

# Identifying regulatory mechanisms of androgen receptor variants by DNA-PKcs

Thesis submitted in partial fulfilment of the requirement of the degree of Doctor of Philosophy



Beth Adamson

Translational and Clinical Research Institute

Faculty of Medical Sciences

Newcastle University

September 2022



## Abstract

The androgen receptor (AR) is a master regulator of prostate cancer (PCa) development and progression, hence current therapies target the AR signalling pathway to inhibit tumour growth. The generation of alternatively spliced AR variants (AR-Vs) is a major resistance mechanism observed in patients who progress to the advanced castrate-resistant PCa (CRPC) stage of disease. In contrast to full-length AR (FL-AR), AR-Vs are constitutively active and drive the growth of PCa without the requirement of activating androgens. Furthermore, AR-Vs are refractory to the current repertoire of AR-targeting therapies hence there is a major drive to develop treatments that can inhibit these aberrantly functioning receptors. Targeting AR-V co-regulatory proteins, that are required for enabling their function, represents a tractable means for inactivating AR-Vs in advanced disease.

DNA-PKcs, a key kinase in the DNA damage response, has been shown to regulate FL-AR transcriptional activity and is upregulated in both PCa and CRPC. Given shared co-regulator dependencies of FL-AR and AR-Vs, and the observation that AR-Vs regulate DNA damage, we hypothesised that DNA-PKcs may influence AR-V activity as a co-regulator. Using proximity biotinylation, we show that DNA-PKcs is a prominent AR-V interacting protein in the presence and absence of DNA damage and demonstrate that DNA-PKcs regulates AR-V transcriptional activity and protein abundance in CRPC cell lines. Furthermore, DNA-PKcs inhibition and depletion has anti-proliferative effects in several CRPC cell lines in the absence of DNA damage.

Global transcriptomic analysis revealed a novel role for DNA-PKcs in the regulation of alternative splicing that is important in the generation of AR-Vs. Interrogation of the regulatory role of splicing associated genes by DNA-PKcs revealed that DNA-PKcs regulates expression of the RNA binding protein, RBMX, which was then validated as a key regulator of AR-V synthesis. In conclusion, targeting DNA-PKcs or RBMX are potential therapeutic options for AR-V positive PCa patients.

## Acknowledgements

First, I would like to thank my supervisor Dr Luke Gaughan for the opportunity to join his group. His support and scientific guidance has been invaluable throughout my PhD. I am so grateful for his mentorship over the course of the project; it has been a fantastic experience working alongside him. I would also like to thank the other members of my supervisory team, Dr Stuart McCracken, Dr Elaine Willmore and Dr Ian Hickson for their input into the project.

A special thanks to the rest of the Solid Tumour Discovery group for their advice and support whenever needed. In particular, I would like to thank fellow PhD students, Laura Walker and Nicholas Brittain for their advice and input into the project as well as their constant reassurance and words of encouragement.

Thank you to the donors of the Ken Bell Scholarship who made this research possible.

I would like to thank my family and friends who have always provided support and been there for me.

Most importantly, I cannot thank my partner, Lewis, enough for his incredible support and patience throughout the past 4 years. His belief in me, always making me laugh and allowing us to start our family with our wonderful dog, Joey, has gotten me through the difficult times and for this, I could not be more grateful.

## Table of Contents

<b>Abstract .....</b>	<b>iii</b>
<b>Acknowledgements .....</b>	<b>iv</b>
<b>Abbreviations.....</b>	<b>xi</b>
<b>List of figures.....</b>	<b>xiv</b>
<b>List of tables.....</b>	<b>xxi</b>
<b>Chapter 1 Introduction .....</b>	<b>1</b>
<b>1.1 The prostate .....</b>	<b>1</b>
<b>1.2 Prostate cancer .....</b>	<b>2</b>
1.2.1 Risk factors.....	3
1.2.2 Genomic landscape.....	3
<b>1.3. The androgen receptor.....</b>	<b>5</b>
1.3.1 N-terminal transactivation domain.....	5
1.3.2 DNA binding domain .....	5
1.3.3 Hinge region.....	6
1.3.4 Ligand binding domain.....	6
<b>1.4 Androgen receptor signalling.....</b>	<b>7</b>
<b>1.5 Androgen receptor signalling regulation .....</b>	<b>9</b>
<b>1.6 Prostate cancer treatment strategies.....</b>	<b>10</b>
1.6.1 Targeting the AR signalling pathway.....	12
1.6.2 Chemotherapy .....	15
<b>1.7 Mechanisms of castration resistance .....</b>	<b>16</b>
1.7.1 Hypersensitivity .....	17
1.7.2 Promiscuous ligand binding.....	18
1.7.3 Outlaw pathway.....	19
1.7.4 The bypass pathway.....	19
1.7.5 Aberrant expression of androgen receptor co-regulators.....	20

<b>1.8 AR splice variants</b> .....	<b>20</b>
1.8.1 Origin of AR-Vs.....	21
1.8.2 Regulation of AR-Vs .....	24
1.8.3 Biological and Clinical relevance .....	26
1.8.4 Therapeutically targeting AR-Vs.....	28
<b>1.9 DNA-PKcs</b> .....	<b>30</b>
1.9.1 DNA-PKcs in DNA repair.....	31
1.9.2 DNA-PKcs in transcriptional regulation.....	33
1.9.3 DNA-PKcs and the AR.....	34
1.9.4 DNA-PKcs inhibitors used in this study .....	35
1.9.5 The androgen receptor and the DNA damage response .....	37
1.9.6 DNA-PKcs as a potential regulator of AR-V activity .....	38
<b>Chapter 2 Aims and objectives</b> .....	<b>39</b>
<b>Chapter 3 Materials and Methods</b> .....	<b>40</b>
<b>3.1 Mammalian cell culture</b> .....	<b>40</b>
3.1.1 Subculturing .....	40
3.1.1 Cell storage .....	41
3.1.2 Compounds.....	41
<b>3.2 siRNA and synthetic guide RNA transfection</b> .....	<b>43</b>
<b>3.3 RNA extraction</b> .....	<b>44</b>
<b>3.4 Reverse transcription</b> .....	<b>44</b>
<b>3.5 Quantitative RT-PCR</b> .....	<b>45</b>
<b>3.6 Agarose gel electrophoresis</b> .....	<b>47</b>
<b>3.7 Chromatin Immunoprecipitation</b> .....	<b>47</b>
<b>3.8 RNA Immunoprecipitation</b> .....	<b>49</b>
<b>3.9 Western Blot</b> .....	<b>51</b>
<b>3.10 Immunofluorescence</b> .....	<b>53</b>

<b>3.11 Cell proliferation assays.....</b>	<b>53</b>
3.11.1 Cell counts.....	53
3.11.2 Live cell imaging.....	54
3.11.3 Sulforhodamine B (SRB) growth assay.....	54
<b>3.12 Cell cycle analysis.....</b>	<b>55</b>
<b>3.13 Plasmids.....</b>	<b>55</b>
3.10.2 Bacterial transformation.....	56
3.10.3 Plasmid extraction.....	56
3.10.4 Plasmid transfection.....	57
<b>3.14 Virus production and transductions.....</b>	<b>57</b>
3.14.1 Virus generation.....	57
3.14.2 Mammalian cell transduction.....	57
<b>3.15 Statistical analysis.....</b>	<b>58</b>
<b><i>Chapter 4 Assessing DNA-PKcs as an AR-V coregulator.....</i></b>	<b>59</b>
<b>4.1 Introduction.....</b>	<b>59</b>
4.1.1 Aims.....	63
<b>4.2 Results.....</b>	<b>64</b>
4.2.1 DNA-PKcs expression is significantly upregulated in PCa and CRPC.....	64
4.2.2 DNA-PKcs inhibition reduces proliferation of CRPC cell lines.....	67
4.2.3 DNA-PKcs inhibition reduces AR target gene expression.....	71
4.2.4 DNA-PKcs inhibition in VCaP cells reduces AR-V protein levels.....	79
4.2.5 DNA-PKcs is recruited to AR target genes in CWR22Rv1-AR-EK cells.....	83
4.2.6 DNA-PKcs inhibition does not affect AR recruitment to target genes.....	84
4.2.7 DNA-PKcs inhibition impacts the cell cycle profile of CWR22Rv1 and CWR22Rv1-AR-EK cells.....	88
4.2.8 Validating DNA-PKcs-mediated regulation of AR-V activity using DNA-PKcs knockdown.....	89
4.2.9 DNA-PKcs inhibition with a more selective inhibitor showed similar effects on proliferation and AR-V target gene expression in CWR22RV1-AR-EK cells.....	103

4.2.10 DNA-PKcs inhibition with AZD7648 impacts proliferation and AR-V target gene expression in CWR22RV1-AR-EK cells .....	106
4.2.11 DNA-PKcs inhibition did not significantly downregulate PSA or KLK2 expression in steroid depleted conditions .....	108
<b>4.3 Discussion.....</b>	<b>111</b>
<b><i>Chapter 5 Investigating global transcriptomic effects of DNA-PKcs inhibition and knockdown in CRPC.....</i></b>	<b>119</b>
<b>5.1 Introduction.....</b>	<b>119</b>
5.1.1 Aims .....	121
<b>5.2 Specific materials and methods .....</b>	<b>122</b>
5.2.1 RNA-sequencing experimental set-up .....	122
5.2.2 RNA-sequencing analysis .....	122
5.2.3 Gene set enrichment analysis .....	123
5.2.4 Immunoprecipitation .....	124
5.2.5 Minigene-based splicing analysis.....	124
5.2.6 Differential splicing analysis.....	125
<b>5.3 Results.....</b>	<b>126</b>
5.3.1 Sample validation.....	126
5.3.2 Quality control of RNA-sequencing samples.....	127
5.3.3 Principal component analysis .....	130
5.3.4 Differential gene expression using DESeq2 .....	130
5.3.5 Gene set enrichment analysis reveals pathways that are impacted in response to DNA-PKcs manipulation.....	137
5.3.6 DNA-PKcs inhibition and knockdown downregulates the androgen response .....	144
5.3.7 DNA-PKcs inhibition and knockdown downregulates genes involved in the spliceosome.....	145
5.3.8 Splicing regulation involving DNA-PKcs and RBMX controls androgen receptor variant synthesis..	150
5.3.9 RBMX is a key regulator of AR-V synthesis .....	153
<b>5.4 Discussion.....</b>	<b>165</b>

<b>Chapter 6 Determining AR-V involvement in DNA damage repair.....</b>	<b>176</b>
<b>6.1 Introduction.....</b>	<b>176</b>
6.1.1 CRISPR/Cas9 genome editing.....	179
6.1.2 Aims .....	182
<b>6.2 Specific materials and methods .....</b>	<b>183</b>
6.2.1 Biotin labelling protocol.....	183
6.2.2 Cytoplasmic nuclear extraction.....	184
6.2.3 Pierce 660nm to determine concentration.....	184
6.2.4 Streptavidin enrichment and initial analysis of biotinylation by western blot .....	185
6.2.5 Preparation of samples for mass spectrometry and mass spectrometry data analysis .....	186
6.2.6 Stable cell line generation.....	187
6.2.7 dsODN generation.....	188
<b>6.3 Results.....</b>	<b>189</b>
6.3.1 Validating sgRNAs .....	190
6.3.2 Knock in cell line attempts .....	196
6.3.3 Generation of an APEX2-V7 plasmid.....	203
6.3.4 The APEX2-V7 fusion protein is functional and is recruited to androgen response elements on chromatin.....	205
6.3.5 Optimisation of the labelling protocol in HEK293 and CWR22Rv1-AR-EK cells .....	209
6.3.6 The androgen receptor was identified as the most abundant biotinylated protein in CWR22Rv1-AR-EK cells by mass spectrometry .....	215
<b>6.4 Discussion.....</b>	<b>233</b>
<b>Chapter 7 Conclusions and future work.....</b>	<b>239</b>
<b>Chapter 8 Appendix .....</b>	<b>245</b>
<b>8.1 Splicing genes differentially expressed in response to DNA-PKcs manipulation .....</b>	<b>245</b>
<b>8.2 dsODN sequences .....</b>	<b>248</b>
<b>8.3 Proteins that are increased and decreased in response to ionising radiation.....</b>	<b>249</b>
<b>8.4 British association for cancer research 60<sup>th</sup> anniversary meeting abstract .....</b>	<b>252</b>

