>
<td>20</td>
<td>FLJ20309</td>
<td>hypothetical protein FLJ20309</td>
</tr>
<tr>
<td>21</td>
<td>HDHD2</td>
<td>haloacid dehalogenase-like hydrolase domain containing 2</td>
</tr>
<tr>
<td>22</td>
<td>INVS</td>
<td>inversin</td>
</tr>
<tr>
<td>23</td>
<td>IRF2BP2</td>
<td>interferon regulatory factor 2 binding protein 2</td>
</tr>
<tr>
<td>24</td>
<td>LIN9</td>
<td>lin-9 homolog (C. elegans)</td>
</tr>
<tr>
<td>25</td>
<td>LMNB1</td>
<td>lamin B1</td>
</tr>
<tr>
<td>26</td>
<td>LSM11</td>
<td>LSM11, U7 small nuclear RNA associated</td>
</tr>
<tr>
<td>27</td>
<td>MAP3K1</td>
<td>mitogen-activated protein kinase kinase kinase 1</td>
</tr>
<tr>
<td>28</td>
<td>METTL8</td>
<td>methyltransferase like 8</td>
</tr>
<tr>
<td>29</td>
<td>MLLT6</td>
<td>myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog); translocated to, 6</td>
</tr>
<tr>
<td>30</td>
<td>NOG</td>
<td>noggin</td>
</tr>
<tr>
<td>31</td>
<td>ODZ4</td>
<td>odz, odd Oz/ten-m homolog 4 (Drosophila)</td>
</tr>
<tr>
<td>32</td>
<td>ORMDL1</td>
<td>ORM1-like 1 (S. cerevisiae)</td>
</tr>
<tr>
<td>33</td>
<td>P15RS</td>
<td>cyclin-dependent kinase 2B-inhibitor-related protein</td>
</tr>
<tr>
<td>34</td>
<td>PDS5A</td>
<td>PDS5, regulator of cohesion maintenance, homolog A (S. cerevisiae)</td>
</tr>
<tr>
<td>35</td>
<td>PHF21A</td>
<td>PHD finger protein 21A</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>36</td>
<td>PHKA2</td>
<td>phosphorylase kinase, alpha 2 (liver)</td>
</tr>
<tr>
<td>37</td>
<td>PIGA</td>
<td>phosphatidylinositol glycan anchor biosynthesis, class A (paroxysmal nocturnal hemoglobinuria)</td>
</tr>
<tr>
<td>38</td>
<td>PNMA1</td>
<td>paraneoplastic antigen MA1</td>
</tr>
<tr>
<td>39</td>
<td>PSCD8B</td>
<td>pleckstrin homology, Sec7 and coiled-coil domains, binding protein</td>
</tr>
<tr>
<td>40</td>
<td>PXMP4</td>
<td>peroxisomal membrane protein 4, 24kDa</td>
</tr>
<tr>
<td>41</td>
<td>RAPGEF2</td>
<td>Rap guanine nucleotide exchange factor (GEF) 2</td>
</tr>
<tr>
<td>42</td>
<td>RNF169</td>
<td>ring finger protein 169</td>
</tr>
<tr>
<td>43</td>
<td>SESN2</td>
<td>sestrin 2</td>
</tr>
<tr>
<td>44</td>
<td>SFNS12IP1</td>
<td>SFRS12-interacting protein 1</td>
</tr>
<tr>
<td>45</td>
<td>SFRS2B</td>
<td>splicing factor, arginine/serine-rich 2B</td>
</tr>
<tr>
<td>46</td>
<td>SFRS3</td>
<td>splicing factor, arginine/serine-rich 3</td>
</tr>
<tr>
<td>47</td>
<td>SOX11</td>
<td>SRY (sex determining region Y)-box 11</td>
</tr>
<tr>
<td>48</td>
<td>SSR1</td>
<td>signal sequence receptor, alpha (translocon-associated protein alpha)</td>
</tr>
<tr>
<td>49</td>
<td>TLR4</td>
<td>toll-like receptor 4</td>
</tr>
<tr>
<td>50</td>
<td>TYW3</td>
<td>tRNA-ylW synthesizing protein 3 homolog (S. cerevisiae)</td>
</tr>
<tr>
<td>51</td>
<td>WDR1</td>
<td>WD repeat domain 1</td>
</tr>
<tr>
<td>52</td>
<td>WDR48</td>
<td>WD repeat domain 48</td>
</tr>
<tr>
<td>53</td>
<td>XP07</td>
<td>exportin 7</td>
</tr>
<tr>
<td>54</td>
<td>ZC3H10</td>
<td>zinc finger CCCH-type containing 10</td>
</tr>
<tr>
<td>55</td>
<td>ZFAND6</td>
<td>zinc finger, AN1-type domain 6</td>
</tr>
<tr>
<td>56</td>
<td>HTR2C</td>
<td>5-hydroxytryptamine (serotonin) receptor 2C</td>
</tr>
<tr>
<td>57</td>
<td>LIN54</td>
<td>lin-54 homolog (C. elegans)</td>
</tr>
<tr>
<td>58</td>
<td>SPTBN1</td>
<td>spectrin, beta, non-erythrocytic 1</td>
</tr>
<tr>
<td>59</td>
<td>ABI1</td>
<td>abl-interactor 1</td>
</tr>
<tr>
<td>60</td>
<td>ADAM12</td>
<td>ADAM metalloepitidase domain 12 (meltrin alpha)</td>
</tr>
<tr>
<td>61</td>
<td>ANKH</td>
<td>ankylosis, progressive homolog (mouse)</td>
</tr>
<tr>
<td>62</td>
<td>APH1A</td>
<td>anterior pharynx defective 1 homolog A (C. elegans)</td>
</tr>
<tr>
<td>63</td>
<td>ARL5B</td>
<td>ADP-ribosylation factor-like 5B</td>
</tr>
<tr>
<td>64</td>
<td>BRPF1</td>
<td>bromodomain and PHD finger containing, 1</td>
</tr>
<tr>
<td>65</td>
<td>CCNL2</td>
<td>cyclin L2</td>
</tr>
<tr>
<td>66</td>
<td>CPEB2</td>
<td>cytoplasmic polyadenylation element binding protein 2</td>
</tr>
<tr>
<td>67</td>
<td>CROP</td>
<td>cisplatin resistance-associated overexpressed protein</td>
</tr>
<tr>
<td>68</td>
<td>CTBP2</td>
<td>C-terminal binding protein 2</td>
</tr>
<tr>
<td>69</td>
<td>DDX5</td>
<td>DEAD (Asp-Glu-Ala-Asp) box polypeptide 5</td>
</tr>
<tr>
<td>70</td>
<td>EHMT1</td>
<td>euchromatic histone-lysine N-methyltransferase 1</td>
</tr>
<tr>
<td>71</td>
<td>EIF3J</td>
<td>eukaryotic translation initiation factor 3, subunit J</td>
</tr>
<tr>
<td>72</td>
<td>EIF4G3</td>
<td>eukaryotic translation initiation factor 4 gamma, 3</td>
</tr>
<tr>
<td>73</td>
<td>EML1</td>
<td>echinoderm microtubule associated protein like 1</td>
</tr>
<tr>
<td>74</td>
<td>FOS</td>
<td>v-fos FBJ murine osteosarcoma viral oncogene homolog</td>
</tr>
<tr>
<td>75</td>
<td>GATA2A</td>
<td>GATA zinc finger domain containing 2A</td>
</tr>
<tr>
<td>76</td>
<td>H3F3B</td>
<td>H3 histone, family 3B (H3.3B)</td>
</tr>
<tr>
<td>77</td>
<td>HMG2B</td>
<td>high-mobility group box 2</td>
</tr>
<tr>
<td>78</td>
<td>HNRNP3</td>
<td>heterogeneous nuclear ribonucleoprotein A3</td>
</tr>
<tr>
<td>79</td>
<td>IKZF2</td>
<td>IKAROS family zinc finger 2 (Helios)</td>
</tr>
<tr>
<td>80</td>
<td>ITCH</td>
<td>itchy E3 ubiquitin protein ligase homolog (mouse)</td>
</tr>
<tr>
<td></td>
<td>Gene Name</td>
<td>Description</td>
</tr>
<tr>
<td>---</td>
<td>-------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>81</td>
<td>KCMF1</td>
<td>potassium channel modulatory factor 1</td>
</tr>
<tr>
<td>82</td>
<td>KCNA4</td>
<td>potassium voltage-gated channel, shaker-related subfamily, member 4</td>
</tr>
<tr>
<td>83</td>
<td>LHFPL4</td>
<td>lipoma HMGIC fusion partner-like 4</td>
</tr>
<tr>
<td>84</td>
<td>LMO7</td>
<td>LIM domain 7</td>
</tr>
<tr>
<td>85</td>
<td>LOC399947</td>
<td>similar to expressed sequence AI593442</td>
</tr>
<tr>
<td>86</td>
<td>MAGI2</td>
<td>membrane associ'd guanylate kinase, WW and PDZ domain containing 2</td>
</tr>
<tr>
<td>87</td>
<td>MAP4K4</td>
<td>mitogen-activated protein kinase 4</td>
</tr>
<tr>
<td>88</td>
<td>MBOAT1</td>
<td>membrane bound O-acyltransferase domain containing 1</td>
</tr>
<tr>
<td>89</td>
<td>MEIS1</td>
<td>Meis homeobox 1</td>
</tr>
<tr>
<td>90</td>
<td>MN1</td>
<td>meningioma (disrupted in balanced translocation) 1</td>
</tr>
<tr>
<td>91</td>
<td>MON2</td>
<td>MON2 homolog (S. cerevisiae)</td>
</tr>
<tr>
<td>92</td>
<td>NDN</td>
<td>necdin homolog (mouse)</td>
</tr>
<tr>
<td>93</td>
<td>NLK</td>
<td>nemo-like kinase</td>
</tr>
<tr>
<td>94</td>
<td>NMNAT2</td>
<td>nicotinamide nucleotide adenyllytransferase 2</td>
</tr>
<tr>
<td>95</td>
<td>OPCML</td>
<td>opioid binding protein/cell adhesion molecule-like</td>
</tr>
<tr>
<td>96</td>
<td>P2RY4</td>
<td>pyrimidinergic receptor P2Y, G-protein coupled, 4</td>
</tr>
<tr>
<td>97</td>
<td>PAX3</td>
<td>paired box 3</td>
</tr>
<tr>
<td>98</td>
<td>PBRM1</td>
<td>polybromo 1</td>
</tr>
<tr>
<td>99</td>
<td>PKD2</td>
<td>polycystic kidney disease 2 (autosomal dominant)</td>
</tr>
<tr>
<td>100</td>
<td>PPP1R8</td>
<td>protein phosphatase 1, regulatory (inhibitor) subunit 8</td>
</tr>
<tr>
<td>101</td>
<td>PRPF40A</td>
<td>PRP40 pre-mRNA processing factor 40 homolog A (S. cerevisiae)</td>
</tr>
<tr>
<td>102</td>
<td>PRPF40B</td>
<td>PRP40 pre-mRNA processing factor 40 homolog B (S. cerevisiae)</td>
</tr>
<tr>
<td>103</td>
<td>PURB</td>
<td>purine-rich element binding protein B</td>
</tr>
<tr>
<td>104</td>
<td>RAP2C</td>
<td>RAP2C, member of RAS oncogene family</td>
</tr>
<tr>
<td>105</td>
<td>RAPH1</td>
<td>Ras association (RalGDS/AF-6) and pleckstrin homology domains 1</td>
</tr>
<tr>
<td>106</td>
<td>RELT</td>
<td>RELT tumor necrosis factor receptor</td>
</tr>
<tr>
<td>107</td>
<td>SENP6</td>
<td>SUMO1/sentrin specific peptidase 6</td>
</tr>
<tr>
<td>108</td>
<td>SLC25A1</td>
<td>solute carrier family 25 (mitochondrial citrate transporter), member 1</td>
</tr>
<tr>
<td>109</td>
<td>SLC2A1</td>
<td>solute carrier family 2 (facilitated glucose transporter), member 1</td>
</tr>
<tr>
<td>110</td>
<td>SLC38A4</td>
<td>solute carrier family 38, member 4</td>
</tr>
<tr>
<td>111</td>
<td>SNX22</td>
<td>sorting nexin 2</td>
</tr>
<tr>
<td>112</td>
<td>SOCS5</td>
<td>suppressor of cytokine signalling 5</td>
</tr>
<tr>
<td>113</td>
<td>ST3GAL5</td>
<td>ST3 beta-galactoside alpha-2,3-sialyltransferase 5</td>
</tr>
<tr>
<td>114</td>
<td>THSD7A</td>
<td>thrombospondin, type I, domain containing 7A</td>
</tr>
<tr>
<td>115</td>
<td>TNNI1</td>
<td>troponin I type 1 (skeletal, slow)</td>
</tr>
<tr>
<td>116</td>
<td>UNK</td>
<td>unkempt homolog (Drosophila)</td>
</tr>
<tr>
<td>117</td>
<td>USP7</td>
<td>ubiquitin specific peptidase 7 (herpes virus-associated)</td>
</tr>
<tr>
<td>118</td>
<td>VEGFA</td>
<td>vascular endothelial growth factor A</td>
</tr>
<tr>
<td>119</td>
<td>VGLL4</td>
<td>vestigial like 4 (Drosophila)</td>
</tr>
<tr>
<td>120</td>
<td>WSB2</td>
<td>WD repeat and SOCS box-containing 2</td>
</tr>
<tr>
<td>121</td>
<td>ZFAND3</td>
<td>zinc finger, AN1-type domain 3</td>
</tr>
<tr>
<td>122</td>
<td>ZFHX4</td>
<td>zinc finger homeobox 4</td>
</tr>
<tr>
<td>123</td>
<td>ZIC1</td>
<td>Zic family member 1 (odd-paired homolog, Drosophila)</td>
</tr>
<tr>
<td>124</td>
<td>ZNF516</td>
<td>zinc finger protein 516</td>
</tr>
<tr>
<td>125</td>
<td>ZNF740</td>
<td>zinc finger protein 740</td>
</tr>
<tr>
<td>126</td>
<td>ZNF827</td>
<td>zinc finger protein 827</td>
</tr>
</tbody>
</table>
127 tcag7.1228 hypothetical protein FLJ25778
128 ABHD13 abhydrolase domain containing 13
129 ANKS6 ankyrin repeat and sterile alpha motif domain containing 6
130 ARHGEF5 Rho guanine nucleotide exchange factor (GEF) 5
131 ASB8 ankyrin repeat and SOCS box-containing 8
132 BLID BH3-like motif containing, cell death inducer
133 BRMS1L breast cancer metastasis-suppressor 1-like
134 C1orf83 chromosome 1 open reading frame 83
135 C2orf15 chromosome 22 open reading frame 15
136 CD84 CD84 molecule
137 CEP350 centrosomal protein 350kDa
138 CSDE1 cold shock domain containing E1, RNA-binding
139 DCP1A DCP1 decapping enzyme homolog A (S. cerevisiae)
140 DIP2C DIP2 disco-interacting protein 2 homolog C (Drosophila)
141 DLX2 distal-less homeobox 2
142 FAM104A family with sequence similarity 104, member A
143 GMEB2 glucocorticoid modulatory element binding protein 2
144 HAO1 hydroxyacid oxidase (glycolate oxidase) 1
145 KIAA1147 KIAA1147
146 KLHL20 kelch-like 20 (Drosophila)
147 LIMCH1 LIM and calponin homology domains 1
148 LPCAT3 lysophosphatidylcholine acyltransferase 3
149 MAF v-maf musculoaponeurotic fibrosarcoma oncogene homolog (avian)
150 MAPK1 mitogen-activated protein kinase 1
151 MLLT4 myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog) translocated to 4
152 MTF1 metal-regulatory transcription factor 1
153 NAV1 neuron navigator 1
154 PISD phosphatidylinositol 4-phosphate 5-kinase D1
155 RALBP1 ralA binding protein 1
156 RAVER1 ribonucleoprotein, PTB-binding 1
157 RPE ribulose-5-phosphate-3-epimerase
158 SDC4 syndecan 4
159 SDK1 sidekick homolog 1, cell adhesion molecule (chicken)
160 SMAD4 SMAD family member 4
161 SOCS1 suppressor of cytokine signalling 1
162 SOCS6 suppressor of cytokine signalling 6
163 SRGAP3 SLIT-ROBO Rho GTPase activating protein 3
164 SSH2 slingshot homolog 2 (Drosophila)
165 TARDBP TAR DNA binding protein
166 TIAM2 T-cell lymphoma invasion and metastasis 2
167 ZAK sterile alpha motif and leucine zipper containing kinase AZK
168 ZNF263 zinc finger protein 263
169 ZNF618 zinc finger protein 618
Figure 30. Expression of PCA3-shRNA2 and primary/mature miRNAs according to nuclear and cytoplasmic localization.

The nuclear:cytoplasmic RNA expression is shown for PCA3-shRNA2 (red) and various primary and mature microRNAs for comparison. For each mature short RNA (including PCA3-shRNA2, ratio 0.60), the majority of the transcript is expressed within the cytoplasm, (ratio <1) in contrast to the primary pri-miR hairpin transcript (ratio >1). The majority of the PCA3 mRNA is within the nucleus (nuclear:cytoplasmic ratio 454.2) (Drayton et al., 2015).
3.3.6.3 The selection and expression of mRNA targets in cell lines

The target mRNAs were annotated with their expression in human PCa samples (Stuart et al., 2004), and were preferentially selected if they’re known to be downregulated in cancer (reflecting our hypothesized targeting by upregulated PCA3-shRNA2: defined as fold change <1.0 and t-test P<0.05) or implicated in PCa biology, and having high predicted binding affinity (e.g. 8-mer seed). The resultant panel (Table 16) included interesting potential targets, such as ETS variant genes 1 and 5 (ETV1 and ETV5), mitogen-activated protein kinase 1 (MAPK31), noggin, N-cadherin, and TEA domain family member 1 (SV40 transcriptional enhancer factor).

Transfection of DU145 cells (chosen as they have low endogenous PCA3-shRNA2 expression) with the PCA3-shRNA2 plasmid and a scrambled RNA sequence was performed. A 1000 fold change seen with PCA3-shRNA2 plasmid represents successful transfection (Figure 31). RNA expression of these 12 predicted targets (Table 17, primers) was measured in PCA3-shRNA2 transfected DU145 cells. Reciprocal knockdown of COPS2 (COP9 signalosome subunit 2), SOX11 (sex determining region Y HMG-box 11), WDR48, TEAD1, and Noggin, suggestive of targeting was identified (Table 16).
Table 16. Selected potential mRNA targets of PCA3-shRNA2.

TargetScan revealed 178 potential mRNA targets of PCA3-shRNA2. A total of 12 mRNA targets were selected for evaluation. These mRNA have high predicted binding affinity, are known to be downregulated in cancer (Stuart et al., 2004) or implicated in prostate oncogenesis. The expression of the 12 mRNAs were measured following transfection of DU145 cells with PCA3-shRNA2 plasmid. Fold change was calculated relative to non-transfected cells (Drayton et al., 2015). "Mer" suffix refers to the number of bases.
Figure 31. Transfection of DU145 with PCA3-shRNA2 and controls.

DU145 cells were transfected with PCA3-shRNA2 (red) or a scrambled RNA sequence (mock). Expression of PCA3-shRNA2 in cells transfected with PCA3-shRNA2 (Red), scrambled RNA sequence and untransfected cells (control) is shown. Bars represent the mean of three independent repeats and standard deviation. A 1000 fold change seen with PCA3-shRNA2 plasmid (red) represents successful transfection (Drayton et al., 2015).
Table 17. Primers and condition used to detect target mRNAs.

TargetScan revealed 178 mRNA targets of PCA3-shRNA2. mRNAs (n=12) with predicted high binding affinity and implicated in prostate oncogenesis were selected for validation. Primers and conditions used for qRT-PCR are shown (Drayton et al., 2015).

<table>
<thead>
<tr>
<th>ID</th>
<th>Name</th>
<th>Fwd Primer</th>
<th>Rev Primer</th>
<th>Annealing Temp</th>
<th>Amplicon Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>KIAA0515</td>
<td>KIAA0515</td>
<td>TGGCTCACCTTCGTCATCTGA</td>
<td>TCATCCTCGGATACTGTGGAA</td>
<td>60 °C</td>
<td>215</td>
</tr>
<tr>
<td>COPS2</td>
<td>COP9 constitutive photomorphogenic homolog subunit 2 (Arabidopsis)</td>
<td>TTTTAGCCAGTTACATGTCG</td>
<td>CTTCCCTCAAGTGCATTTTACCA</td>
<td>60 °C</td>
<td>234</td>
</tr>
<tr>
<td>TEAD1</td>
<td>TEA domain family member 1 (SV40 transcriptional enhancer factor)</td>
<td>GCCCGGAATGATTCAACAG</td>
<td>CAATTGGACACCTTGCA</td>
<td>60 °C</td>
<td>165</td>
</tr>
<tr>
<td>CDH2</td>
<td>cadherin 2, type 1, N-cadherin (neuronal)</td>
<td>TCAGGCGCTGTAGAGGCTT</td>
<td>ATGCAATCCTTCGATAAGACTG</td>
<td>60 °C</td>
<td>94</td>
</tr>
<tr>
<td>ETV1</td>
<td>ets variant gene 1</td>
<td>TGGCAGTTTTGAGTCTTTTC</td>
<td>CGGAGTGAAACGGCTAAGTTTATC</td>
<td>60 °C</td>
<td>170</td>
</tr>
<tr>
<td>ETV5</td>
<td>ets variant gene 5 (ets-related molecule)</td>
<td>CAGTCAACTTCAAGAGGCTTG</td>
<td>TGCTACATGGCTACAAGACGAC</td>
<td>60 °C</td>
<td>168</td>
</tr>
<tr>
<td>INVS</td>
<td>inversin</td>
<td>TGCTCTACAGAGGCTCATG</td>
<td>AGGCAATAACATAGTGTTTCTCT</td>
<td>60 °C</td>
<td>84</td>
</tr>
<tr>
<td>MAP3K1</td>
<td>mitogen-activated protein kinase kinase 1</td>
<td>TCTCACCATTAGGCCCTGAGGA</td>
<td>AGGAAAGATTTAGGCCCTATCTG</td>
<td>60 °C</td>
<td>97</td>
</tr>
<tr>
<td>NOG</td>
<td>noggin</td>
<td>CCATGCAGGAGGAGATCAA</td>
<td>TCGGAAATGATGGGTTACTG</td>
<td>60 °C</td>
<td>337</td>
</tr>
<tr>
<td>SOX11</td>
<td>SRY (sex determining region Y)-box 11</td>
<td>AGGATTTCGATTCACAGG</td>
<td>AGGTCGGAGAAGTTCGCTT</td>
<td>60 °C</td>
<td>121</td>
</tr>
<tr>
<td>WDR1</td>
<td>WD repeat domain 1</td>
<td>TGGGATTCCAGCAATTAGTTGGA</td>
<td>CCAGATAGTTGATTGACCCGAC</td>
<td>60 °C</td>
<td>209</td>
</tr>
<tr>
<td>WDR48</td>
<td>WD repeat domain 48</td>
<td>TGGGACAATAGGCTTTTGGTC</td>
<td>TGTCAGGTTTTCTTAGTTCTGT</td>
<td>60 °C</td>
<td>164</td>
</tr>
</tbody>
</table>
3.3.6.4 Expression of target mRNAs in urinary samples

The mRNA expression of the two strongest (largest reciprocal fold change following PCA3-shRNA2 transfection) candidates (COPS2 and SOX11) was measured in the larger urinary sample cohort (n=471) to look for biologic associations in-vivo. There was a significant inverse correlation between the expression of COPS2 and PCA3-shRNA2 (Figure 3, $r=-0.32$, $P<0.001$) and reduced expression of COPS2 in the PCa samples (Figure 3, fold change, $0.29\pm0.5$; t-test $P<0.001$) was identified when compared with controls. Non-significant lower expression for SOX11 was also seen in cancerous samples (fold change, $0.74\pm1.5$; $P=0.08$) when compared with controls and this mRNA was not significantly correlated with PCA3-shRNA2 expression ($r=-0.1$; $P=0.48$).
COPS2 had the lowest fold change when DU145 cells were transfected with PCA3-shRNA2 and was further validated in the larger urinary cohort (n=471). The expression (qRT-PCR) of COPS2 was inversely correlated to that of PCA3-shRNA2 (Drayton et al., 2015).
Figure 33. The expression of COPS2 in urinary samples.

COPS2 expression (qRT-PCR) was measured in urinary samples (n=471). Expression of COPS2 was lower in the urinary cells of patients with prostate cancer (n=355) when compared to BPH (n=116) controls (Drayton et al., 2015).
3.4 Discussion

Many transcribed RNAs do not encode proteins and these RNAs are termed non-coding RNAs. They are classified according to size and cellular location. Although short RNAs/miRNAs (~20-22-bp in size) have been extensively studied (Catto et al., 2011), little is known about the function of most long ncRNAs. Long ncRNAs may have direct involvement in chromatin remodeling and androgen receptor regulation (Yang et al., 2013), and may be processed into shorter more active ncRNAs. Many miRNAs are clustered together and are derived from single primary transcripts (i.e. miRs-24-2/27a/23a) (Drayton et al., 2014). Rogler et al showed that RNase MRP (a 268-bp ncRNA component of mitochondrial RNA processing endoribonuclease) was the source for two shorter (~20-bp) RNAs involved in the biology of cartilage-hair hypoplasia (Rogler et al., 2014). Therefore, we hypothesized that PCA3 could be a source for short biologically active RNAs.

3.4.1 Identification of PCA3-shRNA2 and expression in urinary samples

We present *in-silico* and *in-vitro* data suggesting that a short RNA hairpin is produced during processing of the PCA3 transcript. This short RNA is located within intron 1 of PCA3, close to a region of high species conservation, suggesting biologic protection. Our short ncRNA expression appeared closely correlated with that for PCA3 in both cell lines and urinary samples. This was expected as our data suggested that the short RNA is derived from the PCA3 transcript. In post-prostate massage (DRE) urinary cell pellets from two large patient cohorts, we found that PCA3-shRNA2 detected PCa with a similar accuracy to PCA3.

Unlike long ncRNAs, short ncRNAs are stable molecules and do not decay with repeated freeze-thawing or prolonged storage at room temperature. We previously showed that short RNAs do not dramatically degrade with freeze-thawing and prolonged storage at room temperature (in plain clean universal containers without RNase inhibitors) (Miah et al., 2012). As such, PCA3-shRNA2 may be a more stable biomarker for PCa than the current PCA3 test. Assays to detect our PCA3-shRNA2 would not be as vulnerable to delays in
handling or variations in stringency in collection, and so should be more reproducible.

3.4.2 The biological role of PCA3-shRNA2

Nuclear and cytoplasmic RNA ratios suggested that PCA3-shRNA2 is abundant in the cytoplasm. In contrast, PCA3 is more abundant in the nucleus. This finding is not entirely clear but may suggest a potential functional mRNA targeting role in the cytoplasm for PCA3-shRNA2. Another method to investigate RNA localization include RNA fluorescence in-situ hybridization (FISH) (Coassin et al., 2014). An unbiased genome-wide computational search through the use of TargetScan (potential mRNA targets), annotation from PCa microarray (genes implicated in PCa) and DAVID (oncogenic significance) identified genes and pathways implicated in the biology of PCa.

Our targeting analysis identified expression changes in COPS2, SOX11, WDR48, TEAD1, and Noggin when PCA3-shRNA2 upregulation was induced. These mRNAs are involved in the regulation of gene transcription, uro-genital tract development, and in cell growth and signalling. Therefore, they appear ideal oncogenic gene candidates. We explored the expression of COPS2 and SOX11 in exfoliated urinary cell pellets. We found that COPS2 expression was inversely correlated to PCA3-shRNA2 and significantly reduced in cancerous urinary samples (p<0.001), suggesting biological validation. Although, SOX11 expression was also inversely correlated to PCA3-shRNA2 and its expression was lower in cancerous urinary samples, this did not reach statistical significance.

COPS2 is a transcription corepressor that underwent a decreased expression in cells with PCA3-shRNA2 upregulation. COPS2 is a component of the COP9 signalosome complex that regulates the ubiquitin conjugation pathway during various cellular and developmental processes, including phosphorylation of p53. COPS2 is abundantly expressed in most human tissues, suggesting an important role in cellular homeostasis, but has not
been studied in depth with respect to human malignancies. SOX11 is a transcription factor belonging to the SRY-related HMG-box (SOX) family. These regulate many biological processes, including haematopoiesis, vasculogenesis, and cardiogenesis during embryonic development (Stovall et al., 2014), and some members are negative regulators of the WNT-beta-catenin-TCF pathway (Katoh, 2002) which is associated with prostate biology. Although Katoh et al reported reduced expression of SOX7 in PCa cells, SOX11 function and expression has not been reported in PCa at the time of our analysis.

Of the other predicted targets, noggin appears particularly interesting because of its association with bone metastasis. Noggin is an antagonist of bone morphogenetic proteins (BMP) (Haudenschild et al., 2004), which has been reported to be downregulated in PCa cells (Schwaninger et al., 2007; Secondini et al., 2011). Noggin loss leads to the development of bone metastases. Therefore, reversal of noggin loss may be used to palliate or reduce the activity of osteolytic malignant disease.

3.5 Conclusions
We found a short RNA (PCA3-shRNA2) that is derived from the PCA3 gene that is probably co-expressed with PCA3. We identified a potential role for this ncRNA in PCa biology. The short RNA may be a more suitable target of the PCA3 biomarker assay.
CHAPTER 5: PCA3-shRNA2 Expression and Eventual Diagnosis of Prostate Cancer
5.1 Background
In Chapter 3 we reported retrospectively our initial and repeat PBx data from the Sheffield cohort of the ProtecT study. Patients with an initial negative biopsy and persistently elevated PSA levels are difficult to manage. Our data showed that 321/920 (34.9%) men were diagnosed with PCa on initial biopsy and 66/248 (26.6%) were diagnosed with PCa on rPBx. These data were consistent with the international data presented in Table 1. Whilst MRI now appears a promising tool in this context, there is an urgent need to identify biomarkers that may inform rPBx decisions.

In Chapter 4 we identified a short RNA within PCA3, which we termed PCA3-shRNA2. We showed that PCA3-shRNA2 was expressed in PCa cell lines and overexpressed in urinary samples obtained from patients with PCa. In addition, we explored the functional roles of this short RNA and found that it targets mRNAs involved in PCa biology (including COPS2 and SOX11).

Our initial analysis of PCA3-shRNA2 used urine samples from men with and without PCa. All specimens were taken at first presentation or diagnosis (Drayton et al., 2015). Although PCA3-shRNA2 appeared to be a promising option as a urinary biomarker, its expression in PBx tissues warranted exploration, since the PROGENSA PCA3 assay is clinically advocated for guiding rPBx in men with an elevated PSA. We aimed to investigate whether PCA3-shRNA obtained from the negative initial PBx has a predictive role in PCa detection on rPBx.

5.2 Methods

5.2.1 Expression of PCA3-shRNA2 in FFPE samples
The pathology database for men with an initial PBx between 1994 and 2010 (to allow follow-up) was searched. An annotation with clinical details, the number and timing of rPBx, and the eventual diagnosis of PCa was performed. Men whose initial PBx did not show cancer were identified, and a matched cohort whose rPBx did or did not find PCa was made.
5.2.2 RNA extraction from prostate biopsies (FFPE)

The formalin fixed paraffin embedded (FFPE) blocks from the initial PBx were retrieved, and cut sections at 10µm thickness were obtained (Sectioning from FFPE blocks was performed by Maggie Glover). One section was stained with H&E to confirm diagnosis and extraction of RNA was performed from the remaining. Paraffin (deparaffinization solution, Qiagen, UK) was removed before lysis with Proteinase K. Samples were treated with DNase to eliminate all genomic DNA, before washing and elution in RNase-free water. Total and miRNA were extracted using miRNeasy FFPE kit (Qiagen, UK) as per manufacturer’s protocol and measured using a 2100 Bioanalyzer (Agilent).

As detailed (Drayton et al., 2015) extracted RNA was subject to real-time quantitative RT-PCR (HT7900 PCR system) using the High-capacity reverse transcription cDNA kit (Applied Biosystems) and the TaqMan microRNA Reverse Transcription kit (Applied Biosystems). RNA expression was determined using qPCR with TaqMan primers for PSA (Assay ID: Hs03063374_m1), PCA3 (Assay ID: Hs03309852_g1) and two custom designed TaqMan assays, PCA3-shRNA2A and PCA3-shRNA2B.

PCA3-shRNA2A
ACTGCACTCCAGCCTGGGCA (Ambion: assay IDs, CSGJ090)

PCA3-shRNA2B
CACTGCACTCCAGCCTGGGCA (Ambion: assay IDs, CSHSNF8)

Expression of PCA3, and PCA3-shRNA2 was normalized to PSA (Clarke et al., 2009) calculated using DCt values (Miah et al., 2012).

PSA forward
5-GCATCAGGAACAAAGCGTG-3
PSA reverse
5-CCTGAGGAATCGATTCTTCA-3
5.3 Results

5.3.1 Patients and FFPE samples
Residual tissue from the first PBx (between 2002 and 2008) of 116 men with an eventual diagnosis of PCa (rPBx between 2002 and 2013) and 94 men without PCa were obtained (Table 18). The two populations were broadly comparable for clinical features. The mean (±SD) age at referral within our cohort was 63.5 (±7.1) for men with cancer and 62.5 (±6.5) years for those with benign PBx. The mean initial PSA was 9.5 (±1.8) and 13.2 (±61.4) in the cancer and benign group respectively. A total of 17/23 (73.4%) men with suspicious findings and 4/7 (57.1%) men with HGPIN on initial PBx were found to have cancer on rPBx. The majority of cancers (n=74, 64%) were detected on the second PBx and Gleason score 3+3=6 was the most common grade amongst these tumours (n= 78, 67.2%). The mean (SD) time to diagnosis of PCa from the initial negative PBx was 29.8 (±36.6) months.
Table 18. Patients and FFPE samples analysed in this report.

A total of 210 men had a negative initial PBx. All men underwent at least one rPBx and 116 men were found to have PCa. The age and PSA in the malignant and matched benign groups, and the Gleason score for PCa detected on rPBx are shown. (Pang et al., 2017).
5.3.2 Expression of PCA3-shRNA2 in prostate biopsies (FFPE)
PSA and PCA3-shRNA2 RNA was detected in all samples, and PCA3 mRNA in 190 (90%) biopsies. Expression of PCA3 and PCA3-shRNA2 was normalized to PSA mRNA, as for the PROGENSA assay. As seen previously, expression of PCA3 and PCA3-shRNA2 were correlated (Pearson’s r=0.69, p<0.01, Figure 34) suggesting co-expression. Interestingly, there appeared no deterioration in RNA yield across the time period that the samples were stored (Figure 35).
Figure 34. Scatterplot of PCA3 and PCA3-shRNA2 RNA expression normalized to PSA mRNA in FFPE benign prostate biopsies.