**Chapter 9 References..... 253**

## Abbreviations

°C:	Degrees Celsius
μL:	Microlitre
μM:	Micromolar
ADT:	Androgen Deprivation therapy
AR:	Androgen receptor
ARE:	Androgen response element
AR-V:	Androgen receptor variant
ATCC:	American type culture collection
Cas9:	CRISPR-associated endonuclease
CDK:	Cyclin-dependent kinase
cDNA:	Complementary DNA
CE:	Cryptic exon
ChIP:	Chromatin immunoprecipitation
CRISPR:	Clustered regularly interspaced palindromic repeats
CRPC:	Castrate resistant prostate cancer
CTC:	Circulating tumour cell
DBD:	DNA binding domain
DDR:	DNA damage response
DEG:	Differentially expressed genes
DHT:	Dihydrotestosterone
DMSO:	Dimethyl sulfoxide
DNA-PKcs:	DNA-dependent protein kinase, catalytic subunit
DSB:	Double stranded break

dsODN:	Double stranded oligodeoxynucleotide
EDTA:	Ethylenediaminetetraacetic acid
FC:	Fold change
FL-AR:	Androgen receptor full length
FM:	Full media
GnRH:	Gonadotropin-releasing hormone
GR:	Glucocorticoid receptor
GSEA:	Gene set enrichment analysis
HNRNP:	Heterogeneous nuclear ribonucleoprotein
HR:	Homologous recombination
HSP:	Heat shock protein
iBAQ:	intensity based absolute quantification
Indels:	Insertions/deletions
IP:	Immunoprecipitation
IP-MS:	Immunoprecipitation-mass spectrometry
IR:	Ionising radiation
kDa:	Kilodaltons
LBD:	Ligand binding domain
LogFC:	Log 2 fold change
mCRPC:	Metastatic CRPC
mM:	Millimolar
mRNA:	messenger RNA
MS:	Mass spectrometry
mSA:	Monomeric streptavidin
NES:	Normalised enrichment score

NHEJ:	Non-homologous recombination
NLS:	Nuclear localisation signal
NT:	Non-treated
NTD:	N-terminal transactivation domain
PBS:	Phosphate buffered saline
PCa:	Prostate cancer
PCR:	Polymerase chain reaction
Pre-mRNA:	Precursor messenger RNA
PSA:	Prostate specific antigen
qPCR:	quantitative PCR
RBMX:	RNA Binding Motif Protein X-Linked
RIME:	Rapid immunoprecipitation of endogenous proteins
RIP:	RNA immunoprecipitation
RNA:	Ribonucleic acid
qRT-PCR:	quantitative real-time PCR
SDEG:	Significantly differentially expressed genes
SDM:	Steroid depleted media
SDS-PAGE:	Sodium dodecyl-sulphate polyacrylamide gel electrophoresis
siRNA:	Small interfering RNA
sgRNA:	Synthetic guide RNA
snRNP:	Small nuclear ribonucleoproteins
TCGA:	The Cancer Genome Atlas
TLCV2:	Tetracycline-inducible lenti-CRISPR version 2
Tm:	Melting temperature
TSS:	Transcriptional start site

## List of figures

Figure 1.1 Anatomy of the human prostate.....	1
Figure 1.2 Average number of new cases of prostate cancer per year and age-specific incidence rates per 100,000 population, Males, UK, 2013-2015. ....	2
Figure 1.3 The androgen receptor.....	7
Figure 1.4 The androgen signalling pathway.....	9
Figure 1.5 The progression of prostate cancer with timing of different available treatment options for patients at each stage of the disease. ....	11
Figure 1.6 Androgen receptor signalling pathway targeting treatments in prostate cancer ..	15
Figure 1.7 Mechanisms of resistance in prostate cancer.....	17
Figure 1.8 The AR gene and sequence of exons and cryptic exons in several androgen receptor variants arising from alternative splicing. ....	24
Figure 1.9 Structure of DNA-PKcs.....	31
Figure 1.10 Overview of non-homologous end joining pathway.....	33
Figure 1.11 Chemical structures of NU7441, NU5455 and AZD7648.....	36
Figure 4.1 <i>PRKDC</i> is altered in a portion of prostate adenocarcinoma and metastatic prostate adenocarcinoma patients.....	65
Figure 4.2 <i>PRKDC</i> expression is significantly upregulated in tumour samples compared to normal and in metastatic PCa compared to localised.....	67
Figure 4.3 NU7441 significantly effects proliferation of CWR22Rv1-AR-EK, CWR22Rv1, VCaP and LNCaP cells.....	70
Figure 4.4 1 $\mu$ M NU7441 significantly reduces PCa cell proliferation in CWR22Rv1-AR-EK and CWR22Rv1 cells. ....	71
Figure 4.5 DNA-PKcs inhibition reduces AR-V mediated transcription in CWR22RV1-AR-EK cells. ....	72

Figure 4.6 DNA-PK inhibition reduces transcription of the AR-V target gene <i>UBE2C</i> in CWR22Rv1 cells after 24 hours with enzalutamide treatment.....	75
Figure 4.7 DNA-PKcs inhibition reduces AR mediated transcription post 24-hour pre-enzalutamide treatment in CWR22Rv1 cells.....	76
Figure 4.8 DNA-PKcs inhibition reduces transcription of some AR-V target genes in VCaP cells. ....	78
Figure 4.9 DNA-PKcs inhibition reduces AR-V protein levels in VCaP cells.....	81
Figure 4.10: DNA-PKcs inhibition reduces AR-V protein levels in VCaP cells, that can be partially blocked with the addition of a proteasome inhibitor.....	81
Figure 4.11 AR CHIP showed no change in recruitment to regulatory regions in AR introns in VCaP cells.....	83
Figure 4.12: DNA-PKcs is recruited to AR regulatory elements in CWR22Rv1-AR-EK cells, that is largely refractory to inhibition with 1 $\mu$ M NU7441.....	84
Figure 4.13 NU7441 does not affect AR recruitment to cis-regulatory elements of some key AR-regulated genes.....	86
Figure 4.14 DNA-PKcs is not dependent on AR presence at AR target genes in CWR22Rv1-AR-EK.....	87
Figure 4.15 DNA-PKcs inhibition does not significantly affect the cell cycle profile of CWR22Rv1-AR-EK and CWR22Rv1 cells after 96 hours treatment at 1 $\mu$ M.....	89
Figure 4.16 DNA-PKcs knockdown did not mirror pharmacological inhibition of DNA-PKcs inhibition on AR gene expression analysis.....	91
Figure 4.17 DNA-PKcs knockdown did not mirror pharmacological inhibition of DNA-PKcs inhibition on AR gene expression analysis.....	92
Figure 4.18: DNA-PKcs knockdown did not mirror pharmacological inhibition of DNA-PKcs inhibition on AR gene expression analysis.....	93
Figure 4.19 DNA-PKcs knockdown showed varying results across the PCa cell lines CWR22Rv1-AR-EK, CWR22Rv1 and VCaP.....	95

Figure 4.20 DNA-PKcs knockdown using the deconvoluted Dharmacon SmartPool showed differing effects on proliferation of CWR22Rv1-AR-EK, CWR22Rv1 and VCaP cells.....	96
Figure 4.21 DNA-PKcs knockdown using the deconvoluted Dharmacon SmartPool showed differing effects on AR target gene expression in CWR22Rv1-AR-EK cells. ....	98
Figure 4.22 DNA-PKcs knockdown using the deconvoluted Dharmacon SmartPool showed differing effects on AR target gene expression in CWR22Rv1 cells. ....	99
Figure 4.23: DNA-PKcs knockdown using the deconvoluted Dharmacon SmartPool showed differing effects on AR target gene expression in VCaP cells.....	100
Figure 4.24 DNA-PKcs knockdown using siDNA-PK2-4 caused a significant decrease in proliferation of CWR22Rv1-AR-EK and CWR22Rv1 cells.....	102
Figure 4.25 DNA-PKcs knockdown using siDNA-PK2-4 downregulated expression of <i>UBE2C</i> , but not <i>PSA</i> , in CWR22Rv1-AR-EK and CWR22Rv1 cell lines. ....	103
Figure 4.26 DNA-PKcs inhibition using NU5455 causes a significant reduction in cell count as a consequence of G1 arrest.....	104
Figure 4.27 DNA-PKcs inhibition with NU5455 caused a significant reduction in AR-V target genes in CWR22RV1-AR-EK cells and VCaP cells. ....	105
Figure 4.28 DNA-PKcs inhibition with AZD7648 significantly reduces CWR22Rv1-AR-EK proliferation and several AR-V target genes. ....	107
Figure 4.29 CWR22Rv1-AR-EK cells are sensitive to DNA-PKcs inhibition with three DNA-PKcs inhibitors.....	108
Figure 4.30 DNA-PKcs inhibition in CWR22Rv1-AR-EK cells grown in steroid-depleted media significantly reduced the expression of AR-V only target genes and proliferation. ....	110
Figure 5.1 A notable number of genes in response to DNA-PKcs inhibition in C4-2 cells overlap with the AR-V transcriptome in CWR22Rv1-AR-EK cells. ....	120
Figure 5.2 RNA-sequencing experimental set up to determine the global transcriptomic effect on CWR22Rv1-AR-EK cells upon DNA-PKcs inhibition with NU7441, NU5455 and AZD7648 and knockdown with an siRNA pool of siDNA-PK2-4. ....	122
Figure 5.3 Interpretation of an enrichment plot from gene set enrichment analysis .....	123

Figure 5.4 Minigene construct diagram with the resulting mRNA for mini-AR-V7 and mini-FL-AR with the primer pair locations for each product. ....	125
Figure 5.5 NU7441, NU5455, AZD7648 and knockdown of DNA-PKcs all show similar effects on CWR22Rv1-AR-EK proliferation after 5 days and selectively impact AR/AR-V target gene expression.....	127
Figure 5.6 FastQC report generated with MultiQC showed the RNA-sequencing data passed the necessary quality controls.....	129
Figure 5.7 Principal component analysis of RNA-sequencing data shows clustering of intra-experimental repeats of each experimental arm.....	130
Figure 5.8 MA plots show the number of up and down regulated genes in response to DNA-PKcs knockdown and inhibition with NU7441, AZD7648 and NU5455.....	132
Figure 5.9 Comparison of overlapping significantly differentially expressed genes in response to DNA-PKcs inhibition shows <i>TWIST1</i> and <i>PDK4</i> are commonly down- and upregulated, respectively.....	133
Figure 5.10 Significantly differentially expressed genes comparisons between the AR-V transcriptome and genes differentially expressed in response to DNA-PKcs inhibition and depletion reveal a significant overlap between AR-V and DNA-PKcs knockdown transcriptomes.....	135
Figure 5.11 Comparisons between common significantly downregulated and upregulated genes between siAR-V and siDNA-PKcs show a considerable number of SDEGs shared between the two transcriptomes. ....	135
Figure 5.12 Heat map showing significant differentially expressed genes in response to DNA-PKcs inhibition with NU5455 and DNA-PKcs depletion.....	136
Figure 5.13 DNA-PKcs regulates several KEGG gene-sets. ....	141
Figure 5.14 DNA-PKcs regulates dysregulation to several hallmark gene sets. ....	143
Figure 5.15 DNA-PKcs depletion and inhibition causes a significant downregulation of the androgen response in CWR22Rv1-AR-EK cells. ....	144