RNA was extracted from FFPE samples and PCA3 and PCA3-shRNA2 expression was measured with qRT-PCR. PCA3-shRNA2 was correlated to PCA3 expression in FFPE benign prostate sample (Pang et al., 2017).

*Red: prostate cancer detected on repeat biopsies; Black: benign
Figure 35. PSA, PCA3 and PCA3-shRNA2 expression in FFPE prostate biopsies stratified by year of collection.

FFPE initial PBx specimens were retrieved and categorized into years they were obtained. PSA, PCA3 and PCA3-shRNA2 expression (qRT-PCR) was measured. The level of PCA3-shRNA2 (also PSA and PCA3) expression did not vary in PBx obtained between 2002 and 2008 (Pang et al., 2017).

* Line in the box represents the median and whiskers represent the highest and lowest DCt value.
5.3.3 PCA3-shRNA2 expression and eventual diagnosis

RNA expression with the eventual diagnosis in each man was compared. We saw upregulation of PCA3 (average 2.1-fold) and PCA3-shRNA2 (average 1.5-fold) in men with an eventual diagnosis of cancer, when compared to those with only benign histology (Table 19). For PCA3, this difference reached statistical difference (t-test p=0.02), but this was not the case for PCA3-shRNA2 (p=0.2, Figure 36). When evaluating PCA3 and PCA3-shRNA2 expression with respect to Gleason scores, PCA3 expression was significantly higher in low Gleason Scores. No differences were observed with PCA3-shRNA2 (Figure 37).
### Eventual diagnosis

<table>
<thead>
<tr>
<th></th>
<th>Eventual diagnosis</th>
<th>Mean DCt* ± St. Dev</th>
<th>Mean DCt* ± St. Dev</th>
<th>Fold change**</th>
<th>t-test p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PCA3-shRNA2</strong></td>
<td>Cancer</td>
<td>-16.53 ± 3.24</td>
<td>-15.99 ± 3.61</td>
<td>1.45</td>
<td>0.20</td>
</tr>
<tr>
<td></td>
<td>Benign</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>PCA3</strong></td>
<td>Cancer</td>
<td>6.40 ± 3.03</td>
<td>7.48 ± 2.86</td>
<td>2.12</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>Benign</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 19.** RNA expression stratified by eventual diagnosis.

PCA3 and PCA3-shRNA2 expression was measured in the initial PBx FFPE specimens obtained from men with or without PCa detected on rPBx. There was a significant elevated PCA3 RNA level (fold change 2.12, p=0.02) from men who were diagnosed with PCa on rPBx (Pang *et al.*, 2017).

* Delta Ct (Dct) normalized to PSA mRNA

** Fold change in patients with cancer versus those with BPH
Figure 36. Box plot of PCA3 and PCA3-shRNA2 expression stratified for eventual diagnosis.

PCA3 and PCA3-shRNA3 expression (qRT-PCT) was measured in FFPE. PCA3 and PCA3-shRNA2 DCt levels were normalized to PSA mRNA. The expression of PCA3 in FFPE of initial PBx were significantly higher in men who were diagnosed with PCa on rPBx compared to benign pathology (Pang et al., 2017).
Figure 37. PSA, PCA3 and PCA3-shRNA2 expression in tissues with respect to Gleason grade.

The expression of PCA3 and PCA3-shRNA2 (normalized to PSA) in FFPE initial PBx specimens was compared across Gleason scores. PCA3 expression is higher in low-grade disease (Gleason 6). No differences in PCA3-shRNA2 expression was seen with respect to Gleason scores (Pang et al., 2017).
5.3.4 The ability of PCA3-shRNA2 to identify disease (PBx FFPE)

To determine predictive role, we calculated the concordance index for each RNA (Figure 38). Whilst, PCA3 was associated with the detection of PCa (C-index 0.61, p=0.01), this was not the case for PCA3-shRNA2 (C-index 0.55, p=0.22).

Figure 38. Predictive ability of PCA3 and PCA3-shRNA2 expression for the eventual diagnosis of prostate cancer.

Receiver operating characteristic curve analysis was used to assess the diagnostic role of PCA3-shRNA2 in FFPE. The concordance index for PCA3 and PCA3-shRNA2 was 0.61 and 0.55 respectively (Pang et al., 2017).
5.4 Discussion

The clinical context of the PCA3 assay is to guide further investigation in men with an elevated serum PSA, whose PBx do not reveal cancer. With this in mind, we undertook a further analysis of PCA3 in rPBx specimens. Whilst the PCA3 assay uses urine samples to measure PCA3, this resource was not available, and others have found that urine PCA3 expression reflects that within tissues (Hessels et al., 2003). Both RNAs and proteins are considered good candidates for biomarkers in FFPE tissues, due to their stability over long periods of time (Sequeiros et al., 2013).

We selected men with and without PCa, matched for age, PSA and duration of rPBx, and extracted RNA from stored FFPE tissue not used during pathological reporting. We identified several key findings. Despite storage at room temperature for 8-14 years, we detected robust expression of PSA and PCA3 and PCA3-shRNA2. The latter was most abundant, in keeping with its stability and ease of detection. We identified correlated expression of the two RNAs expressed from the PCA3 gene confirming our previous observations and suggesting that the short ncRNA is a product of PCA3 transcription (Drayton et al., 2015). These two observations suggest that PCA3-shRNA2 may be incorporated into the PCA3 assay to facilitate easier handling of the biological samples prior to laboratory measurement.

However, our experiments failed to support our primary hypothesis, namely that PCA3-shRNA2 expression was associated with a subsequent diagnosis of PCa. This is in contrast to expression of PCA3 mRNA within our population, and may reflect that this ncRNA is not expressed by PCa, that ncRNAs have a dynamic expression that is less stable than mRNAs over time, or that our experimental design was wrong. With regards to the former, we previously found high malignant expression of PCA3-shRNA2 in three separate cohorts (i.e. PCa cell lines, fresh frozen microdissected prostatic tissues and prostate massage fluids) suggesting this may not be the explanation. With regards to dynamic expression, it is known that one function of short ncRNAs is to epigenetically regulate mRNA expression.
This regulation is dynamic, with ncRNA expression fluctuating depending upon the cellular needs and stress. For example, individual ncRNAs have been found to have both oncogenic and tumour suppressive roles, depending upon the context (Svoronos et al., 2016). Thus, it is plausible that 2-3 years before the diagnosis of cancer, PCA3-shRNA2 expression is not elevated, as the target mRNAs (such as COPS2 and SOX11) do not have, as yet, altered function in the prostate. With regards to experimental design, we powered the study using expression estimates and used FFPE tissues (to replace urine samples). We did find a trend towards upregulation of PCA3-shRNA2 in cancer, suggesting under-powering of the sample size. It may also be that FFPE tissues do not preserve differential expression of all RNAs.

Of note, previous analyses of PCA3 expression in PBx tissues have reported inconsistencies, with upregulation in PCa and no difference between malignant and normal prostate (Klecka et al., 2010; Paziewska et al., 2014; Alinezhad et al., 2016).

Extracting good quality RNA from stored FFPE can be difficult as quality is affected by paraffin, and often FFPE specimens become brittle due to age which makes sectioning difficult resulting in lower yields of useable tissue. The yield and molecular weight of recovered RNA are often low as evident in our current analysis. Although we detected PCA3-shRNA2 in the historic FFPE specimens, we should have used RNA integrity number (RIN) to measure RNA integrity prior to qRT-PCR and also sequencing of the PCR product in order to confirm detection (Schroeder et al., 2006; Bustin et al., 2009). The positive controls used for the urinary/FFPE experiments were PSA (used to normalize PCA3 and PCA3-shRNA2) and U1/GAPDH. Negative control was non-template control (NTC) using NF-water. In addition, the use of controls without reverse transcription enzymes and appropriate primer design (e.g. intron-spanning) are important to exclude the possibility of genomic DNA amplification. Other positive controls that could have been used include RNU 44/48 and hsa-miR-26b/92 as these control RNAs have been shown the least variability.
5.5 Conclusion
We showed stable expression of PSA, PCA3 and PCA3-shRNA2 in historic PBx samples. Whilst PCA3 and PCA3-shRNA2 expression were correlated, only the former was significantly associated with the presence of occult PCa.
CHAPTER 6: N6-Adenosine Methylation and Cancer
6.1 Background
Although PCA3-shRNA2 is expressed in urinary and PBx (FFPE) samples obtained from patients with PCa (Chapter 4 and 5), and overexpressed in the former when compared with benign samples, we found no correlation between localised and advanced disease. In addition, the mechanistic drive for the aberrant RNA expression was unknown. To explore mechanisms of altered RNA expression in PCa, we focused upon RNA methylation as a newly identified epigenetic trait. Whilst it has been known for many years that methylation of N6-adenosine base (m6A) is the most common epigenetic modification of RNA (Wei et al., 1976), technological limitations have prevented in-depth analysis (Pollex et al., 2010). Recent molecular advances have now overcome these limitations and the epigenome wide distribution of m6A has been reported in a human (hepatocellular carcinoma cell line (HepG2)) and mouse (normal liver) cells (Dominissini et al., 2012). The authors identified over 12,000 m6A sites on mRNAs of >7,000 human genes, these sites were highly conserved between human and mouse, and preferentially located within stop codons and long internal exons. They also identified that m6A is a dynamic mark, associated with cell stimuli and cell phenotypes. N6-methyladenosine is known to be regulated by ‘writers’ (METTL3/4/14, WTAP), ‘erasers’ (FTO, ALKBH5) and ‘readers’ (YTHDF1/2/3), and is associated with RNA splicing, export, decay and translation (Jia et al., 2011; Fu et al., 2014). In addition, it has been shown that m6A is associated with a wide range of disease processes including obesity, inflammation and cancer (Leukaemia, prostate, breast, colorectal, gastric cancers) (Maity et al., 2015).

The importance of the m6A in malignancy is reported, but the extent and relative distribution of this event in common cancers is currently unknown (Fu et al., 2014). In addition, to date, mammalian m6A sites have been mapped and characterized in only a small number of mammalian cell lines/tissues. Whether m6A is abundant across common human tumours is unknown.

We hypothesised that the conservation of m6A sites would allow an in-silico analysis to determine the likely extent of m6A within common cancers, and
the potential biological implications of this. Using an *in-silico* approach, we aimed to explore whether m6A was abundant in a range of tumours and compared the expression of mRNAs reported to undergo methylation of N6-adenosine in a sample of common human cancers including prostate, bladder, renal, lung, breast, ovarian, colorectal and gastric cancer. Part two of this project involved identifying m6A within androgen-sensitive (LNCaP) and its metastatic (LNCaP-LN3) sibling cell lines through methylated RNA immunoprecipitation and next generation sequencing (MeRIP-seq).

6.2 Methods

5.2.1 *In-silico* analysis of m6A

5.2.1.1 Selection and annotation of RNA transcriptomic datasets

Array Express (http://www.ebi.ac.uk/arrayexpress/ 2014) was searched for publically deposited RNA expression microarray datasets reporting transcriptomes within prostate, bladder, renal, lung, breast, ovarian, colorectal and gastric cancer. Commercially manufactured, non-custom platforms were filtered (preferentially selecting the Affymetrix HG-U133A/B/plus 2 platforms) to reduce experimental variation between experiments. Processed data in tab-delimited text (*.txt) files were extracted, annotated for RNA locus, sample details and m6A susceptibility. m6A susceptibility was defined using topology by Dominissini et al. (Dominissini et al. 2012); specifically, susceptible loci were RNAs whose N6-adenosine methylation peaked under all experimental conditions (untreated, exposure to ultraviolet-radiation, heat shock, hepatocyte growth factor and interferon-γ) in HepG2 hepatic cell carcinoma and normal human brain cells.

6.2.1.2 RNA Selection

Differential mRNA expression was calculated within each microarray dataset using the Significance Analysis of Microarrays (SAM) method, Student's t-test p<0.001 and FDR <0.05 (Tusher et al., 2001). Fold change (FC) was calculated using median values across each sample type. Comparisons were between malignant and non-malignant tissues within each cancer, between
cancer types and within cancer phenotypes when available (e.g. androgen-sensitive prostate cancer (AS-PCa) and castration-resistant prostate cancer (CRPC)). Differentially expressed mRNAs were ranked according to m6A susceptibility (m6A susceptible, m6A(+); m6A not susceptible, m6A(-)) and the frequency of their aberrant expression (across and within cancer types). To predict oncogenic processes that involve m6A, the DAVID bioinformatics database (http://david.abcc.ncifcrf.gov/ 2014) (Huang et al. 2009) was used to identify the gene ontology (GO) terms that are enriched for m6A predicted transcripts. Coding genes were functionally clustered at ‘high’ stringency, and clusters with an enrichment score of >1.5 were selected for analysis.

6.2.1.3 Statistical analysis
Data were plotted using GraphPad Prism 6 and Microsoft Excel 2016. Parametric continuous data were compared using Student’s t-test and a threshold of p>0.05 was considered significant. Differential mRNA expression was calculated using SAM method (Student’s t-test p<0.001 and FDR <0.05).

6.2.2 Immunoprecipitation and sequencing of m6A
To test our in-silico findings in cultured cells, RNA extraction followed by m6A immunoprecipitation and RNA sequencing was performed. RNA extraction and fragmentation using Prime 5 PerfectPure RNA cultured cell kit (Scientific Laboratory Supplies) was described in Chapter 2.4.1. A total of 2mg of total RNA was extracted from each LNCaP and LNCaP-LN3 cell lines. Fragmentation was performed using Zinc Chloride and fragmented RNA was subject to IP using anti-m6A rabbit polyclonal antibody (Synaptic Systems). Immunoprecipitation with the same amount of input RNA (1mg total RNA) without anti-m6A antibody served as a negative control. RNA-sequencing on the IP libraries was performed using an Illumina Hi-Seq 2500 (Performed by Emilie Jarratt from the Sheffield diagnostic genetic services, Sheffield children’s hospital). Bioinformatic analyses were performed with colleagues (Dr Ian Sudbery and Dr James Bradford) from the University of Sheffield Bioinformatic Hub (Chapter 2.5.1).
6.3 Results

6.3.1 Microarray datasets and sample population

At a total of 47 microarray datasets (Tables 20 and 21) reporting mRNA profiles in 2,405 (range: 172-998) cancer and 1,434 (83-616) control samples were identified and extracted (Lenburg et al., 2003; Dyrskjøt et al., 2004; Jones et al., 2005; Hendrix et al., 2006; Liu et al., 2006; Richardson et al., 2006; Gumz et al., 2007; Moreno et al., 2007; Chandran et al., 2007; Galamb et al., 2008, 2010; Bahrani-Mostafavi et al., 2008; Landi et al., 2008; Arredouani et al., 2009; Gyorffy et al., 2009; Mok et al., 2009; Casey et al., 2009; Yusenko et al., 2009; D’Errico et al., 2009; Hou et al., 2010; Kim et al., 2010; Pau Ni et al., 2010; Shiraishi et al., 2010; Skrzypczak et al., 2010; Wang et al., 2010, 2012, 2013; Zhang et al., 2010; Holbrook et al., 2011; King et al., 2011; Sanchez-Palencia et al., 2011; Chen et al., 2011; Cho et al., 2011; Alhopuro et al., 2012; Jia et al., 2012; Lopez et al., 2012; Okayama et al., 2012; Urquidi et al., 2012; Eftang et al., 2013; Kuner et al., 2013; Clarke et al., 2013; Ryan et al., 2014). These datasets included multiple comparisons between malignant and non-malignant (benign/normal) tissues and two datasets comparing AS-PCa with CRPC (Best et al., 2005; Cai et al., 2013). The average mean (SD) experiment size was 84.6 (±63.7) samples.
A total of 47 datasets were extracted from Array express and analysed. Percentages of m6A susceptible (m6A+) and non-susceptible (m6A-) mRNAs of those with significant aberrant expression (Student’s t-test p<0.001) in cancer are shown.