Figure 5.16 Differential splicing analysis reveals genes that are significantly alternatively spliced in response to DNA-PKcs knockdown .....	147
Figure 5.17 DNA-PKcs regulated genes involved in the spliceosome, which are upregulated in response to darolutamide in VCaP cells.....	149
Figure 5.18 RBMX was significantly upregulated in response to darolutamide in VCaP cells and in tumour vs normal samples. ....	150
Figure 5.19 DNA-PKcs inhibition and depletion decreases expression of several AR-V transcripts.....	151
Figure 5.20 Validation that RBMX identified from RNA-sequencing experiment are downregulated by DNA-PKcs inhibition and knockdown in separate independent experiments in CWR22Rv1-AR-EK cells. ....	152
Figure 5.21 DNA-PKcs mRNA expression is correlated with RBMX mRNA expression. ....	153
Figure 5.22 RBMX depletion significantly reduces AR-V and AR-V target gene expression and impacts PCa cells proliferation. ....	155
Figure 5.23 RBMX depletion significantly reduces AR and AR-V protein and transcript levels. ....	156
Figure 5.24. RBMX overexpression marginally increases AR-V mRNA and AR-V target gene expression.....	157
Figure 5.25 RBMX CRISPR knockout was not achieved in CWR22Rv1-AR-EK-Cas9i cells. ....	159
Figure 5.26 DNA-PKcs is recruited to regions upstream of the <i>RBMX</i> transcriptional start site. ....	160
Figure 5.27 RNA Immunoprecipitation shows RBMX enrichment at FL-AR mRNA transcripts. ....	162
Figure 5.28 RBMX depletion significantly impacts splicing of FL-AR minigene.....	163
Figure 5.29 RBMX depletion does not impact transcription of AR pre-mRNA. ....	164
Figure 5.30 RBMX is involved in co-transcriptional splicing of pre-mRNA.....	171

Figure 5.31 RBMX depletion decreases FL-AR/AR-V and AR-V7 protein levels in CWR22Rv1 cells .....	174
Figure 5.32 Proposed mechanism of DNA-PKcs modulated AR/AR-V splicing. ....	175
Figure 6.1 Mechanism of CRIPSR genome editing .....	181
Figure 6.2 Diagrammatic representation of proximity labelling reaction.....	184
Figure 6.3 Flow chart of proximity labelling and identification protocol.....	185
Figure 6.4 AR gene for insertion of APEX2 at the start codon of AR exon 1.....	190
Figure 6.5 Editing efficiency in CWR22Rv1-AR-EK-Cas9 and CWR22Rv1-AR-EK-iCas9 cell lines using AR Exon 1 targeting sgRNA. ....	192
Figure 6.6 TLCV2 sequence showing location of mSA insertion and sequencing showing successful incorporation of mSA sequence.....	194
Figure 6.7 Validation of a CWR22Rv1-AR-EK-iCas9-mSA inducible cell line .....	195
Figure 6.8 Diagrammatic representation of APEX2 knock-in pipeline .....	196
Figure 6.9 Screening of CWR22Rv1-AR-EK cells after puromycin selection.....	197
Figure 6.10 Removal of puromycin resistance cassette from LCV2 plasmid. ....	198
Figure 6.11 APEX2-AR-V knock-in screening of puromycin-selected clones and biotinylation test of single clone.....	200
Figure 6.12 Cell sorting of RFP positive clones and western blot screening of sorted clones .....	202
Figure 6.13 An APEX2-AR-V7 encoding plasmid was successfully cloned.....	204
Figure 6.14 Successful expression of APEX2-AR-V7 fusion protein in HEK293 cells after transfection with APEX2-AR-V7 plasmid. ....	205
Figure 6.15 Optimising the transfection efficiency of the APEX2-AR-V7 plasmid .....	207
Figure 6.16 FLAG antibody-based ChIP shows APEX2-AR-V7 fusion protein is recruited to androgen response elements on chromatin .....	208

Figure 6.17 Optimisation of the proximity biotinylation protocol in HEK293 and CWR22Rv1-AR-EK cells. ....	211
Figure 6.18 Analysis of $\gamma$ H2AX foci at several time points post 2 Gy irradiation .....	212
Figure 6.19 Cytoplasmic nuclear fractionation of CWR22Rv1-AR-EK cells enriches for nuclear biotinylated proteins and removes highly expressed endogenously biotinylated protein from eluates. ....	213
Figure 6.20 Optimised expression/ labelling protocol of AR-V7 interactome experiment. ...	214
Figure 6.21 Biotinylation of proteins was validated in samples that were subsequently analysed by mass spectrometry .....	217
Figure 6.22 Heat map of identified proteins in each experimental arm across three replicates .....	219
Figure 6.23 Principal component analysis of proteomics data of each of three replicates...	221
Figure 6.24 APEX2-AR-V7 interacting proteins identified by mass spectrometry in the presence and absence of ionising radiation.....	222
Figure 6.25 APEX2-AR-V7 interacting proteins.....	223
Figure 6.26 Proteins with a riBAQ value >1.5-fold higher in response to irradiation. ....	225
Figure 6.27 Proteins with a riBAQ >1.5 fold higher in response to irradiation.....	227
Figure 6.28 The components of the DNA-PK holoenzyme were identified as interactors of AR-V7 .....	228
Figure 6.29 Comparisons of the APEX2-AR-V7 interactome identified using proximity labelling and AR-V7 and FL-AR interactomes identified by RIME.....	229
Figure 6.30 Comparisons of the APEX2-AR-V7 interactome identified using proximity labelling and the published AR-V <sup>567es</sup> interactomes identified by RIME. ....	230
Figure 6.31 Several proteins were identified as part of the AR-V7 and DNA-PKcs interactome .....	231

## List of tables

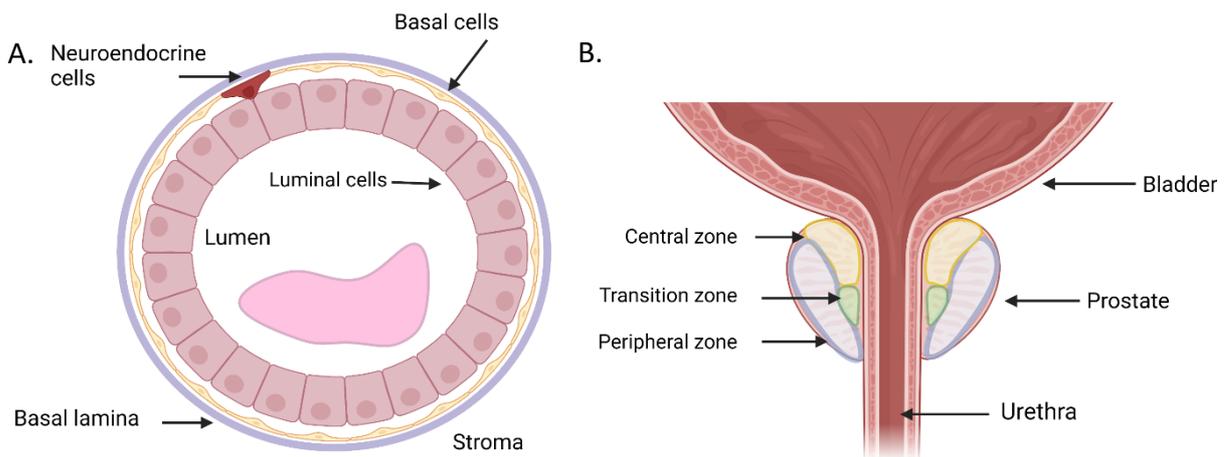
Table 3.1 Cell lines .....	41
Table 3.2 List of compounds used with their respective concentrations .....	42
Table 3.3 Small interfering RNA and synthetic guide RNA sequences.....	43
Table 3.4 Components of Reverse transcription reaction with their respective volumes .....	45
Table 3.5 Components used per well for qRT-PCR with their respective volumes.....	46
Table 3.6 Primer sequences used in qPCR. ....	46
Table 3.7 CHIP solutions and buffers .....	48
Table 3.8 Primer Sequences for CHIP analysis.....	49
Table 3.9 RNA immunoprecipitation buffers .....	51
Table 3.10 Western blot solutions and reagents .....	52
Table 3.11 Antibodies .....	52
Table 5.1 Lists of differentially expressed and significant differentially expressed genes in response to DNA-PKcs inhibition and knockdown in CWR22Rv1-AR-EK cells .....	132
Table 5.2 Number of significantly differentially expressed KEGG pathways in response to DNA-PKcs inhibition and depletion of DNA-PKcs.....	137
Table 5.3 Downregulated KEGG pathways with their normalised enrichment scores in response to three DNA-PKcs inhibitors and depletion of DNA-PKcs .....	137
Table 5.4 Commonly downregulated KEGG pathways between DNA-PKcs inhibition with NU5455 and depletion of DNA-PKcs .....	139
Table 5.5 Number of significantly differentially expressed Hallmark pathways in response to DNA-PKcs inhibition and depletion of DNA-PKcs .....	141
Table 5.6 Common downregulated Hallmark pathways with their normalised enrichment scores in response DNA-PKcs inhibition with NU5455 and depletion of DNA-PKcs.....	142
Table 5.7 Commonly downregulated hallmark pathways between DNA-PKcs inhibition and depletion.....	143

Table 5.8 Overlapping ‘core enriched’ spliceosome genes between NU5455 and siDNA-PKcs. .....	146
Table 5.9 Primer sequences for amplification of genomic DNA for TIDE analysis.....	158
Table 6.1 Primer Sequences to generate the dsODN template for the APEX2 knock in. ....	189
Table 6.2 Primer sequences used to amplify region around sgRNA mediated cut side of <i>AR</i> gene for TIDE analysis.....	191
Table 6.3: Primers used for the amplification of the APEX2 sequence to be inserted into the AR-V7 donor plasmid.....	203
Table 6.4 Top two proteins identified from first CWR22Rv1-AR-EK labelling experiment in each experimental arm. ....	215
Table 6.5 Number of proteins identified by mass spectrometry for each proximity labelling replicate.....	218
Table 6.6 Proteins that abundances are statistically significantly altered in response to irradiation. ....	223
Table 6.7 Top 10 biological processes that are enriched in the list of proteins that are more abundant AR-V7 interactors in response to irradiation. ....	226

# Chapter 1 Introduction

## 1.1 The prostate

The prostate is a small secretory gland located at the base of the bladder and surrounds the urethra. It is part of the male reproductive system, and its primary function is to secrete important components of the seminal fluid which is required to protect and maintain spermatozoa. The typical glandular architecture of the prostate is made up of proximal luminal cells surrounding the glandular lumen which are adjacent to a layer of basal cells interspersed with a smaller number of neuroendocrine cells (Figure 1.1A). These epithelial cells are surrounded by smooth muscle cells. The human prostate can be divided into three sub-anatomical zones; the peripheral, central and transition zones (Figure 1.1B). The peripheral zone is the largest region of the prostate, accounting for approximately 70% of the glandular tissue within the prostate and is the most common site of prostate cancer (PCa). The central zone is smaller than the peripheral zone and surrounds the ejaculatory ducts. The smallest zone is the transition zone which surrounds the urethra and grows in size with age. The transition zone is where benign prostatic hyperplasia (BPH) arises; a common enlargement of the prostate in aging men, with over 90% of men over the age of 80 being affected (Laczko *et al.*, 2005).

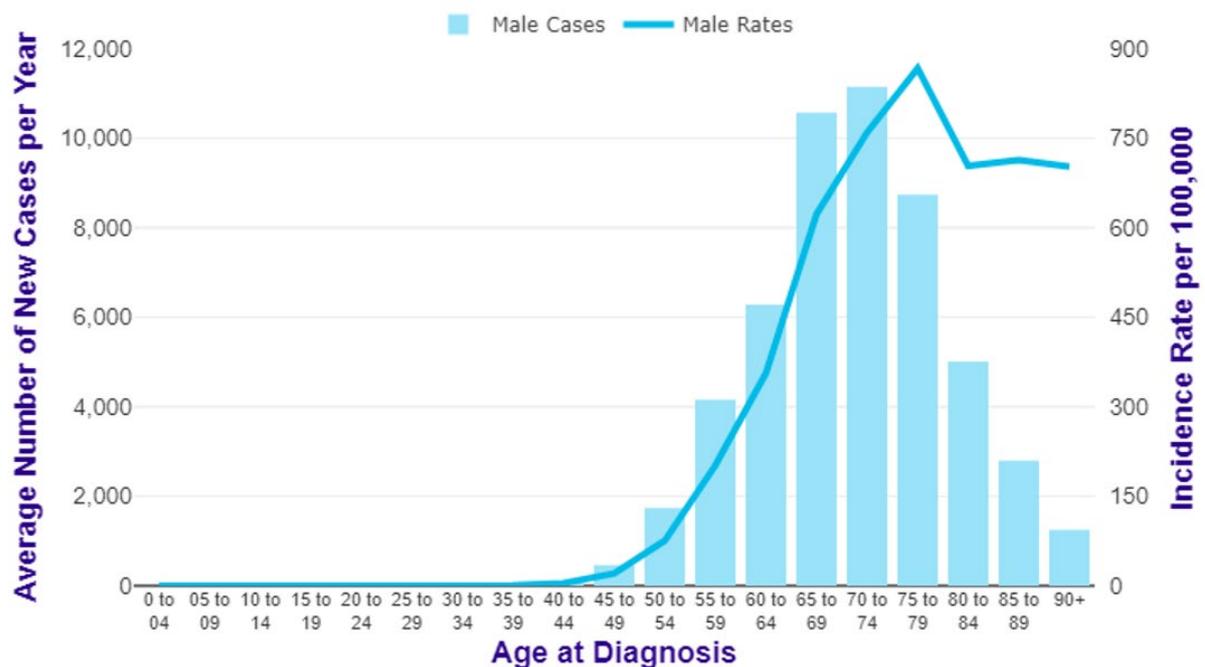


**Figure 1.1 Anatomy of the human prostate**

A. Illustration of the human prostate and its location with the three sub-anatomical zones B. Cellular organisation within the prostate duct. Adapted from (Shen and Abate-Shen, 2010)

## 1.2 Prostate cancer

PCa is the most common male cancer and the second most common cause of cancer-related deaths in males in western countries. In 2017, over 48,000 men were diagnosed with PCa in the UK, equating to 130 men per day (CRUK). Over 11,000 men die from PCa per year, accounting for 13% of all cancer related deaths, making it a major health concern. The lifetime risk of developing PCa in the UK is around 1 in 6 and the incidence has dramatically increased over the past 40 years. This is a result of an aging population, improved healthcare and increased prostate specific antigen (PSA) screening. A major spike in PCa diagnoses between the late 1980s and early 1990s was the result of detection of asymptomatic PCa because of the PSA screening (CRUK, Siegel *et al.*, 2016). The majority of PCa originates in the peripheral zone of the prostate with around 70-80% of tumours arising here. Around 20% of PCa tumours originate in the transition zone and around 5% of PCa tumours originate in the central zone and tend to be more aggressive and more likely to metastasise (Cohen *et al.*, 2008).



**Figure 1.2 Average number of new cases of prostate cancer per year and age-specific incidence rates per 100,000 population, Males, UK, 2013-2015.**

Prostate cancer is the most common male cancer in the UK, with over 48,000 cases each year. Its incidence rate increases with age, with a peak age at diagnosis of 65-69 years. Image taken from cancer research UK (CRUK).

### 1.2.1 Risk factors

The main PCa risk factors include age, ethnic origin and family history. PCa incidence is strongly associated with age; the majority of cases occur in men over 65 with a peak incidence occurring in men aged 75-79 (Figure 1.2). Ethnic background is another major risk factor for PCa; several epidemiological studies have shown that black men have a higher lifetime risk of developing PCa than white men; The Prostate Cancer in Ethnic Subgroups study also showed that in the UK, black men are more likely to be diagnosed on average 5 years younger (Ben-Shlomo *et al.*, 2008). Finally, the lifetime risk of men with a family history of PCa can increase 2.3-2.5-fold if they have an affected first-degree relative (father/brother), and 1.2-fold if they have a mother or sister affected by breast cancer (Kicinski *et al.*, 2011, Chen *et al.*, 2008b).

### 1.2.2 Genomic landscape

The genomic landscape of PCa is complex with several gene mutations and chromosomal rearrangements commonly identified. Many studies have been carried out to define the genetic abnormalities that occur in all stages of PCa in order to better characterise the PCa genome. A 2015 study identified that around 90% of metastatic castrate-resistance PCa (mCRPC) harbour mutations that could be targetable, 62.7% of individuals had mutations in the *androgen receptor (AR)* gene and 65% had aberrations in cancer-related genes that are non-AR related (Robinson *et al.*, 2015).

Gene fusions between the E-twenty-six (ETS) family members: *erythroblast transformation-specific transcription factor (ERG)* and *ETS variant 1 (ETV1)*, and the androgen-regulated promoter of *TMPRSS2* have been detected in up to 79% of PCa samples, first identified by Tomlins *et al.*, (Tomlins *et al.*, 2005) and as such is the most common gene fusion of any human cancer type. The fusion converts ERG transcription factor genes to become androgen-regulated which enhances their expression and enables oncogenesis. *TMPRSS2-ERG* fusions are commonly identified in early-stage disease indicative of a role for altered ERG activity in facilitating PCa initiation. As expected, other oncogenic mutations are required alongside the *TMPRSS2:ERG* fusion to lead to carcinogenesis, such as *p53* and *Rb* inactivation. Tumour suppressors TP53 and RB1 are inactivated by genomic loss in approximately 50% and 20% of advanced PCa patients, respectively (Robinson *et al.*, 2015). *PTEN* loss, through deletion and mutation, occurs in approximately 40% of advanced PCa patients; causing aberrant PI3K/AKT-

mTOR signalling and is therefore correlated to a poorer prognosis and increased likelihood of metastasis (Robinson *et al.*, 2015).

Alterations in DNA damage repair (DDR) genes are common in PCa with over 20% of high-grade tumours harbouring DDR germline and/or somatic mutations in genes such as *BRCA2*, *BRCA1*, *ATM* or *CHEK2* (Lang *et al.*, 2019, Abida *et al.*, 2017). Mutations in this pathway make the tumour particularly sensitive to radiotherapy or drugs targeting other components of the DDR, and more susceptible to cell death by inducing synthetic lethality. The use of Poly (ADP)-ribose polymerase (PARP) inhibitors is an example of exploiting such genetic vulnerability in cancer patients that harbour germline homologous recombination repair mutations, such as *BRCA1* and *BRCA2*. The use of the PARP inhibitor Olaparib in mCRPC patients after progression following anti-androgenic drugs and taxane-based chemotherapy was FDA approved in 2020 (Hussain *et al.*, 2020).

A small percentage of PCa cases have a genetic predisposition linked to a more aggressive disease, such as a mutation in the *BRCA2* gene, which also occurs in 1-2% of sporadic cases (Castro and Eeles, 2012). This is because patients with defective *BRCA2* display impaired DDR leading to genomic instability and increased likelihood of developing cancer. A *BRCA2* germline mutation is the genetic event that causes the highest PCa risk with an ~8.6-fold increased risk of developing PCa before the age of 65 (Kote-Jarai *et al.*, 2011). Importantly, *BRCA2* mutant patients respond well to the PARP inhibitors Olaparib and Rucaparib either as a monotherapy or in combination with anti-androgen therapy (Li *et al.*, 2017, Mateo *et al.*, 2015).

Given the implications of certain mutations that occur in PCa, genetic testing in PCa patients would inform prognosis and facilitate personalised medicine to significantly improve clinical management. This is not routinely done, especially in early-stage PCa patients, however in later-stage PCa patients, molecular characterisation will be critical to stratify patients onto targeted treatments such as PARP inhibitors.

### 1.3. The androgen receptor

The androgen receptor (AR), a member of the steroid hormone family of nuclear receptors, is a 110 kDa protein encoded by the *AR* gene located on Xq11-12. It is a ligand-dependent transcription factor that mediates the action of androgens, such as testosterone, essential for male sexual development and differentiation.

The *AR* gene contains 8 exons that encode a 919 amino acid protein defined by three main functional domains: an N-terminal transactivation domain (NTD), a DNA-binding domain (DBD) and the C-terminal ligand binding domain (LBD) which is connected to the DBD by a hinge region, shown in Figure 1.3.

#### 1.3.1 N-terminal transactivation domain

Exon 1 encodes the entire NTD, which is the largest domain in the AR, making up nearly 60% of the AR protein. This is the most flexible, unstructured region, and the least conserved domain amongst members of the nuclear receptor family (Claessens *et al.*, 2008). This has made defining crystal structures of the NTD difficult and hence our understanding of the 3-dimensional folding of the domain is limited. One of the two transcriptional activation functions (AF) present in the AR is located in the NTD and is termed AF-1. The AF-1 is androgen-independent and is required for the transcriptional activity of the AR. Within the AF-1 are two transcriptional activation units (TAU), TAU-1 and TAU-5, that are essential for AR transcriptional activity (Jenster *et al.*, 1995). Two structurally important motifs are <sup>23</sup>FQNL<sup>27</sup>, contained in TAU1, and <sup>433</sup>WHTLF<sup>427</sup> in TAU-5, which are essential for the N/C-terminal interaction between the AR NTD and LBD that is critical for AR transcriptional competency (Steketee *et al.*, 2002).

#### 1.3.2 DNA binding domain

Exons 2-3 encode the DBD which is the most highly conserved domain among the steroid hormone nuclear receptor family. It consists of two zinc fingers that are required for mediating selective DNA binding at promoter and enhancer regions of canonical AR-regulated genes. The first zinc finger recognises and binds target cognate sequences, termed androgen receptor elements (AREs), contained within *cis*-regulatory elements of target genes. The sequence of three amino acids in the DBD, termed proximal ('P') box, also known as the 'recognition helix'

are highly conserved between other nuclear hormone receptors including the glucocorticoid receptor (GR), progesterone receptor (PR), and mineralocorticoid receptor (MR) and as such this helix recognises similar DNA response elements. However, selective AREs have been identified that allow for the specific activation by the AR over other nuclear hormone receptors (Shaffer *et al.*, 2004, Claessens *et al.*, 1996, Verrijdt *et al.*, 1999). The second zinc finger interacts with the phosphate backbone to stabilise protein-DNA contacts. The distal ('D') box is also present in the DNA binding domain which is important for receptor dimerization (Shaffer *et al.*, 2004).