CRPC, castration-resistant prostate cancer; ca, cancer

### Table 20. Total number of cancer and control samples within each cancer type.

<table>
<thead>
<tr>
<th>Cancer</th>
<th>Number Datasets</th>
<th>Total samples (n, range)</th>
<th>% m6A (+) Mean (range)</th>
<th>% m6A (-) Mean (range)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prostate</td>
<td>7</td>
<td>Ca: 292 (11-68) Control: 294 (8-71)</td>
<td>47.8 (37.6-55.1)</td>
<td>52.2 (45.0-62.4)</td>
</tr>
<tr>
<td>CRPC</td>
<td>2</td>
<td>CRPC: 39 (10-29) AS-PCa: 32 (10-22)</td>
<td>57.2 (56.2-58.2)</td>
<td>42.8 (41.8-43.8)</td>
</tr>
<tr>
<td>Bladder</td>
<td>4</td>
<td>Ca: 291 (10-188) Control: 129 (7-68)</td>
<td>49.9 (42.5-59.4)</td>
<td>50.1 (40.6-57.6)</td>
</tr>
<tr>
<td>Renal</td>
<td>6</td>
<td>Ca: 175 (10-69) Control: 56 (4-23)</td>
<td>40.1 (26.3-54.1)</td>
<td>59.9 (46.0-73.7)</td>
</tr>
<tr>
<td>Breast</td>
<td>6</td>
<td>Ca: 342 (42-104) Control: 232 (7-143)</td>
<td>45.2 (39.0-54.8)</td>
<td>54.8 (45.2-61.1)</td>
</tr>
<tr>
<td>Lung</td>
<td>6</td>
<td>Ca: 481 (30-226) Control: 219 (20-65)</td>
<td>41.7 (27.3-55.4)</td>
<td>58.3 (44.6-72.7)</td>
</tr>
<tr>
<td>Colorectal</td>
<td>5</td>
<td>Ca: 213 (15-81) Control: 140 (8-55)</td>
<td>52.9 (47.8-59.2)</td>
<td>47.1 (40.8-52.2)</td>
</tr>
<tr>
<td>Ovarian</td>
<td>5</td>
<td>Ca: 252 (32-99) Control: 65 (4-35)</td>
<td>48.6 (36.2-56.8)</td>
<td>51.4 (43.2-63.8)</td>
</tr>
<tr>
<td>Gastric</td>
<td>6</td>
<td>Ca: 320 (12-134) Control: 267 (15-134)</td>
<td>45.8 (39.7-53.9)</td>
<td>54.2 (46.1-60.3)</td>
</tr>
<tr>
<td>Total</td>
<td>47</td>
<td>Ca: 2405 (172-998) Control: 1434 (83-616)</td>
<td>46.7 (26.3-59.4)</td>
<td>53.3 (40.6-73.7)</td>
</tr>
</tbody>
</table>
Table 21. Summary of the 47 microarray datasets extracted from Array Express.

The number (%) and mean fold changes (FC) of aberrantly (t-test p<0.001) expressed transcripts that are susceptible (m6A+) and not susceptible (m6A-) to m6A-methylation are shown. mRNAs were also subcategorised into upregulation (FC within top 10% percentile) and downregulation (FC within bottom 10% percentile).

<table>
<thead>
<tr>
<th>Author</th>
<th>Cancer</th>
<th>Cancer (n)</th>
<th>Benign/Normal (n)</th>
<th>Array</th>
<th>m6A+</th>
<th>n</th>
<th>%</th>
<th>Mean (SD) FC</th>
<th>m6A-</th>
<th>n</th>
<th>%</th>
<th>Mean (SD) FC</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wang Y</td>
<td>Prostate</td>
<td>53</td>
<td>71</td>
<td>HG-U133A</td>
<td>Total</td>
<td>1057</td>
<td>2</td>
<td>47.6</td>
<td>1.00 + 0.04</td>
<td>Total</td>
<td>11643</td>
<td>2</td>
<td>52.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p&lt;0.001</td>
<td>3077</td>
<td>29.1</td>
<td>1.00 + 0.06</td>
<td>p&lt;0.001</td>
<td>2555</td>
<td>21.9</td>
<td>1.01 + 0.07</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>FC&gt;90%</td>
<td>927</td>
<td>8.8%</td>
<td>1.07 + 0.04</td>
<td>FC&gt;90%</td>
<td>986</td>
<td>8.5%</td>
<td>1.08 + 0.05</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p&lt;0.001</td>
<td>1074</td>
<td>10.2</td>
<td>0.94 + 0.03</td>
<td>p&lt;0.001</td>
<td>730</td>
<td>6.3%</td>
<td>0.94 + 0.04</td>
<td></td>
</tr>
<tr>
<td>Wang Y</td>
<td>Prostate</td>
<td>68</td>
<td>69</td>
<td>HG-U133 Plus 2.0</td>
<td>Total</td>
<td>1895</td>
<td>3</td>
<td>34.7</td>
<td>1.00 + 0.05</td>
<td>Total</td>
<td>35660</td>
<td>2</td>
<td>65.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p&lt;0.001</td>
<td>5464</td>
<td>%</td>
<td>28.8</td>
<td>1.01 + 0.08</td>
<td>p&lt;0.001</td>
<td>7394</td>
<td>20.7</td>
<td>0.99 + 0.10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>FC&gt;90%</td>
<td>1471</td>
<td>7.8%</td>
<td>1.09 + 0.07</td>
<td>FC&gt;90%</td>
<td>2176</td>
<td>6.1%</td>
<td>1.11 + 0.07</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p&lt;0.001</td>
<td>1222</td>
<td>6.4%</td>
<td>0.92 + 0.03</td>
<td>p&lt;0.001</td>
<td>2999</td>
<td>8.4%</td>
<td>0.90 + 0.04</td>
<td></td>
</tr>
<tr>
<td>Liu P</td>
<td>Prostate</td>
<td>44</td>
<td>13</td>
<td>HG-U133A</td>
<td>Total</td>
<td>1046</td>
<td>4</td>
<td>47.3</td>
<td>1.03 + 0.37</td>
<td>Total</td>
<td>11669</td>
<td>2</td>
<td>52.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p&lt;0.001</td>
<td>409</td>
<td>3.9%</td>
<td>1.08 + 0.63</td>
<td>p&lt;0.001</td>
<td>334</td>
<td>2.9%</td>
<td>1.07 + 0.81</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>FC&gt;90%</td>
<td>134</td>
<td>1.3%</td>
<td>1.81 + 0.56</td>
<td>FC&gt;90%</td>
<td>95</td>
<td>0.8%</td>
<td>2.12 + 0.81</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p&lt;0.001</td>
<td>124</td>
<td>1.2%</td>
<td>0.56 + 0.12</td>
<td>p&lt;0.001</td>
<td>156</td>
<td>1.3%</td>
<td>0.54 + 0.15</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Prostate</td>
<td>44</td>
<td>13</td>
<td>HG-U133B</td>
<td>Total</td>
<td>6691</td>
<td>29.6</td>
<td>1.02 + 0.32</td>
<td>Total</td>
<td>15886</td>
<td>70.4</td>
<td>Liu P 2006</td>
<td></td>
</tr>
<tr>
<td>Researcher</td>
<td>Tissue Type</td>
<td>Sample Code</td>
<td>Expression</td>
<td>FDR</td>
<td>p-value</td>
<td>Fold Change</td>
<td>Expression</td>
<td>FDR</td>
<td>p-value</td>
<td>Fold Change</td>
<td>Expression</td>
<td>FDR</td>
<td>p-value</td>
</tr>
<tr>
<td>------------</td>
<td>-------------</td>
<td>-------------</td>
<td>------------</td>
<td>-----</td>
<td>----------</td>
<td>-------------</td>
<td>------------</td>
<td>-----</td>
<td>----------</td>
<td>-------------</td>
<td>------------</td>
<td>-----</td>
<td>----------</td>
</tr>
<tr>
<td>Chandran R</td>
<td>Prostate</td>
<td>HG-U95Av2,B,C</td>
<td>1346</td>
<td>35.7%</td>
<td>&lt;0.001</td>
<td>1.04 ± 0.22</td>
<td>24223</td>
<td>64.3%</td>
<td>&lt;0.001</td>
<td>2.00 ± 0.46</td>
<td>Chandran R 2007</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kuner R</td>
<td>Prostate</td>
<td>Illumina HumanHT-12 v3.0</td>
<td>1275</td>
<td>26.2%</td>
<td>&lt;0.001</td>
<td>1.00 ± 0.02</td>
<td>35902</td>
<td>73.8%</td>
<td>&lt;0.001</td>
<td>1.09 ± 0.03</td>
<td>Kuner R 2012</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arredouani M</td>
<td>Prostate</td>
<td>HG-U133 Plus 2.0</td>
<td>1895</td>
<td>34.7%</td>
<td>&lt;0.001</td>
<td>1.02 ± 0.37</td>
<td>35660</td>
<td>65.3%</td>
<td>&lt;0.001</td>
<td>1.33 ± 1.09</td>
<td>Arredouani M 2009</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Best CJ</td>
<td>CRPC</td>
<td>AS-PCa</td>
<td>1054</td>
<td>47.5%</td>
<td>&lt;0.001</td>
<td>0.74 ± 0.72</td>
<td>11669</td>
<td>52.5%</td>
<td>&lt;0.001</td>
<td>1.23 ± 1.16</td>
<td>Best CJ 2005</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cai C</td>
<td>CRPC</td>
<td>AS-PCa</td>
<td>8942</td>
<td>49.2%</td>
<td>&lt;0.001</td>
<td>1.68 ± 2.51</td>
<td>9244</td>
<td>50.8%</td>
<td>&lt;0.001</td>
<td>1.85 ± 2.24</td>
<td>Cai C 2013</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>Tissue</td>
<td>Age</td>
<td>Condition</td>
<td>Affymetrix Platform</td>
<td>Total Count</td>
<td>Fold Change</td>
<td>Percent</td>
<td>SE</td>
<td>p-value</td>
<td>Total Count</td>
<td>Fold Change</td>
<td>Percent</td>
<td>SE</td>
</tr>
<tr>
<td>------------------</td>
<td>----------</td>
<td>-----</td>
<td>--------------------------------</td>
<td>------------------------</td>
<td>-------------</td>
<td>-------------</td>
<td>---------</td>
<td>-----------</td>
<td>-------------</td>
<td>-------------</td>
<td>-------------</td>
<td>---------</td>
<td>-----------</td>
</tr>
<tr>
<td>Kim WJ</td>
<td>Bladder</td>
<td>188</td>
<td>68</td>
<td>Illumina Human- v2</td>
<td>1057</td>
<td>47.6%</td>
<td>33.2</td>
<td>1.08 + 0.48</td>
<td>p&lt;0.001</td>
<td>7846</td>
<td>31.4%</td>
<td>48.2</td>
<td>1.00 + 0.04</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1361</td>
<td>21.9%</td>
<td></td>
<td>1.06 + 0.03</td>
<td>p&lt;0.001</td>
<td>827</td>
<td>9.2%</td>
<td>20.9</td>
<td>0.92 + 0.04</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dyrskjøt L</td>
<td>Bladder</td>
<td>41</td>
<td>14</td>
<td>HG-U133A</td>
<td>1895</td>
<td>34.7%</td>
<td>33.2</td>
<td>1.03 + 0.29</td>
<td>p&lt;0.001</td>
<td>1057</td>
<td>47.6%</td>
<td>33.2</td>
<td>1.08 + 0.48</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1361</td>
<td>21.9%</td>
<td></td>
<td>1.06 + 0.03</td>
<td>p&lt;0.001</td>
<td>827</td>
<td>9.2%</td>
<td>20.9</td>
<td>0.92 + 0.04</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urquidi V</td>
<td>Bladder</td>
<td>52</td>
<td>40</td>
<td>HG-U133 Plus 2.0</td>
<td>1938</td>
<td>35.2%</td>
<td>33.2</td>
<td>1.03 + 0.29</td>
<td>p&lt;0.001</td>
<td>1057</td>
<td>47.6%</td>
<td>33.2</td>
<td>1.08 + 0.48</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1361</td>
<td>21.9%</td>
<td></td>
<td>1.06 + 0.03</td>
<td>p&lt;0.001</td>
<td>827</td>
<td>9.2%</td>
<td>20.9</td>
<td>0.92 + 0.04</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zhang Z</td>
<td>Bladder</td>
<td>10</td>
<td>7</td>
<td>HG-U133 Plus 2.0</td>
<td>1938</td>
<td>35.2%</td>
<td>33.2</td>
<td>1.03 + 0.29</td>
<td>p&lt;0.001</td>
<td>1057</td>
<td>47.6%</td>
<td>33.2</td>
<td>1.08 + 0.48</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1361</td>
<td>21.9%</td>
<td></td>
<td>1.06 + 0.03</td>
<td>p&lt;0.001</td>
<td>827</td>
<td>9.2%</td>
<td>20.9</td>
<td>0.92 + 0.04</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gumz M</td>
<td>Renal</td>
<td>10</td>
<td>10</td>
<td>HG-U133A</td>
<td>1057</td>
<td>47.6%</td>
<td>33.2</td>
<td>1.03 + 0.29</td>
<td>p&lt;0.001</td>
<td>7790</td>
<td>44.4%</td>
<td>55.6</td>
<td>1.00 + 0.09</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1361</td>
<td>21.9%</td>
<td></td>
<td>1.06 + 0.03</td>
<td>p&lt;0.001</td>
<td>827</td>
<td>9.2%</td>
<td>20.9</td>
<td>0.92 + 0.04</td>
</tr>
</tbody>
</table>

* p-values are reported as p<0.001, p<0.01, or p<0.05 for fold change (FC) greater than 90% (FC>90%) or less than 10% (FC<10%).
<table>
<thead>
<tr>
<th>Renal</th>
<th>10</th>
<th>10</th>
<th>HG-U133B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>6691</td>
<td>29.6%</td>
<td>1.11 + 1.83</td>
</tr>
<tr>
<td>p&lt;0.001</td>
<td>1603</td>
<td>15.2%</td>
<td>1.06 + 2.63</td>
</tr>
<tr>
<td>FC&gt;90%</td>
<td>452</td>
<td>4.3%</td>
<td>3.99 + 4.07</td>
</tr>
<tr>
<td>FC&lt;10%</td>
<td>379</td>
<td>3.6%</td>
<td>0.34 + 0.12</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Yusenko M</th>
<th>Renal</th>
<th>62</th>
<th>5</th>
<th>HG-U133 Plus 2.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>1895</td>
<td>34.7%</td>
<td>1.07 + 0.97</td>
<td>Total</td>
</tr>
<tr>
<td>p&lt;0.001</td>
<td>1212</td>
<td>6.4%</td>
<td>1.97 + 1.61</td>
<td>p&lt;0.001</td>
</tr>
<tr>
<td>FC&gt;90%</td>
<td>559</td>
<td>2.9%</td>
<td>2.81 + 2.04</td>
<td>FC&gt;90%</td>
</tr>
<tr>
<td>FC&lt;10%</td>
<td>51</td>
<td>0.3%</td>
<td>0.28 + 0.14</td>
<td>FC&lt;10%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Jones J</th>
<th>Renal</th>
<th>69</th>
<th>23</th>
<th>HG-U133A</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>1057</td>
<td>47.6%</td>
<td>1.25 + 0.74</td>
<td>Total</td>
</tr>
<tr>
<td>p&lt;0.001</td>
<td>6659</td>
<td>63.0%</td>
<td>1.42 + 0.65</td>
<td>p&lt;0.001</td>
</tr>
<tr>
<td>FC&gt;90%</td>
<td>1164</td>
<td>11.0%</td>
<td>2.21 + 2.17</td>
<td>FC&gt;90%</td>
</tr>
<tr>
<td>FC&lt;10%</td>
<td>411</td>
<td>3.9%</td>
<td>0.51 + 0.19</td>
<td>FC&lt;10%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Lenburg M</th>
<th>Renal</th>
<th>12</th>
<th>4</th>
<th>HG-U133A</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>1057</td>
<td>47.6%</td>
<td>1.24 + 1.55</td>
<td>Total</td>
</tr>
<tr>
<td>p&lt;0.001</td>
<td>5</td>
<td>0.0%</td>
<td>1.06 + 0.59</td>
<td>p&lt;0.001</td>
</tr>
<tr>
<td>FC&gt;90%</td>
<td>1</td>
<td>0.0%</td>
<td>1.95</td>
<td>FC&gt;90%</td>
</tr>
<tr>
<td>FC&lt;10%</td>
<td>2</td>
<td>0.0%</td>
<td>0.62 + 0.06</td>
<td>FC&lt;10%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Renal</th>
<th>12</th>
<th>4</th>
<th>HG-U133B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>6691</td>
<td>29.6%</td>
<td>1.09 + 0.85</td>
</tr>
</tbody>
</table>

196
<p>| | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>p&lt;0.001</td>
<td>p&lt;0.001</td>
<td>p&lt;0.001</td>
<td>p&lt;0.001</td>
<td>p&lt;0.001</td>
<td>p&lt;0.001</td>
<td>p&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>FC&lt;10%</td>
<td>FC&lt;10%</td>
<td>FC&lt;10%</td>
<td>FC&lt;10%</td>
<td>FC&lt;10%</td>
<td>FC&lt;10%</td>
<td>FC&lt;10%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bong I Breast 43</td>
<td>1057</td>
<td>0.5%</td>
<td>1.60 + 1.34</td>
<td>69</td>
<td>0.4%</td>
<td>1.91 + 2.22</td>
<td></td>
</tr>
<tr>
<td></td>
<td>606</td>
<td>0.5%</td>
<td>2.66 + 0.99</td>
<td>27</td>
<td>0.2%</td>
<td>4.05 + 2.24</td>
<td></td>
</tr>
<tr>
<td></td>
<td>220</td>
<td>0.2%</td>
<td>0.33 + 0.17</td>
<td>26</td>
<td>0.2%</td>
<td>0.42 + 0.10</td>
<td></td>
</tr>
<tr>
<td></td>
<td>352</td>
<td>3.3%</td>
<td>0.52 + 0.15</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FC&gt;90%</td>
<td>43</td>
<td>0.2%</td>
<td>1.60 + 1.34</td>
<td>69</td>
<td>0.4%</td>
<td>1.91 + 2.22</td>
<td></td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>0.2%</td>
<td>2.66 + 0.99</td>
<td>27</td>
<td>0.2%</td>
<td>4.05 + 2.24</td>
<td></td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>0.2%</td>
<td>0.33 + 0.17</td>
<td>26</td>
<td>0.2%</td>
<td>0.42 + 0.10</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Casey T Breast 56</td>
<td>1057</td>
<td>47.6%</td>
<td>0.96 + 0.56</td>
<td>11643</td>
<td>52.4%</td>
<td>1.14 + 2.94</td>
<td></td>
</tr>
<tr>
<td></td>
<td>606</td>
<td>5.7%</td>
<td>1.17 + 1.80</td>
<td>500</td>
<td>4.3%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>220</td>
<td>2.1%</td>
<td>2.27 + 2.26</td>
<td>131</td>
<td>1.1%</td>
<td>2.99 + 5.34</td>
<td></td>
</tr>
<tr>
<td></td>
<td>352</td>
<td>3.3%</td>
<td>0.52 + 0.15</td>
<td>346</td>
<td>3.0%</td>
<td>0.48 + 0.18</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clarke C Breast 104</td>
<td>1895</td>
<td>34.7%</td>
<td>1.00 + 0.16</td>
<td>35660</td>
<td>65.3%</td>
<td>1.20 + 19.08</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5279</td>
<td>27.9%</td>
<td>1.06 + 0.29</td>
<td>6679</td>
<td>18.7%</td>
<td>1.04 + 0.32</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2098</td>
<td>11.1%</td>
<td>1.31 + 0.25</td>
<td>2505</td>
<td>7.0%</td>
<td>1.34 + 0.24</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1264</td>
<td>6.7%</td>
<td>0.71 + 0.13</td>
<td>1769</td>
<td>5.0%</td>
<td>0.65 + 0.16</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lopez FJ Breast 54</td>
<td>1895</td>
<td>34.7%</td>
<td>1.00 + 0.10</td>
<td>35660</td>
<td>65.3%</td>
<td>17.0%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3874</td>
<td>20.4%</td>
<td>0.97 + 0.17</td>
<td>6072</td>
<td>0.99 + 0.18</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>907</td>
<td>4.8%</td>
<td>1.19 + 0.11</td>
<td>1521</td>
<td>4.3%</td>
<td>1.19 + 0.12</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1466</td>
<td>7.7%</td>
<td>0.81 + 0.08</td>
<td>1812</td>
<td>5.1%</td>
<td>0.77 + 0.11</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Richardson Breast 43</td>
<td>1895</td>
<td>34.7%</td>
<td>1.20 + 19.08</td>
<td>35660</td>
<td>65.3%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3874</td>
<td>20.4%</td>
<td>0.97 + 0.17</td>
<td>6072</td>
<td>0.99 + 0.18</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>907</td>
<td>4.8%</td>
<td>1.19 + 0.11</td>
<td>1521</td>
<td>4.3%</td>
<td>1.19 + 0.12</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1466</td>
<td>7.7%</td>
<td>0.81 + 0.08</td>
<td>1812</td>
<td>5.1%</td>
<td>0.77 + 0.11</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bong I 2009 Breast 43</td>
<td>1057</td>
<td>0.5%</td>
<td>1.60 + 1.34</td>
<td>69</td>
<td>0.4%</td>
<td>1.91 + 2.22</td>
<td></td>
</tr>
<tr>
<td></td>
<td>606</td>
<td>0.5%</td>
<td>2.66 + 0.99</td>
<td>27</td>
<td>0.2%</td>
<td>4.05 + 2.24</td>
<td></td>
</tr>
<tr>
<td></td>
<td>220</td>
<td>0.2%</td>
<td>0.33 + 0.17</td>
<td>26</td>
<td>0.2%</td>
<td>0.42 + 0.10</td>
<td></td>
</tr>
<tr>
<td></td>
<td>352</td>
<td>3.3%</td>
<td>0.52 + 0.15</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Author</td>
<td>Tumor</td>
<td>Sample Size</td>
<td>Platform</td>
<td>Total Count</td>
<td>%</td>
<td>SEM</td>
<td>p-value</td>
</tr>
<tr>
<td>--------</td>
<td>-------</td>
<td>-------------</td>
<td>----------</td>
<td>-------------</td>
<td>----</td>
<td>-----</td>
<td>---------</td>
</tr>
<tr>
<td>Chen DT</td>
<td>Breast</td>
<td>42</td>
<td>143</td>
<td>HG-U133 Plus 2.0</td>
<td>1895</td>
<td>34.7</td>
<td>1.00 + 0.06</td>
</tr>
<tr>
<td>Landi M</td>
<td>Lung</td>
<td>58</td>
<td>49</td>
<td>HG-U133A</td>
<td>1057</td>
<td>47.6</td>
<td>1.00 + 0.02</td>
</tr>
<tr>
<td>Girard A</td>
<td>Lung</td>
<td>30</td>
<td>20</td>
<td>HG-U133A</td>
<td>1057</td>
<td>47.6</td>
<td>1.00 + 0.05</td>
</tr>
<tr>
<td>Girard B</td>
<td>Lung</td>
<td>30</td>
<td>20</td>
<td>HG-U133B</td>
<td>6691</td>
<td>29.6</td>
<td>1.00 + 0.04</td>
</tr>
<tr>
<td>Hou J</td>
<td>Lung</td>
<td>91</td>
<td>65</td>
<td>HG-U133 Plus 2.0</td>
<td>1895</td>
<td>34.7</td>
<td>-0.24 +</td>
</tr>
<tr>
<td>Study</td>
<td>Tissue</td>
<td>HG-U133 Plus 2.0</td>
<td>HG-U133A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------------</td>
<td>--------</td>
<td>------------------</td>
<td>----------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sanchez-Palencia</td>
<td>Lung</td>
<td>46</td>
<td>45</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>p&lt;0.001</td>
<td>1040</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>FC&gt;90%</td>
<td>18</td>
<td>0.1%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>FC&lt;10%</td>
<td>359</td>
<td>1.9%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>1895</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>p&lt;0.001</td>
<td>8903</td>
<td>47.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>FC&gt;90%</td>
<td>2922</td>
<td>15.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>FC&lt;10%</td>
<td>1718</td>
<td>9.1%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>35660</td>
<td>65.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sanchez-Palencia A</td>
<td>2010</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Okayama H</td>
<td>Lung</td>
<td>226</td>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>p&lt;0.001</td>
<td>1895</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>FC&gt;90%</td>
<td>8134</td>
<td>42.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>FC&lt;10%</td>
<td>1491</td>
<td>7.9%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>FC&lt;10%</td>
<td>1140</td>
<td>6.0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>35660</td>
<td>65.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Okayama H 2012</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Galamb O</td>
<td>Colorectal</td>
<td>15</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>p&lt;0.001</td>
<td>1895</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>FC&gt;90%</td>
<td>672</td>
<td>3.5%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>FC&lt;10%</td>
<td>366</td>
<td>1.9%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>FC&lt;10%</td>
<td>83</td>
<td>0.4%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>35660</td>
<td>65.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Galamb 2008</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alhopuro P</td>
<td>Colorectal</td>
<td>34</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>p&lt;0.001</td>
<td>1057</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>FC&gt;90%</td>
<td>2471</td>
<td>23.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>FC&lt;10%</td>
<td>1207</td>
<td>11.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>FC&lt;10%</td>
<td>423</td>
<td>4.0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>11643</td>
<td>52.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Alhopuro P 2011</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Name</td>
<td>Tissue</td>
<td>Sample Size</td>
<td>Platform</td>
<td>Total Genes</td>
<td>Genes FC&gt;90%</td>
<td>Genes FC&lt;10%</td>
<td>p-Value</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
<td>-------------</td>
<td>-------------------</td>
<td>-------------</td>
<td>--------------</td>
<td>--------------</td>
<td>---------</td>
</tr>
<tr>
<td>Galamb O</td>
<td>Colorectal</td>
<td>27</td>
<td>HG-U133 Plus 2.0</td>
<td>1895</td>
<td>3</td>
<td>3</td>
<td>3.18</td>
</tr>
<tr>
<td></td>
<td></td>
<td>38</td>
<td></td>
<td>6026</td>
<td>p&lt;0.001</td>
<td>p&lt;0.001</td>
<td>1.43</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1842</td>
<td>FC&gt;90%</td>
<td>FC&lt;10%</td>
<td>9.7%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>871</td>
<td>FC&gt;90%</td>
<td>FC&lt;10%</td>
<td>4.6%</td>
</tr>
<tr>
<td>Skrzypczak</td>
<td>Colorectal</td>
<td>81</td>
<td>HG-U133 Plus 2.0</td>
<td>1301</td>
<td>2</td>
<td>2</td>
<td>39.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>24</td>
<td></td>
<td>5138</td>
<td>p&lt;0.001</td>
<td>p&lt;0.001</td>
<td>10.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1317</td>
<td>FC&gt;90%</td>
<td>FC&lt;10%</td>
<td>5.2%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>679</td>
<td>FC&gt;90%</td>
<td>FC&lt;10%</td>
<td>0.0%</td>
</tr>
<tr>
<td>Ryan BM</td>
<td>Colorectal</td>
<td>56</td>
<td>HG-U133A</td>
<td>1057</td>
<td>2</td>
<td>2</td>
<td>17.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>55</td>
<td></td>
<td>1880</td>
<td>p&lt;0.001</td>
<td>p&lt;0.001</td>
<td>7.9%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>836</td>
<td>FC&gt;90%</td>
<td>FC&lt;10%</td>
<td>5.9%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>627</td>
<td>FC&gt;90%</td>
<td>FC&lt;10%</td>
<td>0.0%</td>
</tr>
<tr>
<td>Moreno CS</td>
<td>Ovarian</td>
<td>33</td>
<td>HG-U95Av2</td>
<td>6010</td>
<td>1</td>
<td>1</td>
<td>0.0%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10</td>
<td></td>
<td>2498</td>
<td>p&lt;0.001</td>
<td>p&lt;0.001</td>
<td>0.0%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>439</td>
<td>FC&gt;90%</td>
<td>FC&lt;10%</td>
<td>4.2%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>615</td>
<td>FC&gt;90%</td>
<td>FC&lt;10%</td>
<td>5.8%</td>
</tr>
<tr>
<td>Hendrix ND</td>
<td>Ovarian</td>
<td>99</td>
<td>HG-U133A</td>
<td>1057</td>
<td>2</td>
<td>2</td>
<td>23.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
<td></td>
<td>2498</td>
<td>p&lt;0.001</td>
<td>p&lt;0.001</td>
<td>0.0%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>439</td>
<td>FC&gt;90%</td>
<td>FC&lt;10%</td>
<td>4.2%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>615</td>
<td>FC&gt;90%</td>
<td>FC&lt;10%</td>
<td>5.8%</td>
</tr>
<tr>
<td>Mok SC</td>
<td>Ovarian</td>
<td>53</td>
<td>10</td>
<td>HG-U133 Plus 2.0</td>
<td>Total</td>
<td>1895</td>
<td>53.1%</td>
</tr>
<tr>
<td>--------</td>
<td>---------</td>
<td>----</td>
<td>----</td>
<td>----------------</td>
<td>-------</td>
<td>------</td>
<td>--------</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td>24.6%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>228</td>
<td>2.9%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>663</td>
<td>3.5%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>King ER</th>
<th>Ovarian</th>
<th>35</th>
<th>6</th>
<th>HG-U133 Plus 2.0</th>
<th>Total</th>
<th>1895</th>
<th>53.1%</th>
<th>1.22 + 1.79</th>
<th>Total</th>
<th>35660</th>
<th>65.3%</th>
<th>2.21 + 2.87</th>
<th>King ER 2011</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td>20.0%</td>
<td>2.79 + 5.17</td>
<td></td>
<td>4263</td>
<td>12.0%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3794</td>
<td>9.3%</td>
<td>4.24 + 7.32</td>
<td></td>
<td>1122</td>
<td>3.1%</td>
<td>4.41 + 4.94</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1758</td>
<td>0.1%</td>
<td>0.29 + 0.07</td>
<td></td>
<td>28</td>
<td>0.1%</td>
<td>0.26 + 0.08</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mostavi ZB</th>
<th>Ovarian</th>
<th>32</th>
<th>35</th>
<th>HG-Focus</th>
<th>Total</th>
<th>3933</th>
<th>81.7%</th>
<th>1.21 + 2.88</th>
<th>Total</th>
<th>4813</th>
<th>55.0%</th>
<th>1.42 + 3.00</th>
<th>Mostavi ZB 2008</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td>19.5%</td>
<td>2.16 + 7.68</td>
<td></td>
<td>582</td>
<td>12.1%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>765</td>
<td>6.3%</td>
<td>5.10 + 13.04</td>
<td></td>
<td>138</td>
<td>2.9%</td>
<td>4.11 + 5.31</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>247</td>
<td>0.1%</td>
<td>0.44 + 0.15</td>
<td></td>
<td>285</td>
<td>5.9%</td>
<td>0.39 + 0.17</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Wang Q</th>
<th>Gastric</th>
<th>12</th>
<th>15</th>
<th>HG-U133 Plus 2.0</th>
<th>Total</th>
<th>1895</th>
<th>34.7%</th>
<th>1.04 + 1.06</th>
<th>Total</th>
<th>35660</th>
<th>65.3%</th>
<th>0.75 + 0.96</th>
<th>Wang Q 2012</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td>2.5%</td>
<td>1.03 + 1.27</td>
<td></td>
<td>727</td>
<td>2.0%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>479</td>
<td>0.6%</td>
<td>2.57 + 1.91</td>
<td></td>
<td>99</td>
<td>0.3%</td>
<td>2.67 + 1.46</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>110</td>
<td>0.9%</td>
<td>0.36 + 0.12</td>
<td></td>
<td>423</td>
<td>1.2%</td>
<td>0.32 + 0.14</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Eftang LL</th>
<th>Gastric</th>
<th>20</th>
<th>20</th>
<th>Illumina Human-12 v3</th>
<th>Total</th>
<th>1290</th>
<th>35.7%</th>
<th>1.00 + 0.03</th>
<th>Total</th>
<th>23248</th>
<th>64.3%</th>
<th>0.94 + 0.13</th>
<th>Eftang LL 2013</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>9</td>
<td>4.6%</td>
<td>1.00 + 0.08</td>
<td></td>
<td>690</td>
<td>3.0%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>589</td>
<td>2.4%</td>
<td>1.07 + 0.04</td>
<td></td>
<td>209</td>
<td>0.9%</td>
<td>1.09 + 0.07</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>306</td>
<td>2.0%</td>
<td>0.92 + 0.05</td>
<td></td>
<td>450</td>
<td>1.9%</td>
<td>0.87 + 0.09</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Type</td>
<td>Count</td>
<td>95% CI</td>
<td>Count</td>
<td>95% CI</td>
<td>Count</td>
<td>95% CI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>----------</td>
<td>-------</td>
<td>--------</td>
<td>-------</td>
<td>--------</td>
<td>-------</td>
<td>--------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wang G</td>
<td>Gastric</td>
<td>134</td>
<td>134</td>
<td>HG-U133A</td>
<td>Total 1057</td>
<td>47.6%</td>
<td>1.00 + 0.05</td>
<td>Total 11643</td>
<td>52.4%</td>
<td>0.99 + 0.07</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p&lt;0.001</td>
<td>4645</td>
<td>12.8%</td>
<td>1.02 + 0.06</td>
<td>p&lt;0.001</td>
<td>4951</td>
<td>10.8%</td>
<td>1.09 + 0.06</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>FC&gt;90%</td>
<td>1357</td>
<td>6.8%</td>
<td>1.08 + 0.04</td>
<td>FC&gt;90%</td>
<td>783</td>
<td>6.7%</td>
<td>1.06 + 0.05</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p&lt;0.001</td>
<td>721</td>
<td></td>
<td>0.93 + 0.04</td>
<td>p&lt;0.001</td>
<td>1263</td>
<td></td>
<td>0.92 + 0.06</td>
<td></td>
</tr>
<tr>
<td>Holbrook JD</td>
<td>Gastric</td>
<td>50</td>
<td>49</td>
<td>Illumina Human-12 v3</td>
<td>Total 1291</td>
<td>35.7%</td>
<td>1.32 + 43.89</td>
<td>Total 23244</td>
<td>64.3%</td>
<td>4.08 + 110.14</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p&lt;0.001</td>
<td>2352</td>
<td>18.2%</td>
<td>1.68 + 13.08</td>
<td>p&lt;0.001</td>
<td>2015</td>
<td>8.7%</td>
<td>13.67 + 229.93</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>FC&gt;90%</td>
<td>442</td>
<td>3.4%</td>
<td>5.20 + 27.25</td>
<td>FC&gt;90%</td>
<td>460</td>
<td>2.0%</td>
<td>-2.32 + 15.42</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p&lt;0.001</td>
<td>54</td>
<td>0.4%</td>
<td>-7.07 + 34.73</td>
<td>p&lt;0.001</td>
<td>298</td>
<td>1.3%</td>
<td>0.85 + 0.10</td>
<td></td>
</tr>
<tr>
<td>Cho JY</td>
<td>Gastric</td>
<td>65</td>
<td>19</td>
<td>Illumina HumanWG-6 v3</td>
<td>Total 1179</td>
<td>32.8%</td>
<td>1.00 + 0.05</td>
<td>Total 24169</td>
<td>67.2%</td>
<td>11.9%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p&lt;0.001</td>
<td>2033</td>
<td>17.2%</td>
<td>1.03 + 0.09</td>
<td>p&lt;0.001</td>
<td>2873</td>
<td>4.4%</td>
<td>1.10 + 0.07</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>FC&gt;90%</td>
<td>1127</td>
<td>9.6%</td>
<td>1.09 + 0.06</td>
<td>FC&gt;90%</td>
<td>153</td>
<td>4.4%</td>
<td>0.85 + 0.10</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p&lt;0.001</td>
<td>430</td>
<td>3.6%</td>
<td>0.90 + 0.06</td>
<td>p&lt;0.001</td>
<td>1055</td>
<td>4.4%</td>
<td>0.85 + 0.10</td>
<td></td>
</tr>
<tr>
<td>D'Errico M</td>
<td>Gastric</td>
<td>39</td>
<td>30</td>
<td>HG-U133 Plus 2.0</td>
<td>Total 1895</td>
<td>34.7%</td>
<td>0.95 + 0.61</td>
<td>Total 35660</td>
<td>65.3%</td>
<td>13.8%</td>
<td>0.95 + 1.24</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p&lt;0.001</td>
<td>4086</td>
<td>21.6%</td>
<td>1.37 + 1.15</td>
<td>p&lt;0.001</td>
<td>4918</td>
<td>2.9%</td>
<td>2.34 + 2.14</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>FC&gt;90%</td>
<td>1806</td>
<td>9.5%</td>
<td>2.03 + 1.43</td>
<td>FC&gt;90%</td>
<td>1032</td>
<td>2.9%</td>
<td>2.34 + 2.14</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p&lt;0.001</td>
<td>475</td>
<td>2.5%</td>
<td>0.40 + 0.11</td>
<td>p&lt;0.001</td>
<td>1792</td>
<td>5.0%</td>
<td>0.37 + 0.13</td>
<td></td>
</tr>
</tbody>
</table>
6.3.2 The proportion of N6-adenosine methylated RNAs
Dominissini et al. found that up to 12,769 mRNAs were susceptible to the m6A modification (Dominissini et al. 2012). Annotation of our mRNA datasets using these loci found that m6A(+) mRNAs represented between 26.3% (renal) and 59.4% (bladder) of those with significant aberrant expression (p<0.001 between cancer and controls) in cancer (Table 20, and presented as m6A:non-m6A ratios in Figure 39). Colorectal, bladder and prostate cancers have more predicted m6A(+) (ratio >1), whilst ovarian, gastric (p=0.02), breast (p=0.01), lung (p=0.01), renal (p=0.009) cancers have significantly more m6A(-) mRNAs (ratio <1).