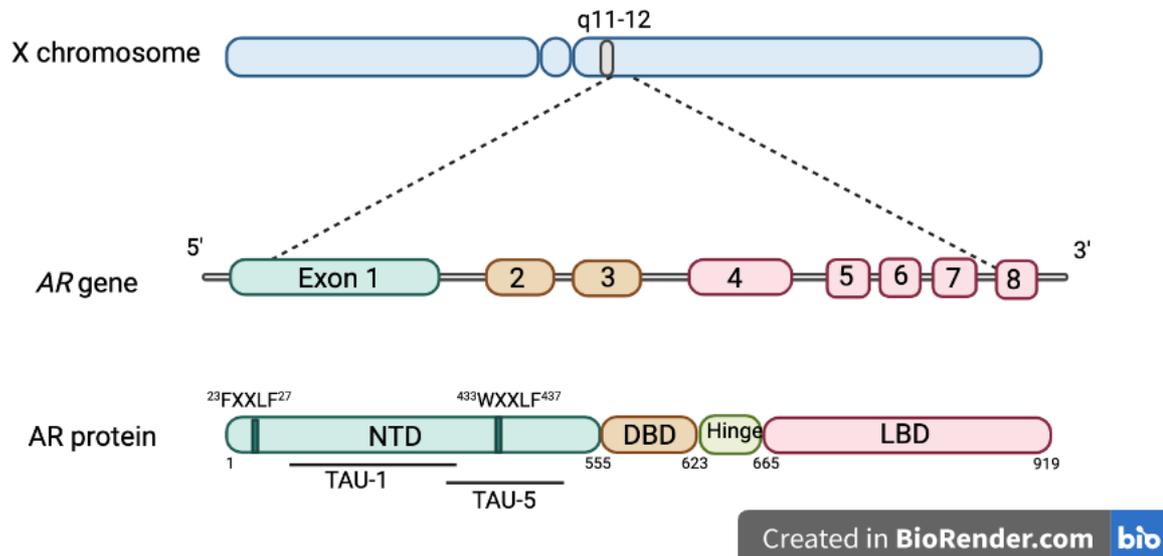
### 1.3.3 Hinge region

The 3' portion of exon 3 and 5' portion of exon 4 encode the hinge region which links the DBD and LBD and is involved in regulating multiple AR activities, including nuclear translocation, transcriptional potency and degradation of the receptor. Translocation to the nucleus is mediated by a nuclear localisation signal (NLS) which spans parts of the DBD and hinge region (Zhou *et al.*, 1994). The hinge region is also involved in N- and C- terminal interaction, DNA binding and the recruitment of co-activators (Haelens *et al.*, 2007). It is also a target site for post-translational modifications such as methylation, acetylation and ubiquitination (Coffey and Robson, 2012). For example, lysine 632 is methylated by histone methyltransferase SET9 and this increases transcriptional activity of the AR by enhancing chromatin recruitment (Gaughan *et al.*, 2011). Located between the hinge and LBD is a nuclear export signal, responsible for AR exportation upon ligand withdrawal (Gong *et al.*, 2012).

### 1.3.4 Ligand binding domain

Exons 4-8 encode the carboxy-terminal LBD which is composed of 11  $\alpha$ -helices and 1  $\beta$ -sheet, which is distinct from other nuclear receptors that contain an additional  $\alpha$ -helices, helix 2. The ligand binding pocket is made up of helices 3, 5, 10 and 11 which, upon ligand (testosterone, di-hydrotestosterone) binding, undergoes a conformational change whereby the highly flexible helix 12 repositions over the ligand binding pocket to present a hydrophobic co-activator interacting surface (Heery *et al.*, 1997). It is this region of the LBD that binds short amphipathic LXXLL motifs, contained within numerous co-activators, such as the SRC-1 and p300, to enable AR co-activation (Heery *et al.*, 1997). The second AF, AF-2, is located in this domain and is responsible for interactions between several transcriptional co-regulatory

proteins, as well as the AR N-terminus to stabilise bound androgens which is required for full transcriptional activity of the AR (Warnmark *et al.*, 2003).



**Figure 1.3 The androgen receptor.**

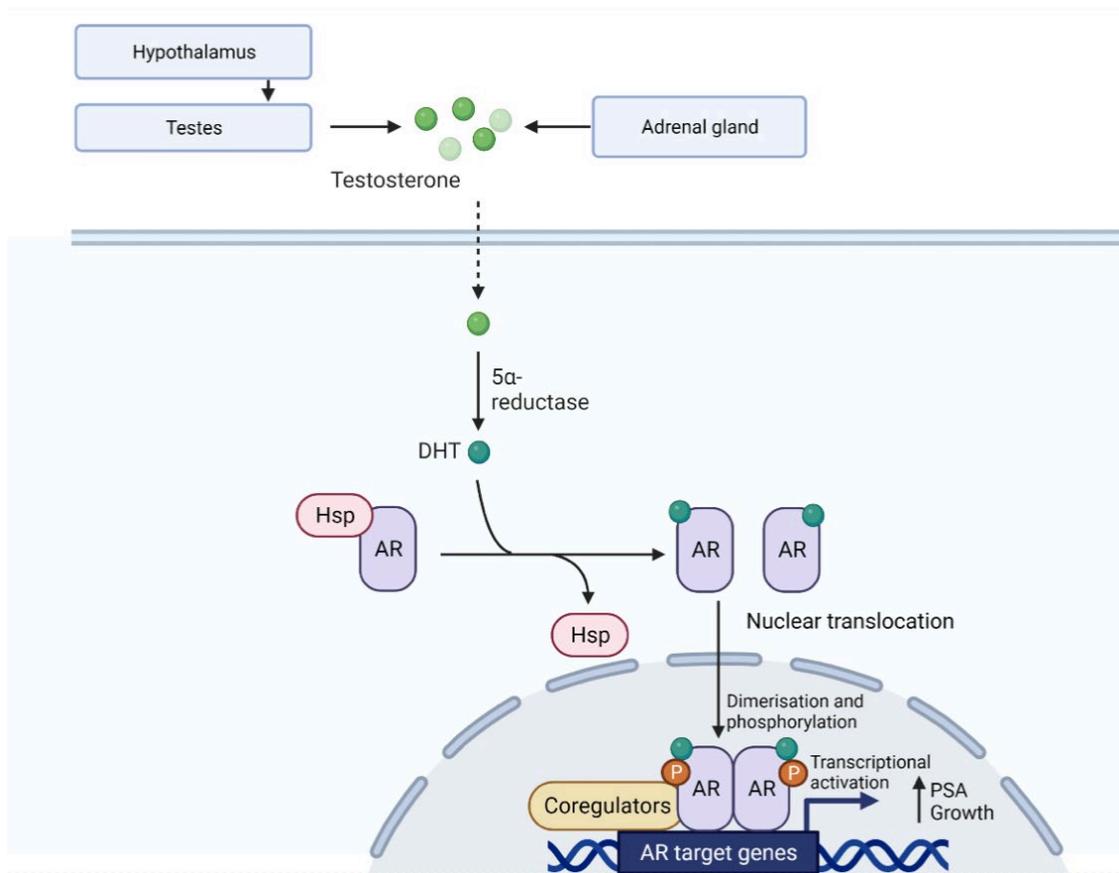
Diagrammatic representation of the *AR* gene with its location on the X chromosome, and below, the AR protein showing its main structural domains. The *androgen receptor* gene (alternatively known as *NR3C4*) is located on the long arm of the X chromosome (q11-12). It is composed of eight exons and encodes the androgen receptor protein that is 919 amino acids. The protein includes an N-terminal transactivation domain (NTD), DNA binding domain (DBD), hinge region and ligand binding domain (LBD). The NTD contains two transcriptional activation units (TAU-1 and TAU-5) that are critical for AR transcriptional activity.

#### 1.4 Androgen receptor signalling

Prostate growth and development are reliant upon AR signalling which is controlled by the hypothalamic-pituitary-gonadal (HPG) axis. The HPG axis is activated by the release of gonadotropin releasing hormone (GnRH) from the hypothalamus which stimulates secretion of luteinising hormone (LH) from the anterior pituitary gland. LH subsequently stimulates synthesis of testosterone in the testes which in turn inhibits the further secretion of GnRH and LH via a negative feedback loop (Tilbrook and Clarke, 2001). Testosterone is secreted by Leydig cells in the testes and is the major circulating androgen. The testes accounts for the production of approximately 90% of circulating testosterone, whilst the other 5-10% is synthesised by the adrenal glands. The majority of testosterone is bound to albumin or sex hormone-binding globulin (SHBG) in the bloodstream, with unbound testosterone able to diffuse into cells of the prostate (Rosner *et al.*, 1991). The AR signalling axis drives male development and is

essential for prostate gland formation and normal function, as well as being fundamental in PCa initiation and progression.

Prior to androgen binding, the AR is maintained in an inactive state in the cytoplasm by binding to heat shock proteins (HSPs) (Figure 1.4). Testosterone enters prostate cells and is converted to the highly potent 5 $\alpha$ -dihydrotestosterone (DHT) by 5 $\alpha$ -reductase. Upon DHT binding to the AR LBD, HSPs are displaced, driving an intramolecular interaction between the N- and C-termini of the monomeric AR. This conformational change exposes the NLS which facilitates translocation of the AR into the nucleus, where it dimerises, via the DBD, hinge and intermolecular N-C-terminal interactions, permitting binding to AREs in promoter and enhancer regions of AR-regulated genes, including *PSA* and *TMPRSS2* (Rosner *et al.*, 1991). Co-recruitment of numerous transcriptional co-regulator proteins to sites of AR binding facilitates recruitment of RNA polymerase II and subsequent transcription of proximal genes involved in prostate growth, homeostasis and transformation. The production of PSA is an indicator AR activity hence it is used as a biomarker in active surveillance of PCa progression.



**Figure 1.4 The androgen signalling pathway**

Testosterone synthesised in the testes or adrenal gland is transported to target tissues. Upon entering prostate cells, testosterone is converted to the more potent metabolite DHT by 5 $\alpha$ -reductase and binds to the ligand binding domain of the androgen receptor, causing dissociation of HSPs. The AR then translocates to the nucleus, dimerises, and binds to androgen response elements at the promoter regions of AR target genes such as prostate specific antigen.

### 1.5 Androgen receptor signalling regulation

As well as regulation by androgen binding, over 150 co-regulatory molecules have been identified to date that have been classified as co-activators or co-repressors of AR transcriptional activity. Although varied in function, co-activators typically facilitate: (i) AR-DNA contacts; (ii) association with the RNA polymerase II machinery; and (iii) subsequent transcription of target genes. In contrast, co-repressors work to antagonise AR-mediated transcription at numerous levels. A number of co-regulatory proteins are responsible for catalysing post-translational modifications of the AR, such as ubiquitination, phosphorylation, acetylation, methylation and sumoylation (Coffey and Robson, 2012) that control AR transcriptional competency. Importantly, many of these proteins become dysregulated in PCa

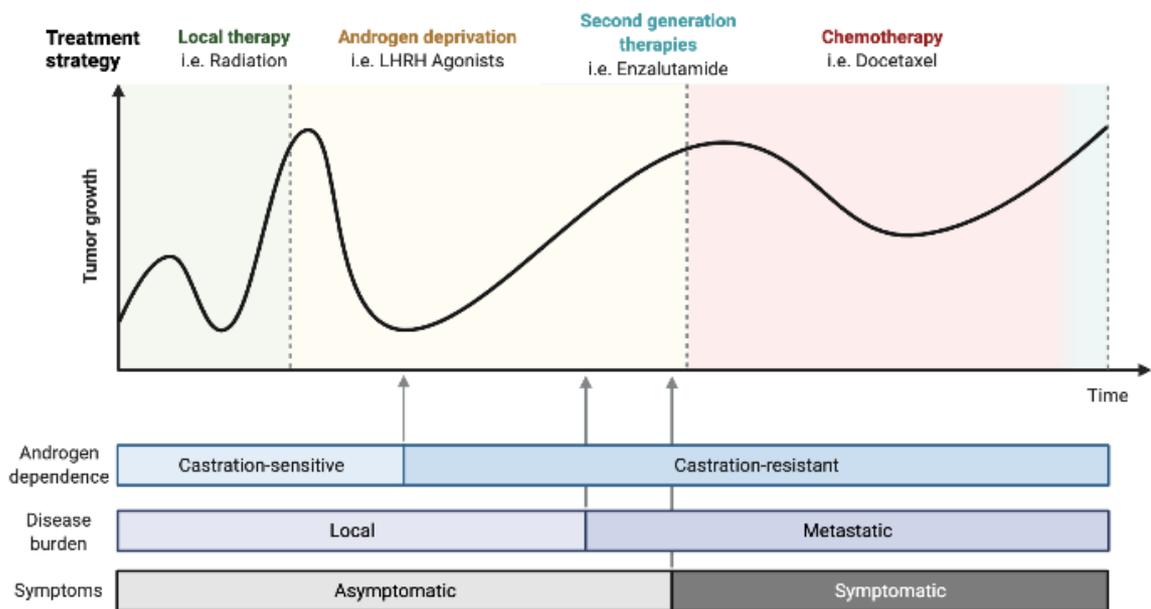
leading to aberrant AR activity and are therefore important in relation to the development of therapies. Phosphorylation accounts for most post-translational modifications of the AR with at least 19 phosphorylation sites having been identified; most of which are located in the NTD. Proteins such as members of the cyclin-dependent kinase (CDK) family and AKT are examples of kinases that phosphorylate the AR to modulate activity, specifically at serine 81 (Gordon *et al.*, 2010) (Lin *et al.*, 2001). This promotes AR protein stability and nuclear localisation. It has been shown that inhibition of a broad range of CDKs (CDK1, CDK2, CDK5, CDK7 and CDK9) with roscovitine diminishes AR phosphorylation at serine 81 and potentiates AR degradation, a process that can be rescued with the addition of the proteasome inhibitor MG132 (Hsu *et al.*, 2011). CDK7 is involved in the activation of the multi-subunit Mediator (MED1) complex that is critical to the regulation of transcription through mediating RNA polymerase II interactions with transcription factors (Allen and Taatjes, 2015). CDK7 was shown to phosphorylate the MED1 coactivator complex to promote AR interaction and subsequent transcription of AR-target genes. As such, inhibition of CDK7 caused loss of both MED1 and AR recruitment to chromatin in VCaP and LNCaP cell lines (Rasool *et al.*, 2019).