6.3.3 Fold changes of m6A susceptible RNAs
All m6A(+) and m6A(-) fold change (FC) and p-value (Malignant Vs Benign) were plotted and visualised through volcano plots (Figures 40 (prostate) and 41 (other cancers)). Only mRNA shared in all datasets within each individual cancer were included. Little difference was seen between m6A(+) and m6A(-) within each cancer type or across all eight cancers.

The expression of m6A(+) and m6A(-) were upregulated (FC>1), except for lung cancer, which had a mean FC of 0.77 in both m6A(+) and m6A(-). Little difference was seen across the cancer types and the overall mean FC was not significant between aberrantly expressed m6A(+) and m6A(-) (Figure 42).
Figure 39. Proportion of differentially expressed RNA susceptible to N6-adenosine methylation.

For each cancer, the number of differentially expressed mRNAs is presented as a ratio of those susceptible to m6A divided by those not susceptible to this modification. Within PCa (blue), there is slightly more m6A(+) compared to m6a(-).

*p-value= %m6A Vs %non-m6A (Student’s t-test)
mRNAs shared in all seven prostate cancer and two castration-resistant prostate cancer datasets (n=9) were analysed. Genes with aberrant expressions (p<0.01) are shown. a) all mRNAs, n= 3097; b) m6A(+) only, n= 1993.

**Figure 40.** Volcano plots of a) all mRNAs and b) m6A susceptible mRNAs within prostate cancer.
Figure 41. Volcano plots of expression of all mRNAs and m6A(+) within a) all cancers, b) bladder, c) kidney, d) breast, e) lung, f) ovarian, g) colorectal, h) gastric cancer.

a)  

*All 571 genes were m6A+*
b) mRNAs shared in all 4 bladder datasets (n=6656)

m6A(+) shared in all 4 bladder datasets (n=3773)
mRNAs shared in all 6 kidney datasets (n=4328)

m6A(+) shared in all 6 kidney datasets (n=2365)
mRNAs shared in all 6 breast datasets (n=13397)

m6A(+) shared in all 6 breast datasets (n=6099)
mRNAs shared in all 6 lung datasets (n=4328)

m6A(+) shared in all 6 lung datasets (n=2365)
mRNAs shared in all 5 ovarian datasets (n=6707)

m6A(+) shared in all 5 ovarian datasets (n=3228)
mRNAs shared in all colorectal datasets (n=10079)

m6A(+) shared in all 5 colorectal datasets (n=5610)
mRNAs shared in all 6 gastric datasets (n=10369)

m6A(+) shared in all 6 gastric datasets (n=5414)
Figure 42. Average fold change of aberrantly expressed RNAs within each cancer and across all cancer types.

The mean (SD, Standard Deviation) fold changes between m6A(+) and m6A(-) within each cancer and across all 8 cancers was not significant.

*p-value= m6A Vs non-m6A (Student’s t-test); m6A(+), N6-methyladenosine susceptible; m6A(-), N6-methyladenosine non-susceptible
6.3.4 Fold changes of the most differentially expressed RNAs

In an attempt to identify potentially the most important differentially expressed m6A susceptible RNAs, those with the highest or lowest (>90% or <10% thresholds) FC in each cancer were selected. The mean percentage of differentially expressed up- and downregulated m6A(+) mRNAs ranged between 12.5-70.4% and 0-66.7% respectively (Figure 43, presented as m6A:non-m6A ratios). With regards to upregulated RNAs, colorectal (p<0.001) and gastric cancer (p=0.004) had significantly more m6A(+) upregulated mRNAs, and prostate (p=0.02) and renal (p=0.002) cancers had significantly more m6A(-) RNAs (Figure 43a). In contrast, when analysing downregulated mRNAs, all cancers, specifically breast (p=0.02), colorectal (p<0.001), lung (p<0.001), renal (p<0.001) and gastric (p<0.001) cancers had significantly more m6A(-) mRNAs (Figure 43b). However, the mean up- or downregulated FC did not differ between the m6A susceptible and non-susceptible groups (Figure 44).
Figure 43. Proportion of differentially expressed RNA susceptible to N6-adenosine methylation within a) top and b) bottom 10%-fold change percentile.

For each cancer, the number of differentially expressed mRNAs is presented as a ratio of those susceptible to m6A divided by those not susceptible to this modification. There were more upregulated (a) and downregulated m6A(-) RNAs compared to m6A(+) in PCa (blue).
<table>
<thead>
<tr>
<th>Cancer</th>
<th>Prostate</th>
<th>Breast</th>
<th>Bladder</th>
<th>Ovarian</th>
<th>Colorectal</th>
<th>Lung</th>
<th>Renal</th>
<th>Gastric</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>%m6A:%non-m6A Mean (SD)</td>
<td>0.94±0.42</td>
<td>0.85±0.18</td>
<td>0.81±0.80</td>
<td>0.69±0.69</td>
<td>0.55±0.21</td>
<td>0.53±0.18</td>
<td>0.50±0.12</td>
<td>0.40±0.16</td>
<td>0.67±0.41</td>
</tr>
<tr>
<td>p-value</td>
<td>0.22</td>
<td><strong>0.02</strong></td>
<td>0.15</td>
<td>0.47</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>
a) Upregulated Fold change

<table>
<thead>
<tr>
<th>Cancer Type</th>
<th>m6A(+) Mean (SD)</th>
<th>m6A(-) Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prostate</td>
<td>1.76 ±0.65</td>
<td>1.65 ±0.52</td>
</tr>
<tr>
<td>Bladder</td>
<td>2.91 ±0.80</td>
<td>3.75 ±0.72</td>
</tr>
<tr>
<td>Renal</td>
<td>3.82 ±6.18</td>
<td>2.30 ±2.28</td>
</tr>
<tr>
<td>Lung</td>
<td>1.94 ±1.38</td>
<td>1.90 ±1.09</td>
</tr>
<tr>
<td>Breast</td>
<td>3.54 ±2.27</td>
<td>3.41 ±2.24</td>
</tr>
<tr>
<td>Ovarian</td>
<td>1.66 ±0.72</td>
<td>2.29 ±1.54</td>
</tr>
<tr>
<td>Colorectal</td>
<td>2.17 ±1.61</td>
<td>3.66 ±4.95</td>
</tr>
<tr>
<td>Gastric</td>
<td>2.42 ±2.55</td>
<td>2.58 ±2.20</td>
</tr>
<tr>
<td>All</td>
<td>2.74 ±2.74</td>
<td>2.58 ±2.20</td>
</tr>
</tbody>
</table>

P-Value m6A(+) V m6A(-): 0.89
Figure 44. The mean fold changes in a) up and b) downregulated N6-methyladenosine susceptible RNAs in each cancer type.

The mean fold change (up- or downregulated FC) did not vary between m6A(+) and m6A(-) in the 8 common cancers (p>0.05).
6.3.5 Functional annotation of m6A susceptible RNAs

6.3.5.1 Individual cancers

To determine the impact of m6A in each cancer we selected differentially expressed mRNAs, susceptible to m6A and common to ≥50% of the datasets within each individual cancer (number of predicted m6A transcripts: PCa/CRPC, n=733; bladder, n=1,239; renal, n=1,708; breast, n=2,505; lung, n=4,281; colorectal, n=1,759; ovarian, n=2,859; gastric, n=1,588 (Figures 45 (Prostate) and 46 (other cancers)). We analysed their functional roles using gene set enrichment analysis (GSEA) within DAVID (https://david.ncifcrf.gov/). DAVID (enrichment score (ES) ≥1.5) identified tumour specific and shared pathways between the individual cancers. Within PCa, predicted m6A(+) candidates are involved in a range of oncogenic pathways including regulation of transcription and apoptosis, and nuclear division (Figure 47). Common clusters that are shared in ≥50% of datasets from each cancer type include regulation of apoptosis, nucleotide-binding, regulation of transcription, and protein transport/localization (Figure 48).
Figure 45. Functional annotation of aberrantly expressed genes shared in prostate cancer.

Genes shared in at least 50% of the nine PCa/CRPC datasets were clustered (DAVID). The number of genes (m6A(+) and m6A(-)) and significance within each cluster is shown. m6A(+) in PCa appear to have oncogenic roles including regulation of apoptosis, nuclear division and regulation of transcription.
Figure 46. Functional annotation of aberrantly expressed m6A(+) and m6A(-) in a) bladder, b) kidney, c) breast, d) lung, e) ovarian, f) colorectal, g) gastric cancer.
g) Gastric Cancer m6A(+) n=1588
- Negative regulation of transcription: 81 genes, p=0.008
- Protein localization: 112 genes, p=0.002
- Nuclear mRNA splicing: 25 genes, p=6.2E-04
- mRNA processing: 41 genes, p=4.0E-04
- Nuclear import: 22 genes, p=6.7E-05
- Vasculature development: 45 genes, p=6.0E-05
- Regulation of apoptosis: 112 genes, p=4.7E-05
- Regulation of ubiquitin-protein ligase activity: 25 genes, p=1.5E-07
- Protein complex assembly: 84 genes, p=6.4E-07
- Nucleotide-binding: 229 genes, p=4.2E-15
- Nuclear division: 74 genes, p=1.7E-22

Gastric Cancer m6A(-) n=429
- Gluconeogenesis: 5 genes, p=0.003
- Chemotaxis: 11 genes, p=0.003
- Fibronectin type II: 10 genes, p=0.04
- Regulation of cell migration: 11 genes, p=0.008
- Nucleotide processing: 11 genes, p=0.02
- Glucose metabolic process: 11 genes, p=0.004
- Regulation of cell division: 6 genes, p=0.005
- Polysaccharide binding: 14 genes, p=1.7E-04
- Vasculature development: 20 genes, p=1.2E-05
- Hydroxylation: 8 genes, p=1.5E-04
- Extracellular matrix: 28 genes, p=2.9E-12
Figure 47. Annotation clusters common in each cancer type.

Clusters common in all cancers were searched. A total of four clusters were found to be shared in all of the eight common cancers. These are all important pathways in oncogenesis.
6.3.5.2 Between cancers

To look at features common to cancer, we then selected aberrantly expressed mRNAs susceptible to m6A that were shared in at least one dataset from each cancer. In total, 689 mRNAs were selected, and GSEA identified key roles in oncogenic pathways such as cell cycle regulation (78 genes, ES= 27.5, p=6.3E-29), mRNA processing (30 genes, ES= 3.8, p=1.2E-04) and apoptosis (52 genes, ES= 2.5, p=0.003). Again, when analysing m6A(-) transcripts, functional clusters were less oncogenic and none were related to RNA activity (Figure 48). A total of 3 clusters were shared in both the m6A(+) and m6A(-) groups and there were significantly more m6A(+) RNAs involved in two of the processes; cell cycle process, p=0.002 and organelle lumen activity p<0.001 (Figure 49 and Table 22).
Figure 48. Functional annotation of aberrantly expressed genes shared in all cancer.

Genes shared in at least one dataset from each cancer were clustered. The number of genes and significance within each cluster is shown. m6A(+) common in all right cancers appear to function in mechanism associated with oncogenesis including apoptosis, mRNA processing/transport, nuclear division and cell cycle regulation.
Functional activities common in both m6A susceptible and non-susceptible transcripts were searched. A total of three processes were found. There were significantly more m6A susceptible transcripts involved in cell cycle process \((p=0.002)\) and organelle lumen activity \((p=0.0006)\).

**Figure 49.** Functional activities common in both m6A susceptible and non-susceptible transcripts.
<table>
<thead>
<tr>
<th>Cancer</th>
<th>Clusters</th>
<th>m6A(+)</th>
<th>m6A(-)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prostate</td>
<td>Extracellular matrix</td>
<td>27</td>
<td>32</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>LIM domain</td>
<td>9</td>
<td>9</td>
<td>0.582</td>
</tr>
<tr>
<td></td>
<td>Vesicle</td>
<td>49</td>
<td>32</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>Myofibril</td>
<td>16</td>
<td>9</td>
<td>0.767</td>
</tr>
<tr>
<td>Bladder</td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Renal</td>
<td>Vacuole</td>
<td>42</td>
<td>15</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>SH2 domain</td>
<td>23</td>
<td>9</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>Membrane organization</td>
<td>65</td>
<td>17</td>
<td>0.288</td>
</tr>
<tr>
<td></td>
<td>Cell fraction</td>
<td>148</td>
<td>32</td>
<td>0.6</td>
</tr>
<tr>
<td></td>
<td>Positive regulation apoptosis</td>
<td>88</td>
<td>16</td>
<td>0.892</td>
</tr>
<tr>
<td>Breast</td>
<td>Nucleotide-binding</td>
<td>430</td>
<td>230</td>
<td>0.0001</td>
</tr>
<tr>
<td></td>
<td>Regulation cell migration</td>
<td>39</td>
<td>23</td>
<td>0.201</td>
</tr>
<tr>
<td></td>
<td>Vasculature development</td>
<td>51</td>
<td>33</td>
<td>0.271</td>
</tr>
<tr>
<td>Lung</td>
<td>Cell fraction</td>
<td>169</td>
<td>119</td>
<td>0.0001</td>
</tr>
<tr>
<td></td>
<td>SH3 domain</td>
<td>40</td>
<td>29</td>
<td>0.035</td>
</tr>
<tr>
<td></td>
<td>Nucleotide-binding</td>
<td>430</td>
<td>160</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>Protein kinase activity</td>
<td>120</td>
<td>67</td>
<td>0.105</td>
</tr>
<tr>
<td></td>
<td>Regulation cell migration</td>
<td>39</td>
<td>24</td>
<td>0.17</td>
</tr>
<tr>
<td></td>
<td>Vasculature development</td>
<td>51</td>
<td>33</td>
<td>0.073</td>
</tr>
<tr>
<td></td>
<td>Protein transport</td>
<td>140</td>
<td>80</td>
<td>0.052</td>
</tr>
<tr>
<td></td>
<td>Phosphorylation</td>
<td>156</td>
<td>78</td>
<td>0.31</td>
</tr>
<tr>
<td>Colorectal</td>
<td>Organelle lumen</td>
<td>384</td>
<td>147</td>
<td>0.0001</td>
</tr>
<tr>
<td></td>
<td>Nucleotide process</td>
<td>34</td>
<td>22</td>
<td>0.782</td>
</tr>
<tr>
<td></td>
<td>RNA polymerase activity</td>
<td>11</td>
<td>8</td>
<td>0.992</td>
</tr>
<tr>
<td>Ovarian</td>
<td>Protein kinase activity</td>
<td>159</td>
<td>86</td>
<td>0.0003</td>
</tr>
<tr>
<td></td>
<td>Serine/theorine kinase activity</td>
<td>117</td>
<td>56</td>
<td>0.0002</td>
</tr>
<tr>
<td>Gastric</td>
<td>Vasculature development</td>
<td>45</td>
<td>20</td>
<td>0.08</td>
</tr>
<tr>
<td></td>
<td>Nucleotide process</td>
<td>29</td>
<td>11</td>
<td>0.437</td>
</tr>
<tr>
<td></td>
<td>Regulation cell migration</td>
<td>28</td>
<td>11</td>
<td>0.384</td>
</tr>
<tr>
<td>All Cancers</td>
<td>Vesicle</td>
<td>44</td>
<td>23</td>
<td>0.819</td>
</tr>
<tr>
<td></td>
<td>Cell cycle process</td>
<td>88</td>
<td>20</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>Organelle lumen</td>
<td>169</td>
<td>49</td>
<td>0.0006</td>
</tr>
</tbody>
</table>

**Table 22.** Functional clusters common in both m6A(+) and m6A(-).
The number of m6A(+) and m6A(-) involved in each shared activity and the associated significance is shown (p-value).
6.3.6 Immunoprecipitation of m6A

Our *in-silico* analysis revealed that m6A is abundant in both PCa and CRPC. Although m6A has been profiled in several cell lines including hepatic cell carcinoma (HepG2) (Dominissini *et al.*, 2012) and human embryonic kidney cells (HEK293T) (Meyer *et al.*, 2012), its distribution in PCa cell lines is unknown. We profiled m6A in LNCaP and LNCaP-LN3 cells through MeRIP-seq.

6.3.6.1 Validation of post-fragmentation RNA size

Validation of RNA size was performed by running fragmented 0.5μg RNA on an agarose gel for 30 mins (Figure 50).

![Image of agarose gel]

**Figure 50.** Validation of RNA size.

Total RNA extracted from LNCaP and LNCaP-LN3 was fragmented with ZnCl$_2$ for 4 mins at 94°C. The desired RNA size of ~100nt (white arrows) was confirmed on a 1.5% (wt/vol) agarose gel.
6.3.6.2 Immunoprecipitation quality control
Before generating IP libraries and performing RNA-seq on the samples, the success of IP was confirmed by performing qRT-PCR on methylated transcripts identified from Dominissini et al’s analysis (Dominissini et al., 2012). Methylated transcripts present in the m6A-antibody IP samples relative to the bead-only (null) IP sample was evidence for successful IP.

A total of 100µg RNA extracted from each cell line LNCaP and LNCaP-LN3 was fragmented to ~100nt and immunoprecipitated with m6A-antibody. cDNA was synthesized through RT-PCR and m6A susceptible genes (CALM3 and DDX39) were measured with qPCR in relation to controls (GAPDH and U1- not m6A susceptible). Results demonstrated a 1.16 and 1.43-fold enrichment for m6A susceptible CALM3 and DDX39 compared to control (GAPDH/U1) respectively (Figure 51).
Figure 51. Concentration of Genes whose transcripts are known to undergo N6-adenosine methylation.

A low scale immunoprecipitation (IP) was performed to test for success prior to performing a IP with RNA sequencing. Total RNA was extracted from PCa cell lines and fragmented to ~100nt using ZnCl₂. Transcripts known to undergo N6-adenosine methylation according to the list published by Dominissini et al was screened (Dominissini et al., 2012). The expression (qRT-PCR) of CALM3 and DDX39 was measured and was higher in the m6A-antibody samples (Blue) compared to bead-only (null) control (Grey), suggesting successful IP.
6.3.7 RNA-sequencing of m6A-IP libraries

Following validation, cDNA libraries were prepared and RNA-sequencing was performed. The bioinformatics analyses were performed with Dr. Ian Sudbery and Dr. James Bradford (Bioinformatic Hub, University of Sheffield).

6.3.7.1 MeRIP-seq read quality control

MeRIP-seq on the IP samples obtained 65 million reads from the two cell lines LNCaP and LNCaP-LN3 (Table 23 and Figure 52). For each cell line, there were three samples, 1) input control (Fragmented and untreated RNA), 2) IP with m6A-antibody and 3) IP without m6A-antibody (negative control). The %GC is slightly lower on m6a-seq sample than on RNA-seq samples and that %GC for LN3-input isn’t particularly higher than for m6A pulldown as would be expected if m6a was enriched for A-bases compared to input. In addition, m6A samples had more reads than input samples. This is similar to how data generally appear with ChIP-seq.

<table>
<thead>
<tr>
<th>Sample</th>
<th>%GC</th>
<th>Length</th>
<th>Reads</th>
</tr>
</thead>
<tbody>
<tr>
<td>LN3-Input 1</td>
<td>44</td>
<td>108</td>
<td>1,946,337</td>
</tr>
<tr>
<td>LN3-Input 2</td>
<td>44</td>
<td>108</td>
<td>1,824,055</td>
</tr>
<tr>
<td>LN3-m6A 1</td>
<td>43</td>
<td>108</td>
<td>5,137,133</td>
</tr>
<tr>
<td>LN3-m6A 2</td>
<td>43</td>
<td>108</td>
<td>4,772,878</td>
</tr>
<tr>
<td>LN3-RNA 1</td>
<td>57</td>
<td>108</td>
<td>9,192,621</td>
</tr>
<tr>
<td>LN3-RNA 2</td>
<td>58</td>
<td>108</td>
<td>8,626,487</td>
</tr>
<tr>
<td>LNCaP-Input 1</td>
<td>47</td>
<td>108</td>
<td>2,854,219</td>
</tr>
<tr>
<td>LNCaP-Input 2</td>
<td>47</td>
<td>108</td>
<td>2,670,629</td>
</tr>
<tr>
<td>LNCaP-m6A 1</td>
<td>43</td>
<td>108</td>
<td>4,087,778</td>
</tr>
<tr>
<td>LNCaP-m6A 2</td>
<td>43</td>
<td>108</td>
<td>3,879,241</td>
</tr>
<tr>
<td>LNCaP-RNA 1</td>
<td>58</td>
<td>108</td>
<td>10,606,616</td>
</tr>
<tr>
<td>LNCaP-RNA 2</td>
<td>58</td>
<td>108</td>
<td>10,017,420</td>
</tr>
</tbody>
</table>

**Table 23.** LNCaP and LNCaP-LN3 MeRIP-seq read quality control.

A total of 65 million reads were obtained from LNCaP and LNCaP-LN3. Each sample were split into two lanes (duplicates). The samples included, input RNA (fragmented and untreated), m6A (IP with m6a-antibody), ‘RNA’ control (IP without m6a-antibody).
Figure 52. Plot of base qualities for LNCaP-LN3-input.

This plot shows the base qualities for LNCaP-LN3 input. a) Reads were of a relatively good quality, with only a small drop off in quality towards the end of the read. b) A large number of uncalled bases at the end of reads and a highly unusual peak in uncalled bases at position four was seen. Therefore, the first and last four bases from each read were trimmed to make reads of 100bp long.
6.3.7.2 MeRIP-seq read mapping

Reads were mapped using HISAT (Hierarchical Indexing for Spliced Alignment of Transcripts) splice aware short read mapper. Samples were mapped in paired-end mode and data from separate lanes was merged after mapping. The mapping rate for LNCaP and LNCaP-LN3 was 66% and 68% respectively (Table 24).

<table>
<thead>
<tr>
<th>TRACK</th>
<th>PAIRS_TOTAL</th>
<th>PAIRS_MAPPED</th>
<th>% MAPPED</th>
</tr>
</thead>
<tbody>
<tr>
<td>LN3-INPUT</td>
<td>3,770,378</td>
<td>2,323,267</td>
<td>62%</td>
</tr>
<tr>
<td>LN3-RNA</td>
<td>17,819,030</td>
<td>5,067,952</td>
<td>28%</td>
</tr>
<tr>
<td>LN3-M6A</td>
<td>9,909,958</td>
<td>6,772,149</td>
<td>68%</td>
</tr>
<tr>
<td>LN1-INPUT</td>
<td>5,524,836</td>
<td>2,809,170</td>
<td>51%</td>
</tr>
<tr>
<td>LN1-RNA</td>
<td>20,623,849</td>
<td>5,892,628</td>
<td>29%</td>
</tr>
<tr>
<td>LN1-M6A</td>
<td>7,966,981</td>
<td>5,281,282</td>
<td>66%</td>
</tr>
</tbody>
</table>

Table 24. LNCaP and LNCaP-LN3 reads mapping rate.

The pairs mapped and percentage mapped for each sample are shown. For LNCaP and LNCaP-LN3, 66% and 68% were mapped to unique loci.

6.3.7.3 Identification of m6A sites

N6-methyladenosine sites (peak calling) were identified at a 5% FDR threshold using MACS2 (Model-based Analysis of ChIP-Seq) software. There were 132 and 218 peaks for LNCaP and LNCaP-LN3 respectively (Table 25). Peaks were centered on a strong ‘A’ base motif (Figure 53).
Table 25. m6A peak calling.

The lengths and number of peaks for the two PCa cell lines are shown. We found 218 and 132 peaks in the LNCaP-LN3 and LNCaP samples respectively.

<table>
<thead>
<tr>
<th>TRACK</th>
<th>AVG(LENGTH)</th>
<th>MAX(LENGTH)</th>
<th>MIN(LENGTH)</th>
<th>NUMBER</th>
</tr>
</thead>
<tbody>
<tr>
<td>LN3 M6A</td>
<td>70,460</td>
<td>786,100</td>
<td>632</td>
<td>218</td>
</tr>
<tr>
<td>LNCAP M6A</td>
<td>125,100</td>
<td>799,200</td>
<td>835</td>
<td>132</td>
</tr>
</tbody>
</table>

**Figure 53.** Sequence logo.

MEME was used to identify the top motifs for each sample. These motifs have strong, central ‘A’ residues.
6.3.8 Expression of m6A transcripts in prostate cancer cell lines
A total of 85 genes were significantly enriched within the m6A fraction and differentially expressed across cell type. Forty-four percent were specific to LNCaP-LN3 and included interesting candidates involved in NOTCH signalling (DTX2), regulation of the androgen (and oestrogen) receptor expression (CDK11), G protein signalling (Opsin-1), DNA repair (PARG) and Y Chromosome located genes implicated in testis/prostate carcinogenesis (TSPY1/3) (Figure 54). Gene-set-enrichment analysis identified significant alterations of pathways specific to the detection of external stimuli and RNA splicing in LNCaP-LN3 cells (p<0.003), suggesting a role in bypassing androgen-dependent signalling.
Figure 54. Methylated transcripts that are differentially expressed in prostate cancer cell lines.

MeRIP-seq was performed in LNCaP and LNCaP-LN3 cell lines. Following bioinformatic analysis, a total of 37 m6A-susceptible transcripts were aberrantly expressed in LNCaP-LN3. These included candidates (red) involved in oncogenic pathways, such as PARG.


CICP27, LINC00623, PDE4DIP, C2orf78, ANKRD36, RP11-1223D19.1, SPDY8BP, RP11-706015.5, TIX28P1, RBM18, RP11-14G6E13.4, CTD-2314B22.3, BMS1P18, RP11-261B23.2, STRC, NOMO1, METTL9, AC007952.4, KYNUP1, ALYREF, GGT3P

Poly (ADP-ribose) glycohydrolase (PARG) has previously been investigated within the Sheffield Department of Oncology and has been shown to be involved in DNA repair (Fathers et al., 2012). One response to DNA damage is the synthesis of poly (ADP-ribose) (PAR) through poly ADP-ribosylation of numerous target proteins, including proteins involved in the RNA transcription export complex (TREX) (Jungmichel et al., 2013; Zhang et al., 2013). PAR is transient, once other repair proteins have localised to the site of damage, PAR must be removed before repair can take place. The PAR polymerase (PARP) enzymes responsible have been implicated in the repair of both single and double-strand breaks. PARG is the endo-exoglycohydrolase that cleaves glycosidic bonds, reversing the action of PARP enzymes and returning proteins to their native site. Therefore, the balanced level of PAR is driven by PARP and PARG enzymes (Figure 55). Recently, a phase 2 trial has shown that treatment with the addition of PARP inhibitor Olaparib in mCRPC who had defects in DNA-repair genes led a to higher response rate (Mateo et al., 2015).

The levels of the product of PARP, PAR (Western blot performed by Dr Helen Bryant, Figure 56a), and PARP1 (Figure 56b) and PARG enzymes (qRT-PCR) in LNCaP and LNCaP-LN3 were measured (Figure 56c). The levels of PAR were reduced in LNCaP-LN3 and the levels of PARP1 and PARG increased in LNCaP-LN3.
Figure 55. The role of PARG and PARP in DNA repair.

In response to stress or DNA damage PARP enzyme mediates PAR synthesis which recruits DNA repair proteins. PARP adds one or many ribose moieties onto target proteins synthesising long branches PAR chains from nicotinamide adenine dinucleotide (NAD+). In order for repair proteins to function at full potential, PAR needs to be removed by PARG enzyme through cleavage of the glucosidic bonds.

PAR, poly (ADP-ribose); PARP, PAR polymerase; PARG, PAR Glycohydrolase
Figure 56. Levels of PARP activity and mRNA in LNCaP and LNCaP-LN3.

The expression of PAR protein and PARG mRNA was measured. a) western blot showed reduced levels of PAR seen in LNCaP-LN3 compared to LNCaP with a tubulin loading control (Performed by Dr Helen Bryant); b) qRT-PCR for PARP1 enzyme mRNA indicated increased fold change in LNCaP-LN3; c) qRT-PCR for PARG enzyme indicated increased fold change in LNCaP-LN3.
6.4 Discussion

6.4.1 *In-silico* analysis of N6-methyladenosine
The methylation of N6-adenosine base has been known since the 1970s, however, new technological advances now allow in-depth analyses of this modification (N. Liu *et al.*, 2014). The recent discovery of the reversibility of m6A has revived its interest in the field of cancer epigenetics (Jia *et al.*, 2011). Dominissini *et al* (Dominissini *et al.*, 2012) revealed the abundance of m6A in mRNAs at a transcriptome-wide level and provided reference for the current analysis.

A cohort of aberrantly (Student’s t-test p<0.001) expressed mRNAs in common eight solid tumours that are susceptible to N6-adenosine methylation has been derived. These mRNAs were subcategorized into up- (top 10% percentile) and downregulatory (bottom 10% percentile) forms. The numbers of predicted m6A appear to vary amongst cancer types, with the highest seen in bladder and colorectal cancer, and lowest seen in lung and renal cancer. When comparing the ratio of m6A against non-m6A susceptible RNAs within and across cancers, colorectal, bladder and prostate cancers have a ratio of >1 suggesting that they have a higher proportion of m6A susceptible RNAs compared to non-susceptible RNAs. However, the mean fold changes (cancer/control) between m6A and non-m6A susceptible RNAs did not differ within each cancer or across all cancers.

To identify the most important m6A susceptible mRNAs, we selected ones within the highest and lowest fold changes and found that colorectal, bladder and gastric cancers had the highest percentage of upregulatory m6A susceptible mRNAs, in contrast all the cancers had more downregulatory m6a non-susceptible mRNAs. On analysing mean fold changes of predicted m6As, lung had the highest and bladder the lowest upregulatory FC, and bladder the highest and gastric the lowest downregulatory FC. These findings of varying percentages of m6A susceptible RNAs, percentages of up- and downregulatory RNAs and varying ranges of FC are interesting, but
the explanation for these differences are currently unclear and warrant further evaluation. Although, there are common oncogenic mechanisms across cancers such as apoptosis and cell cycle regulation, some individual cancers possess unique oncogenic mechanisms such as androgen regulation in PCa (Jerónimo et al., 2011). The uniqueness of individual cancers may explain the variations seen in the current analysis. Nonetheless, it is clear that the distribution of m6A is abundant within the cancer transcriptome and can be as high as 59% (bladder).

Shared genes that encode m6A susceptible transcripts within each cancer type and across cancers were analysed using gene enrichment software (DAVID). These shared genes appear to be involved in oncogenic mechanisms such as regulation of apoptosis, cell cycle, transcription, mRNA processing and transport of proteins.

At the time of analysis, around five independent studies reporting m6A MeRIP-seq data have been described (Dominissini et al., 2012; Meyer et al., 2012; Fustin et al., 2013; Hess et al., 2013; X. Wang et al., 2014). The origin of RNAs came from HCC, mouse liver, human brain, HEK293T, mouse brain, U2OS and HeLa cells. Studies focused on identifying the distribution of m6A and regulators of this modification at a molecular level. With the increasing interest in evaluating the roles of m6A in disease processes, Lui et al combined data from the five independent studies and created an integrated methyltranscriptome database (MeT-DB, http://compgenomics.utsa.edu/methylation). This database comprises a collection of m6A predicted sites and a genome browser to enable visualization and comparison of m6a in different contexts (Liu et al. 2015). Our focus was different in the sense that we predicted m6A sites within eight common cancers by annotating RNA-seq data obtained from human benign and malignant tissues, with the aim to identify predicted m6a sites specific to cancers and unravelling associated oncogenic mechanisms.

The exact role of m6A is unclear, but involves a complex machinery regulated by ‘writers’ ‘erasers’ and ‘readers’ which affects RNA splicing,
export, stability/decay and translation. The outcomes are associated with diseases ranging from benign such as obesity, diabetes and epilepsy, to malignant such as breast, prostate and colorectal cancer (Niu et al., 2013; N. Liu et al., 2014; Maity et al., 2015). With regards to cancer, FTO (Fat mass and obesity-associated protein) is a m6A ‘eraser’ (demethylase) (Jia et al., 2011) and allelic variants/mutation of FTO has been shown to be associated with high incidences of prostate cancer risk (Machiela et al., 2012), breast (Kaklamani et al., 2011), colorectal and stomach cancer (Linnebacher et al., 2010). In contrast, increasing levels of m6A through increasing the methyl donor SAM (S-adenosylmethionine), results in growth inhibition of breast (Pakneshan et al., 2004), colon (Guruswamy et al., 2008) and gastric cancer (Zhao et al., 2010).