Bromodomain (BD) containing 4 (BRD4) is another example of an AR co-regulatory protein that is a potential therapeutic target in PCa. Inhibition of BRD4 with JQ1 blocks the transcriptional activity of the AR, and microarray analysis in several PCa cell lines showed significant reduction in canonical AR target gene expression upon JQ1 treatment (Asangani *et al.*, 2014). Furthermore, JQ1 treatment diminished growth of several AR positive CRPC cell lines and the VCaP tumour xenograft model *in vivo* (Asangani *et al.*, 2014).

## 1.6 Prostate cancer treatment strategies

Androgens play a vital role in the proliferation, differentiation and metastasis of PCa; hence treatment is majorly focussed on inhibiting AR activity. This is achieved, in part, by starving the AR of activating ligands, testosterone and DHT, and by inhibiting their synthesis; a process known as androgen deprivation therapy (ADT). This has been well established since Charles Huggins first reported that depleting androgens by orchiectomy caused a regression in metastatic PCa (Huggins and Hodges, 1972). It is common for anti-androgens to be used in combination with ADT to achieve complete androgen blockade. Anti-androgens inhibit

androgen signalling through directly inhibiting the AR. Despite an initial positive response to ADT, this is not curative, and many patients relapse and progress to CRPC within approximately 1-2 years, which is largely refractory to current treatments. Due to the vast clinical heterogeneity in PCa, response to treatment and survival can vary greatly (Shoag and Barbieri, 2016). A representation of PCa progression with the corresponding treatment options is summarised in Figure 1.5.



Created in [BioRender.com](https://BioRender.com)

**Figure 1.5** The progression of prostate cancer with timing of different available treatment options for patients at each stage of the disease.

Prostate cancer has a good prognosis if treated early when the cancer is localised to the prostate. Multiple criteria are assessed upon diagnosis, including Gleason score and serum PSA levels that can help determine the treatment pathway for the patient. For confined disease, surgery and radiotherapy are effective treatment strategies. Active surveillance of PSA levels determines when a patient goes on to androgen deprivation therapy. Patients respond well initially, however, over time, tumours progress and patients may be offered second-generation anti-androgens such as enzalutamide. Tumours invariably become castration resistant and metastatic to which chemotherapy can be used as there are currently no effective curative treatments.

Which treatment strategy the patient undergoes depends on their PCa risk group which is determined by analysing PSA levels, Gleason score and tumour stage. A Gleason score, originating in the 1970s by Dr Gleason (Gleason, 1988), is a combined score from 2-10 determined from biopsies from two different locations in the prostate. The score increases to

reflect how undifferentiated the tumour is and how likely it is to progress and metastasise. Low risk patients have a PSA level of less than 10 ng/mL, a Gleason score of less than 6 and a tumour stage of T1-2a meaning it is still confined to the prostate. Medium risk patients have a PSA level of 10-20 ng/mL, Gleason score 7 and/or T2b/c tumour stage, meaning the tumour has grown but has not metastasised. High risk patients have a PSA level of 20 ng/mL, Gleason score of >8 and/or stage T3a-4, meaning the tumour has metastasised (Partin *et al.*, 1997, D'Amico *et al.*, 1998).

For low-risk PCa patients, active surveillance can be sufficient, avoiding unnecessary harmful treatment and surgery, with a study reporting 78.6% overall survival after a 6.8 year mean follow up (Klotz *et al.*, 2010). For men with early-stage PCa, where the tumour is confined to the prostate, patients may be offered transurethral resection of the prostate (TURP) or radical prostatectomy and/or radiotherapy. For the minority of men that are diagnosed with advanced high-risk, aggressive disease, they will require multiple treatments such as surgery, radiotherapy in combination with hormonal- and chemo- therapy (Chang *et al.*, 2014). For patients with recurrent PCa, systemic or salvage treatments in an attempt to control the disease are available, however patients will typically progress to lethal, metastatic PCa (Antonarakis *et al.*, 2012).

#### 1.6.1 Targeting the AR signalling pathway

##### Targeting androgen synthesis

ADT is a first-line treatment for men with advanced PCa and aims to significantly reduce testosterone levels. Current ADT includes surgical castration or the use of luteinising hormone-releasing hormone (LHRH) agonists or antagonists causing medical castration, used alone or in combination with anti-androgens. LHRH agonists initially stimulate the release of LH, causing a large increase in testosterone levels. However, through a negative feedback loop, this leads to the downregulation of LH secretion, via a decrease in GnRH, ultimately lowering circulating testosterone to castrate levels. Examples of LHRH agonists include leuprolide, bruserelin and goserelin (Walker *et al.*, 1983, Leuprolide Study, 1984). Importantly, LHRH antagonists, such as Degarelix, have been developed more recently that lower testosterone levels without the initial increase in testosterone as observed with LHRH agonists

and are also faster acting making them the preferred choice for current ADT (Klotz *et al.*, 2008).

Abiraterone acetate is a cytochrome P450 enzyme 17 $\alpha$ -hydroxylase-17,20-lyase (CYP17) inhibitor which inhibits androgen synthesis in the adrenal glands, testes and intra-tumourally (Agarwal *et al.*, 2010). Even in castrate conditions, the production of androgens occurs via 17 $\alpha$ -hydroxylation of pregnenolone and progesterone to form the precursors hydroxypregnenolone and hydroxyprogesterone which are then converted to the testosterone precursors dehydroepiandrosterone and androstenedione by CYP17 (Rehman and Rosenberg, 2012). In the COU-AA-003 trial, in CRPC patients previously treated with docetaxel, it was reported that abiraterone acetate resulted in significant anti-tumour activity with a >50% PSA decline in 51% of patients (Reid *et al.*, 2010).

Unfortunately, the initial positive effects of ADT are typically short-lived as the cancer becomes resistant to therapy and is able to proliferate in the absence or in low-levels of androgens. Men usually become resistant to ADT after 2-3 years and progress to CRPC which ultimately causes death after just 16-18 months (Karantanos *et al.*, 2013).

#### Direct AR targeting

Rather than blocking androgen production to in-turn block the AR signalling pathway, there are drugs that directly target the AR to inhibit its transcription of target genes. First-generation AR targeting compounds include bicalutamide and flutamide (Schellhammer *et al.*, 1995). However, responsiveness to these therapies is short-lived and resistance develops. For example, in response to bicalutamide, a W741L or W741C mutation in the LBD of the AR gene causes bicalutamide to act as an agonist and therefore drives AR signalling (Hara *et al.*, 2003). As the mechanisms underlying PCa progression and therapy resistance have become better understood, second generation anti-androgen agents have been developed to target advanced CRPC. Examples of these therapies include enzalutamide and the more recently developed apalutamide and darolutamide.

Enzalutamide specifically binds to the AR LBD and has a higher AR binding affinity than first generation anti-androgens, such as bicalutamide (Tran *et al.*, 2009). Enzalutamide inhibits AR signalling by reducing its translocation into the nucleus and prevents AR binding to AREs and activating transcription (Tran *et al.*, 2009). It was tested in multiple CRPC cell line models, such

as LNCaP, in which the AR was overexpressed to mimic AR gene amplification, that occurs in approximately 50% of CRPC patients (Chen *et al.*, 2008a, Lapointe *et al.*, 2004). Enzalutamide decreased the expression of AR target genes such as PSA in response to synthetic androgen *in vitro* and suppressed proliferation of LNCaP xenografts (Tran *et al.*, 2009). There have been multiple clinical trials using enzalutamide which have demonstrated a significant improvement in overall survival, metastasis-free survival and PSA progression. In one of the first key clinical trials, AFFIRM, enzalutamide improved the median overall survival by 4.8 months in post-docetaxel treated mCRPC patients (Scher *et al.*, 2012). Later trials in men with non-metastatic CRPC, enzalutamide significantly increased the median metastasis-free survival to 36.6 months versus 14.7 months in patients in the control arm of the trial (Hussain *et al.*, 2018).

Apalutamide and darolutamide are the most recently developed second-generation antiandrogens that also bind to the LBD of the AR preventing AR translocation and DNA binding (Clegg *et al.*, 2012). Critically, unlike first-generation anti-androgens, apalutamide does not undergo an antagonist to agonist switch in the occurrence of AR overexpression. Clinical trials revealed apalutamide significantly improved progression-free survival in metastatic castration-sensitive PCa patients and metastasis-free survival in non-metastatic CRPC in combination with ADT (Chi *et al.*, 2019, Smith *et al.*, 2018).

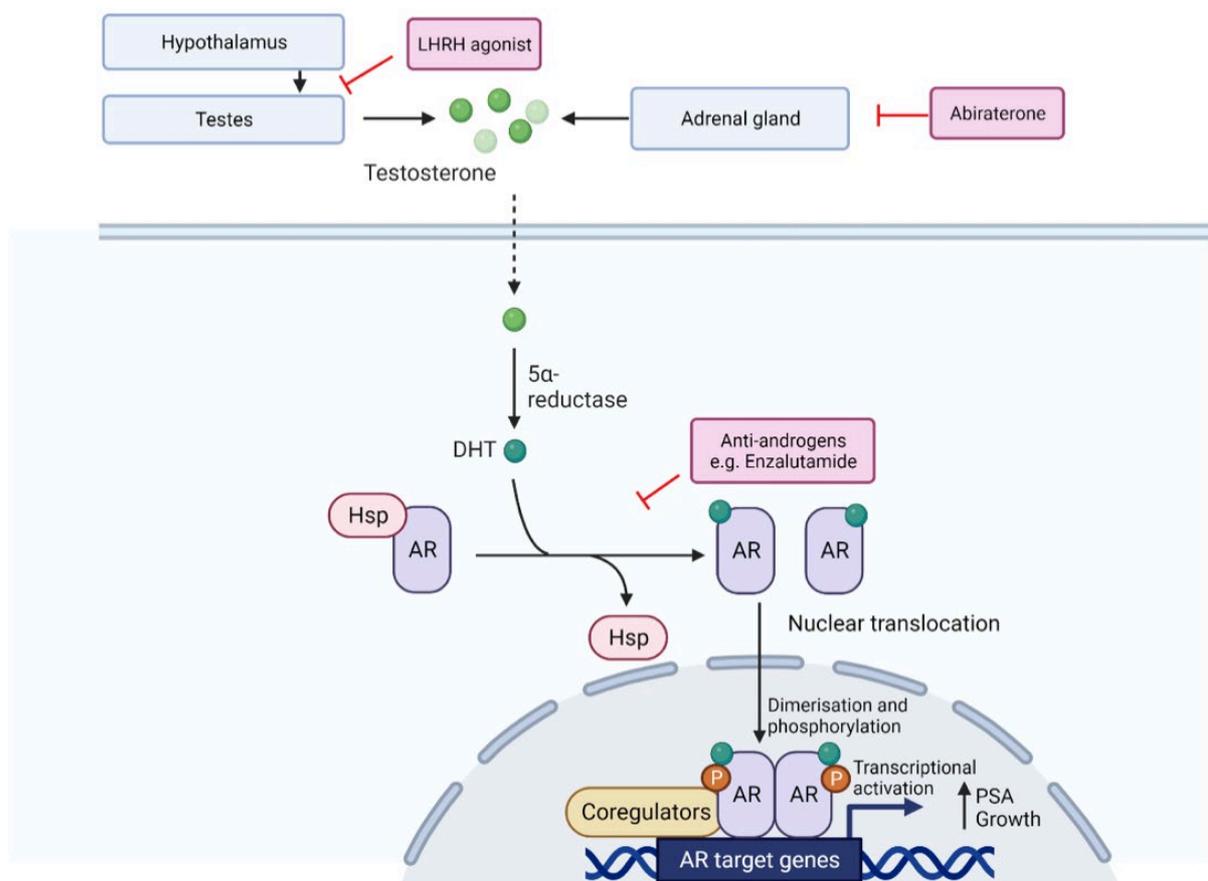
Darolutamide is another AR antagonist with increased potency over enzalutamide and apalutamide and importantly remains antagonistic over common mutants of the AR that arise in response to other anti-androgens (Moilanen *et al.*, 2015). It was approved for use in non-metastatic CRPC patients in 2019 after a successful clinical trial with a 21.9-month improvement in metastasis-free survival (Fizazi *et al.*, 2019). In 2022 results of a clinical trial in metastatic castration-sensitive PCa patients showed the risk of death was 35% lower in the darolutamide arm compared to placebo in combination with ADT and docetaxel (Smith *et al.*, 2022).