Although our understanding of m6A biology is increasing, its distribution in common cancers are unknown. The current in-silico analysis describes the variation in levels of suspected N6-adenosine methylated RNAs within and across eight common solid tumours. In addition, the functional analysis provides information on the roles of these RNAs in oncogenesis. Knowing that m6A is reversible and its evolving role in cancer, tackling this modification in our identified list of susceptible RNAs may alter important oncogenic pathways. The m6A machinery is diverse and methylated transcripts may directly or indirectly affect oncogenesis through its regulators.

6.4.2 N6-methyladenosine profiling in prostate cancer cell lines
The m6A profiling studies discussed above included a range of cell lines, however, these do not include PCa cell lines. To the best of our knowledge, at the time of evaluation, the current analysis was the first to map out the distribution of m6A in PCa cell lines. Due to time constraint and funding availability, the MeRIP-seq experiment was performed once. Triplicates would be performed in the near future by members from the Catto laboratory. However, interesting findings were observed and are worth discussing.
LNCaP-LN3 are metastatic siblings of the same cell line LNCaP. We attempted to profile m6A in PCa cell lines and identify differentially expressed RNAs that may explain castration-resistant or metastatic potentials. Validation of the IP protocol was performed on a low-scale experiment using 300µg of total RNA before proceeding to using 1mg. IP peaks/motifs were centred around ‘A’ residues and ‘GC’ depleted. Several oncogenic candidates were confined to LNCaP-LN3 (i.e. DTX2, CDK11, Opsin-1, PARG, TSPY).

DTX2 is expressed in PCa and is involved in activating the NOTCH signalling pathway which has recently been shown to promote CRPC (Stoyanova et al., 2016). CDK11 is related to cell cycle arrest and apoptosis in a kinase-dependent manner. In addition, it has been shown that abnormal expression of CDK11 in PCa tissue led to the dysfunction of apoptosis and inhibition of metastasis of AR-positive PCa cells (Chi et al., 2014).

PARG is susceptible to N6-adenosine methylation in LNCaP-LN3 according to our profiling results. We showed that PARG was overexpressed in LNCaP-LN3 compared to LNCaP, which may be a response to the overexpressed PARP1 activity in LNCaP-LN3 (Figure 52). As discussed PAR, PARG and PARP are involved in a complex DNA damage/repair mechanism. These candidates are of interest in PCa, as Mateo et al conducted a phase 2 trial in which men (n=50) with mCRPC were treated with PARP inhibitor, Olaparib. All men received prior treatment with Docetaxel. The study showed that men who had defects in DNA repair genes had a higher response rate (reduction in PSA and tumour cells) to Olaparib (Mateo et al., 2015). Olaparib is already approved for treating ovarian cancers with BRCA1/2 mutations, and appears to be a potential PCa treatment agent (Ledermann et al., 2014).

TSPY has been shown to be expressed in prostate tissue, regulated by androgens and be involved in testicular cancer and PCa (Lau, 1999). TSPY is one of >250 cancer testes antigens (CTA) encoding a protein found in
testicular tissue that maybe involved in spermatogenesis (Almeida et al., 2009). CTAs are proteins that are normally only expressed in human testes but are aberrantly expressed in some cancers. They have tumour-restricted expression characteristics with a strong immunogenicity, which together, have made CTAs an attractive target for a possible diagnostic and/or prognostic biomarker (Maxfield et al., 2015).

6.4.3 N6-methyladenosine machinery
The discoveries of m6A mediators (writers, erasers, readers) and the profiling of m6A on the transcriptome enabled further evaluation of the role of m6A in human diseases from different angles including, biological processes (oncogenic, immunological, metabolic processes), pathway levels (p53-mediated pathways), signalling and machinery pathways (spliceosome, nuclear export) and molecular levels (interactions with proteins, mRNAs and subsequent gene expression).

The stability of mRNAs is affected by m6A methylation, and a negative correlation exists between m6A and mRNA abundance (shorter half-life). The presence of readers affects mRNA transport, storage, stability, splicing and translation. Our in-silico and in-vitro analyses have shown abundance of m6A within PCa (including PARG, CDK, TSPY), and these methylated RNAs may directly or indirectly affect the signalling pathways and mRNA processing activities mentioned. Epigenetic DNA and histone modifications affect mostly transcriptional processes, however, it is the end product of proteins that determines biological phenotype and therefore, the post-transcriptional regulation of protein synthesis may prove to be significant.

Several processes have been linked to m6A, for example, reduced m6A levels leads to prolonged nuclear retention of circadian RNAs (period circadian clock 2, Per2 and aryl hydrocarbon receptor nuclear translocator-like, Arntl). The circadian clock is linked to many physiological processes in the human body and it has been shown that DNA methylation is associated
with the circadian clock influencing clock gene expression and oncogenesis. This may also be the case for RNA methylation (Joska et al., 2014).

6.5 Conclusion
Our *in-silico* analysis identified RNAs in eight common solid tumours that are potentially susceptible to N6-adenosine methylation. These RNAs appear to be involved in important oncogenic and signalling pathways. MeRIP-seq confirmed the presence of m6A in PCa cell lines and revealed differentially expressed candidates involved in prostate oncogenesis. m6A modification has been shown to be reversible and implicated in oncogenesis, therefore are attractive as potential targets for investigating biomarkers and therapeutic agents.
CHAPTER 7: Discussion
Prostate cancer is the most common cancer in men, with a rising incidence in the UK. Over, 40,000 new cases are diagnosed, and over 10,000 deaths occur every year in the UK (CRUK, 2016). Many PCa detected are clinically insignificant, hence screening may represent over-diagnosis and subsequent over-treatment of insignificant disease (Illic et al., 2013). The Prostate, Lung, Colon and Ovaries (PLCO) screening trial failed to show any evidence of survival benefit (Andriole et al., 2009) and although the European Randomised Study of Screening for Prostate Cancer (ERSPC) showed a reduction in PCa mortality over 14 years, the risk of over-diagnosis was substantial (Hugosson et al., 2010). The Scandinavian Prostate Cancer Group-4 (SPCG-4) and the Prostate Cancer Intervention Versus Observation Trial (PIVOT) RCTs of radical prostatectomy versus watchful waiting have shown that the improved overall survival was restricted to high-risk, clinically significant PCa (Wilt, 2012; Bill-Axelson et al., 2014).

Prostate cancer is a heterogeneous disease and in order to minimise over-diagnosis and over-treatment, the optimal diagnostic pathway should be able to reliably differentiate between men with and without clinically significant disease. Diagnostic strategies include a combination of PBx techniques and a selection criteria for (r)PBx through imaging (i.e. mpMRI) and biomarkers (i.e PCA3). In addition, once PCa is diagnosed, it is important to be able to predict progression to advanced disease, such as CRPC or metastatic PCa, and also to monitor for recurrence following radical treatment.

7.1 Repeat prostate biopsy outcomes
We reported unpublished data on rPBx outcomes in a cohort of Sheffield men within the national ProtecT study. To the best of our knowledge, this is the largest UK based rPBx set of data at the time of analysis. A literature review was conducted on rPBx outcomes and our results (19.6%) are consistent with the international data (7.5-26.2%). It is important to emphasise that the relatively low PCa detection rate on rPBx in men with suspicious risk (raised PSA) results in high costs attributed to healthcare resources, and most importantly leave potential high-risk disease undetected.
In addition, performing unnecessary rPBx put men at risk of PBx complications (Rosario et al., 2012; Loeb et al., 2013a). Many thoughts have been put into better selecting men for rPBx or increasing the accuracy of diagnostic tests (Kirby et al., 2012). Saturation biopsy or TPM-Bx are PBx techniques discussed before that can increase the detection of anterior tumours (Ekwueme et al., 2013; Li et al., 2014). However, the latter requires general anaesthesia and is more invasive than the traditional TRUS-PBx.

With the addition of mpMRI guidance, this may (PROMIS study) further improve outcomes. The PROMIS study showed that using mpMRI a quarter of men may avoid a primary PBx and performing mpMRI prior to TRUS-Bx up to 18% more clinically significant PCa may be diagnosed compared with TRUS-PBx alone (El-Shater Bosaily et al., 2015; Ahmed et al., 2017).

Majority of PCa detected on rPBx within our cohort were low-grade Gleason 6, with ~4.5% being high-grade. This confirms our existing knowledge on detecting high volumes of insignificant disease on rPBx. Which again, indicates that we need better selection criteria for rPBx. With regards to predictors, our analysis is consistent with the literature in the sense that, PSA, PSA velocity, lower %free PSA and ASAP are positive predictors of PCa.

### 7.2 PCA3-shRNA2 and prostate cancer

Knowing the issues associated with identifying PCa on rPBx, apart from radiological alternatives, urgent biomarkers are also needed. Prostate cancer is a heterogeneous disease, the roles of biomarkers are not just to diagnose PCa, but also used to predict/detect disease progression (CRPC, metastasis) and monitor for recurrence (BCR) following radical treatment (Haese et al., 2008). The current PCA3 test is promising, and now new predictive tests are available such as, 4K and PHI. Since PCA3 is a ncRNA and these RNAs are implicated in many biological processes by targeting mRNAs, and that long ncRNA are processed into smaller more active RNAs, we hypothesised that PCA3 may encode a shorter segment. Indeed, we identified PCA3-shRNA2, which is embedded in intron 1 of PCA3. PCA3-shRNA2 was upregulated in a testosterone manner and appeared to be overexpressed in urinary samples.
obtained from men with PCa, although this was not the case for FFPE initial PBx specimens. PCA3 function was unclear when this work was conducted and we identified a list of mRNA targets of our short-PCA3 including COPS2, SOX11, Noggin, WDR48, TEAD1, which all have roles in oncogenesis. On validation, the expression of COPS2 and SOX11 was negatively correlated to PCA3-shRNA2 in urinary samples from men with PCa, and that the former was significantly underexpressed in PCa urinary samples (Drayton et al., 2015). This piece of work identified a new RNA within PCA3 that has similar predictive role and function in part by targeting mRNAs involved in PCa oncogenesis. However, urinary PCA3-shRNA2 failed to differentiate localised disease from advanced disease. We also showed that PCA3-shRNA2 was detectable in historic FFPE specimens stored for up to eight years, representing high stability of micro/short RNAs (Pang et al., 2017). Knowing the stability of short RNAs, PCA3-shRNA2 may be a more suitable target of the current PROGENSA PCA3 assay.

7.3 N6-adenosine methylation and prostate cancer

Over a hundred post-transcriptional modifications of RNAs have been described. The most common being methylation of the N6-adenosine residue. The finding of the reversible nature of m6A revived its interest in biological research. Since this finding, m6A has been mapped throughout the RNA transcriptome (Dominissini et al., 2012). It has also been shown that m6A is regulated by ‘writers’, ‘erasers’ and ‘readers’ which affects RNA exportation, storage, processing, degradation and translation (Niu et al., 2013; Fu et al., 2014; Cao et al., 2016). In addition, m6A appears to be associated with a number of biological processes including adipogenesis, stem cells renewal, spermatogenesis, development, immune responses and oncogenesis (Niu et al., 2013; Maity et al., 2015).

Since our PCA3-shRNA2 failed to differentiate advanced/metastatic PCa from localised disease, we focused upon this new trait of epigenetics, RNA methylation of N6-adenosine. No association has been documented at the time of analysis between m6A and PCa. Hence, we performed a preliminary
*in-silico* analysis to predict the distribution of m6A in PCa. We showed abundance of m6A in PCa and across seven other common solid tumours, and identified ones that were upregulated and downregulated. In addition, methylated RNAs common across PCa microarray datasets analysed were subject to functional annotation. Many of these methylated RNAs were found to be involved in oncogenic pathways including regulation of transcription, angiogenesis and nuclear division.

Following our *in-silico* analysis, we profiled m6A in PCa cell line through MeRIP-seq. We identified differentially expressed m6A transcripts in metastatic LNCaP cell lines, including PARG, CDK and TSPY. PARG inhibitors have recently been show to improve outcomes in men with mCRPC with known defects in DNA repair genes (Mateo et al., 2015), hence we focused on evaluating PARG along with PAR and PARP which are all involved in the DNA repair mechanism. We found differences in the expression of these three products (increased PARP and PARG, and decreased PAR in LNCaP-LN3) in metastatic PCa cell lines, which may suggest a mechanism for PARG-m6A in castration-resistance and metastasis.

Due to time constraint, the MeRIP-seq was not repeated. Ideally, this needs to be performed in triplicates. Differentially expressed methylated mRNAs common in triplicate experiments should be validated in LNCaP and LNCaP-LN3 cell lines, followed by validation in our patient urinary and PBx samples. RNAs found to undergo N6-adenosine methylation could then be explored functionally by performing apoptosis and cell proliferation assays.
7.4 Conclusions
Repeat PBx are over-performed resulting in over-treatment of insignificant disease. PCA3-shRNA2 is a small RNA expressed by PCA3 and is detected in urinary and PBx samples obtained from men with PCa. This short-RNA is overexpressed in both urinary and initial PBx samples, although the latter did not reach statistical significance. PCA3-shRNA targets numerous mRNAs involved in prostate oncogenesis. Knowing the stability, ease of detection and high activity of short-RNAs, PCA3-shRNA2 may be a more suitable biomarker than the current PCA3 assay. MeRIP-seq revealed abundance of m6A in PCa cell lines, and differentially expressed candidates are involved in PCa biology. m6A is reversible, dynamic and is involved in oncogenesis. These features make m6A potential biomarkers and therapeutic targets.
References


Carver, B. S., Chapinski, C., Wongvipat, J., *et al.* (2011) ‘Reciprocal feedback regulation of PI3K and androgen receptor signaling in PTEN-
deficient prostate cancer.’, Cancer Cell, 19(5), pp. 575–86.


Dudziec, E., Miah, S., Choudhry, H. M. Z., et al. (2011) ‘Hypermethylation of
CpG islands and shores around specific microRNAs and mirtrons is associated with the phenotype and presence of bladder cancer.', Clinical Cancer Research, 17(6), pp. 1287–96.


Karkera, J., Steiner, H., Li, W., et al. (2011) ‘The anti-interleukin-6 antibody siltuximab down-regulates genes implicated in tumorigenesis in prostate


attenuates miRNA-mediated repression by promoting miRISC dissociation from the target RNA.', *Nucleic Acids Research*, 40(11), pp. 5088–100.


biopsy.', PloS One, 8(2), p. e57480.


integrity number for assigning integrity values to RNA measurements’, *BMC Molecular Biology*. BioMed Central, 7(1), p. 3.


Tonttila, P. P., Lantto, J., Pääkkö, E., et al. (2016) ‘Prebiopsy Multiparametric Magnetic Resonance Imaging for Prostate Cancer Diagnosis in Biopsy-naive Men with Suspected Prostate Cancer Based on Elevated Prostate-specific Antigen Values: Results from a Randomized Prospective Blinded Controlled


Appendices

ProtecT consent form

ProtecT study
(Pretaste testing for cancer and treatment)

Professor Fredrik Handed
Section of Urology
University of Sheffield
0114 2712154

Dr Jenny Donovan
Dept. of Social Medicine
University of Bristol
0117 9287214

Professor David Neal
Dept. of Surgery
University of Cambridge
01223 331 940

Consent form 1

ProtecT study entry

Please initial box

1. I confirm that I have read and understand the information sheet dated .......... for the above study and have had the opportunity to ask questions

2. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason, without my medical care or legal rights being affected

3. I agree to take part in the above study and to have blood taken for PSA test

4. I confirm I agree to having blood taken today for other tests

5. I confirm I agree to studies in the future on prostate cancer
   (a) involving checking my GP and/or hospital records

   (b) involving tests on the sample of frozen blood

   (c) and I understand that the information held by the NHS, and records maintained by the General Register Office may be used to keep in touch with me and follow-up my health status

Name of participant ......................... Date ................ Signature .........................

Name of researcher ......................... Date ................ Signature .........................

Name of person taking consent (if different from researcher)

................................................. Date ................ Signature .........................

1 copy for participant; 1 copy for researcher
**ProMPT consent form**

Sheffield Teaching Hospitals NHS Foundation Trust

**ProMPT study**

(_Prostate cancer mechanisms of progression and treatment_)  
(Version 1.1 – 01/03/07)

ProtecT study identification number: .................  

Please initial boxes

1. I confirm that I have read and understand the information sheet (dated ............ version ...) for the above study.  

2. I give permission for samples of blood, tissue and urine to be obtained from me. I understand that the blood, tissue and urine will be used to extract genetic material (DNA) and other factors. I give permission for the analysis of these materials using techniques to identify genes and other factors which may influence prostate cancer. I understand that these samples will be kept for future analysis in conjunction with this research.  

3. I give permission, where appropriate, for surplus biopsy and pathology specimens to be used in this study.  

4. I give permission, whenever appropriate, for additional biopsy tissue to be taken at the time of routine testing, to be used for research purposes in this study.  

5. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving reason, without my medical care or legal rights being affected.  

<table>
<thead>
<tr>
<th>Name of participant</th>
<th>Date</th>
<th>Signature</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Name of person taking consent</th>
<th>Date</th>
<th>Signature</th>
</tr>
</thead>
<tbody>
<tr>
<td>I copy for participant; 1 for researcher; 1 to be kept with hospital notes</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Consent form for ProtecT participants attending annual clinics  

(ABRIDGED form for previous ProMPT)
Copies of Publications Arising from Thesis