Unfortunately, despite the initial positive response to second-generation anti-androgens, like ADT, the majority of men develop resistance to these therapies highlighting the importance for development of novel and more effective therapies. For example, enzalutamide resistance

mechanisms include the F877L mutation in the AR gene and GR upregulation (Balbas *et al.*, 2013, Arora *et al.*, 2013).

### 1.6.2 Chemotherapy

Chemotherapy is used in the later stages of PCa, usually when patients have progressed after the previously discussed AR-targeted therapies. Chemotherapeutic drugs such as docetaxel, a microtubule inhibitor, have been approved for use in mCRPC patients, as it improves symptoms and quality of life, although this leads to just a 2.9 month prolongation in OS (Berthold *et al.*, 2008). Cabazitaxel, a second generation taxane which also inhibits microtubule function leading to cell death, is another chemotherapeutic option for mCRPC patients and is approved for use after docetaxel resistance (de Bono *et al.*, 2010).



**Figure 1.6 Androgen receptor signalling pathway targeting treatments in prostate cancer**

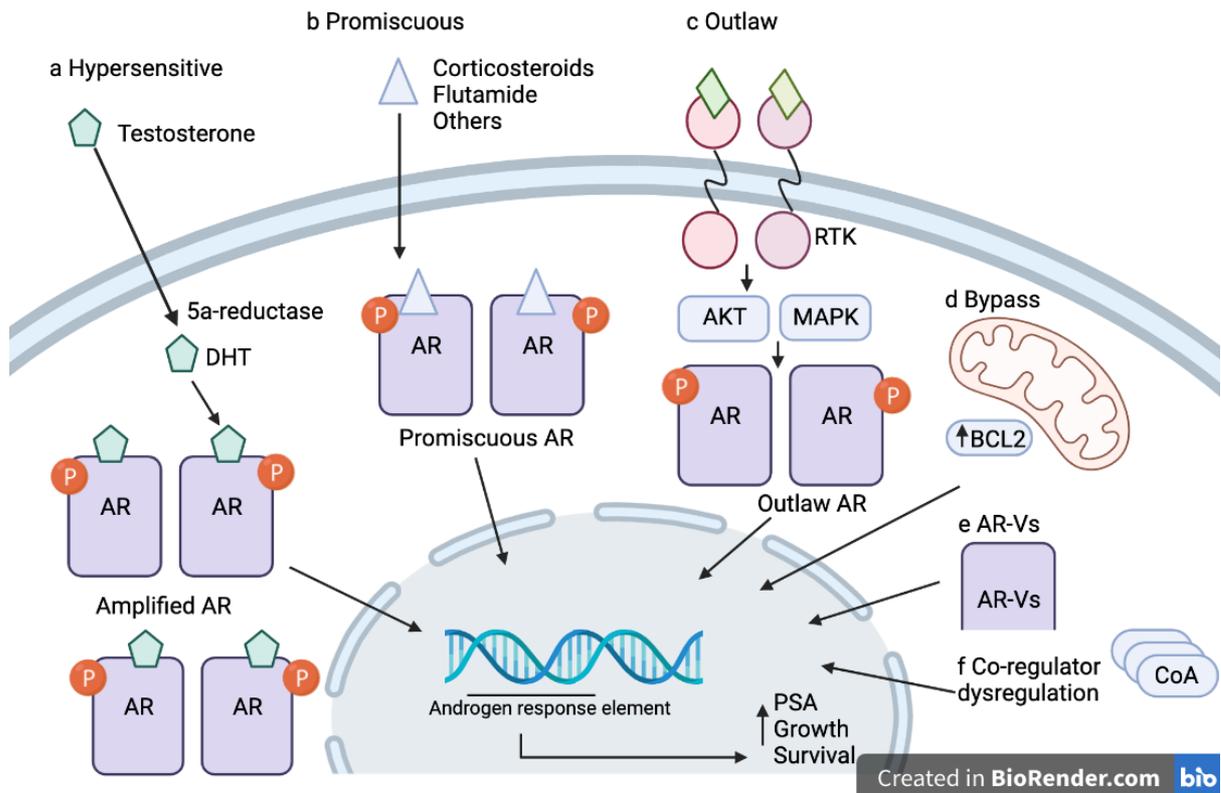
Targeting the AR signalling pathway is the main way to treat prostate cancer because of its critical role in PCa growth and progression. Abiraterone and LHRH agonists aim to inhibit testosterone synthesis from either the adrenal glands or the testes. Anti-androgens directly target the AR LBD to inhibit its activity and impact the transcription of target genes.

## 1.7 Mechanisms of castration resistance

Although androgen ablation is initially effective, many men progress to CRPC and die as a result of a lack of durable and effective therapy for this advanced disease stage (Kirby *et al.*, 2011). The pathways by which CRPC develops have been partly elucidated, providing an insight into how PCa cells become resistant to current therapies. A full understanding of these pathways will lead to the development of novel therapies for CRPC preventing the death of many. Second generation therapies have some effect in prolonging the overall survival of patients with advanced disease but are in no means curative (Orme *et al.*, 2022).

Genetic instability within PCa tumour cells is likely an important enabler of acquired resistance, which allows selection of cell populations that adapt to therapeutic challenge and hence increase the likelihood of survival and further DNA mutations (Feldman and Feldman, 2001). There are conflicting studies describing the contribution of genetic instability to the evolution of tumour cells in PCa; some suggest the tumours harbour mutations *de novo* to survive ADT prior to initial treatment (Cher *et al.*, 1996), whilst others suggest genetic aberrations are dependent upon ADT which provides a selective pressure to push the selection of resistant cells to repopulate the tumour (Taplin *et al.*, 1995, Taplin *et al.*, 1999).

AR hypersensitivity to low level androgens as a consequence of *AR* gene amplification, promiscuous ligand binding driven by AR LBD mutations, or activation of the AR independently of ligands have been proposed to be common mechanisms of resistance to ADT in PCa. These pathways are not mutually exclusive and can arise together to facilitate AR signalling in the absence of androgens during ADT (Robinson *et al.*, 2015).



**Figure 1.7 Mechanisms of resistance in prostate cancer.**

PCa cells have several mechanisms of resistance to several therapies. (a) Amplification of the *AR* gene is common, and this enables AR activation in castrate levels of androgen. (b) Mutations in the *AR* gene enables the AR to become activated to other ligands and some AR targeting drugs. (c) The outlaw pathway describes the process by which the AR becomes activated by other signalling pathways such as through receptor tyrosine kinase (RTK) activation. (d) Bypass pathways enable activation of AR target genes independently of AR signalling. (e) The generation of AR splice variants (AR-Vs), which are constitutively active truncated forms of the AR, which do not require ligands to drive transcription of canonical target genes. (f) Amplification of AR co-activator or downregulation of AR co-repressor regulatory proteins can cause disease progression through increased AR transcriptional activity. Figure adapted from (Feldman and Feldman, 2001).

### 1.7.1 Hypersensitivity

Hypersensitivity arises when levels of circulating androgens are reduced as a consequence of ADT. In this environment, cells develop a hypersensitivity to low testosterone levels through *AR* gene amplification which increases intracellular abundance of AR protein, enabling activation of the pathway despite the low levels of androgen. More than 50% of CRPC patients have *AR* amplification; the majority of which have previously undergone hormone therapy and whose tumours have metastasised (Taylor *et al.*, 2010). Higher levels of DHT due to a substitution mutation (V89L) in the gene encoding 5α-reductase, resulting in increased

catalytic activity of the enzyme, also makes the tumour cells hypersensitive to low androgen levels. This is due to more testosterone being converted to DHT and therefore, increased activation of the AR. This is more common in African-American ethnic groups, putting them more at risk of early-onset and more aggressive PCa (Makridakis *et al.*, 2000). Conversely, an alternative substitution (A49T) in the gene encoding 5 $\alpha$ -reductase results in reduced activity of the enzyme, common among Asians, possibly reducing the risk of PCa (Makridakis, 1997).

These tumours are not androgen independent as they still depend on androgens to form AR-ligand complexes to drive survival and proliferation, but this occurs just at lower levels of androgens.

#### 1.7.2 Promiscuous ligand binding

Promiscuous ligand binding occurs in tumours that have acquired gain of function mutations at specific positions in the *AR* gene encoding the LBD. These mutations enhance the promiscuity of the AR to allow non-androgen steroid hormones and direct AR antagonists to bind and activate the AR pathway (Feldman and Feldman, 2001). This leads to PCa cells being able to proliferate when there are low levels of androgens caused by ADT. Point mutations such as the T878A missense mutation in the LBD of the *AR* gene has been identified in approximately 25% of tumour samples (Gaddipati *et al.*, 1994). This mutant AR binds to various non-androgen hormones, including cortisone, and the AR antagonist flutamide, and uses them as agonists to drive AR signalling. There are other AR mutations that cause receptor activation in the presence of next generation anti-androgens, including the F877L mutation which confers agonism to enzalutamide (Balbas *et al.*, 2013). Identified using a mutagenesis screen in LNCaP cells, the presence of the F877L mutation causes enzalutamide to induce AR nuclear translocation and recruitment of AR to the enhancer regions of target genes (Balbas *et al.*, 2013). There have been compounds generated to target the AR F877L mutant to overcome enzalutamide resistance, such as JNJ-63576253 and darolutamide to treat CRPC patients (Zhang *et al.*, 2021, Moilanen *et al.*, 2015). Similar to the hypersensitivity pathway, it is thought ADT causes a selective pressure for the tumour cells that harbour promiscuous ligand binding-causing mutations to proliferate and result in ADT-resistant cell populations (Taplin *et al.*, 1999).

### 1.7.3 Outlaw pathway

Whilst the hypersensitive and promiscuous binding pathways require ligand binding for AR activation, the outlaw pathway does not. This is a ligand-independent pathway in which the AR is stimulated by activation of receptor tyrosine kinase pathways in response to direct phosphorylation by growth factors including insulin-like growth-factor 1 (IGF-1), keratinocyte growth factor (KGF) and epidermal growth factor (EGF) (Culig *et al.*, 1994). Furthermore, the inflammatory cytokine, interleukin 6 (IL-6) levels in patients were shown to be associated with anti-androgen resistance, with significantly lower IL-6 levels in ADT-sensitive patients compared to *de novo* ADT-resistant patients (Pal *et al.*, 2019). This is through its downstream activation of IL6/STAT3/AR signalling and inhibition of STAT3 with niclosamide was found to downregulate AR signalling. The same study found synergistic effects coupling enzalutamide treatment with niclosamide through enhanced AR protein degradation and therefore reduced AR driven transcription (Liu *et al.*, 2015).

### 1.7.4 The bypass pathway

Similarly to the outlaw pathway, the bypass pathway does not rely on ligand binding to the AR, as this involves bypassing the AR for activation and therefore proliferation and survival. Apoptosis is inhibited in the absence of androgens by the anti-apoptotic activity of Bcl-2. The Bcl-2 protein is known to be associated with the development of B-Cell lymphoma, and is frequently over-expressed in CRPC (Colombel *et al.*, 1993). Other tumour-suppressors and oncogenes could be involved in the bypass pathway and a greater understanding of these mechanisms could lead to more targeted therapies.

### Nuclear receptor cross talk

The GR, another member of the nuclear hormone receptor family, has been shown to drive transcription of several AR target genes such as *PSA*, and its expression is increased in response to anti-androgens (Arora *et al.*, 2013). The study by Arora *et al.*, demonstrated that the GR could bypass AR inhibition and therefore drive resistance to AR targeting therapies. Treatment with enzalutamide in LNCaP xenografts significantly upregulated expression of the GR gene, *NRC31*, as well as several GR target genes. The GR and AR have similar cistromes as demonstrated by chromatin immunoprecipitation-sequencing and have a significantly overlapping gene signature in enzalutamide resistant cells (Arora *et al.*, 2013), supporting the

hypothesis that GR can bypass AR and drive transcription of AR target genes. However, because glucocorticoids are given to PCa patients to mitigate side effects of anti-androgens such as abiraterone, this could provide a favourable environment for tumour cells, leading to PCa progression.

#### 1.7.5 Aberrant expression of androgen receptor co-regulators

As previously discussed, the AR has many co-activator and co-repressor proteins that act to regulate the transcriptional function of the AR (Lonergan and Tindall, 2011). Many AR co-regulators can become upregulated in CRPC and lead to aberrant activation of the AR. In benign prostate tissue, these regulators assist in the activation or suppression of the AR signalling axis, but in CRPC they could promote transcriptional activity and proliferation (Debes *et al.*, 2003). Alternatively, downregulation of co-repressor protein expression could have a similar effect on proliferation and survival (Li *et al.*, 2001). An integrative genome profiling of PCa study reported 20% overexpression of co-activators in primary PCa which increased to a 63% overexpression in CRPC (Taylor *et al.*, 2010). Important AR-coregulatory proteins that harbour genomic gain of function mutations in CRPC circulating tumour cells are FOXA1, NCOA2 and BRD4, whilst GATA2 was an example of a protein with common genomic loss in CRPC (Gupta *et al.*, 2017).

#### 1.8 AR splice variants

More recently, AR splice variants (AR-Vs) have been identified in PCa as an additional resistance mechanism to ADT and other AR-targeting treatments. They are constitutively nuclear and active, C-terminally truncated forms of the AR that promote androgenic signalling in castrate conditions. Recent studies have indicated that they are expressed in response to ADT in over 65% of PCa patients and are associated with PCa progression (Watson *et al.*, 2010, De Laere *et al.*, 2017, Welti *et al.*, 2016). Therefore, elucidating the mechanisms of AR-V generation and regulation is crucial to develop intuitive methods to inhibit AR-V signalling and more effectively treat CRPC patients.

AR-Vs contain the NTD and DBD but lack the LBD of the AR and are therefore refractory to current treatments as ADT does not affect their activity, nor do second generation anti-

androgens such as abiraterone and enzalutamide. Tepper *et al.*, was the first to identify two AR species in the CWR22Rv1 cell line, derived from the CWR22 relapsed PCa xenograft (Sramkoski *et al.*, 1999, Tepper *et al.*, 2002). They identified a C-terminally truncated AR variant 75-80 kDa in size that was primarily located in the nucleus. The cell line showed a lesser response to anti-androgen treatment when compared to the LNCaP cell line. Since then, over 20 species of AR-Vs have been identified in cell lines and patient samples. They are androgen-independent meaning they can promote growth and survival in low androgen levels (Hu *et al.*, 2009).

### 1.8.1 Origin of AR-Vs

There have been several mechanisms suggested for the generation of AR-Vs, including AR genomic rearrangements, protein cleavage of FL-AR, and alternative mRNA splicing. They are thought to be produced when there are low levels of androgen present driving growth in the absence of ligand to provide a mechanism of resistance during treatment for PCa.

Before genomic rearrangements were discovered, Ceraline *et al.*, identified a nonsense mutation creating a stop codon in exon 4 generating an AR-V (Ceraline *et al.*, 2003). Later, proteolytic cleavage of the AR gene in the hinge region by calpain 2 leading to generation of an 80 kDa AR-V lacking the LBD was identified by Libertini *et al.* Inhibition of calpains in CWR22Rv1 cells reduced expression of the AR-V, but this has not been validated in clinical PCa specimens (Libertini *et al.*, 2007, Ware *et al.*, 2014).

The sequences of the alternate transcripts were obtained from CWR22Rv1 cells using 3'-Rapid amplification of cDNA ends (RACE), revealing cryptic exons (CEs) that are integrated into AR-V mRNA. As shown in Figure 1.8, the majority of AR-V transcripts contain exons 1, 2 and 3 (with the exception of AR-V3) followed by a CE generating alternative 3' sequences (Dehm *et al.*, 2008). This finding opposed the earlier suggestion that AR-Vs arise by proteolytic cleavage by calpains. Shortly after, Hu *et al.* identified 7 AR-Vs (AR-V1-7), using *in silico* BLAST sequence analysis of the AR gene in CWR22Rv1 cells, which incorporated distinct CEs. They identified 6 in intron 1, 3 in intron 2 and 3 in intron 3; the latter termed CE1, CE2 and CE3 are all located downstream from exon 3. CE3 is incorporated into the most extensively studied AR-V, AR-V7 which is encoded by ARex1/2/3/CE3 (Figure 1.8). The group also reported that AR-V1 and AR-V7 were the most abundant variants in their hormone-resistant patient samples and reported

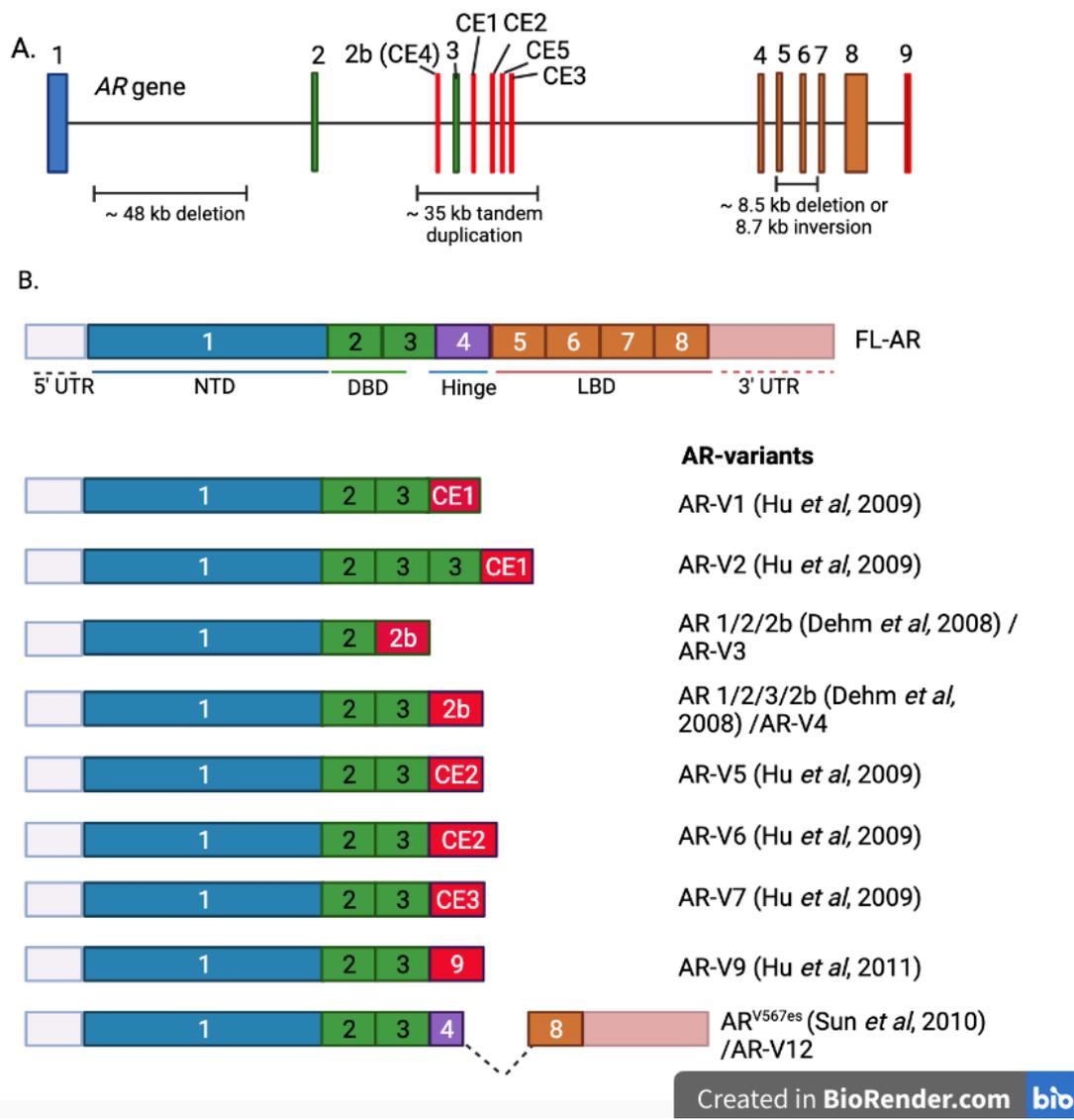
a 20-fold increase in expression compared to hormone sensitive specimens (Hu *et al.*, 2009). AR-V12 or AR<sup>V567es</sup> is generated through exon skipping in which exons 5, 6 and 7 are excluded. Importantly, this AR-V has also been identified in metastatic PCa patient tissue (Sun *et al.*, 2010).

Li *et al.*, then discovered intragenic rearrangements associated with the expression of AR-Vs in CWR22Rv1 cells. They reported a 35 kb AR segment that several previously described alternative AR exons were contained in (Li *et al.*, 2011). The same group also identified genomic rearrangements in human xenograft models including deletions and inversions located in exons 5, 6 and 7 (Li *et al.*, 2012). To validate these findings in the clinical setting, Henzler *et al.* performed DNA-sequencing on patient tissue from early and late stage CRPC and therapy-naïve PCa. This revealed a class of AR gene alterations, termed AR-genomic structural rearrangements, that were present in a third of CRPC patients (Henzler *et al.*, 2016) However, there was no difference in AR-V7 mRNA levels in patients who were positive or negative for the rearrangement events. This was likely due to sub-clonality and heterogeneity of tumour samples that would confound results. A more recent study, using whole-genome sequencing and structural rearrangement analysis of the AR locus in PCa patients, validated that AR gene rearrangements promote AR-V expression. These rearrangements were modelled in cell lines which caused high expression of AR-Vs and led to enzalutamide resistance (Li *et al.*, 2020).

Alternative splicing of FL-AR mRNA involves inclusion or exclusion of exons resulting in alternate AR-Vs, thought to be induced by ADT, is another, more prominent mechanism of AR-V synthesis (Liu *et al.*, 2014a).

Alternative splicing of pre-messenger RNA (pre-mRNA) is an essential biological process responsible for the production of multiple transcripts from single genes. This process involves recruitment of splicing factors to 5' and 3' splice sites at pre-mRNA transcripts causing exclusion of intronic regions and then re-ligation of exons. Alternative splicing is commonly dysregulated in cancer, and in PCa it has been shown to be responsible for the synthesis of multiple AR-V transcripts (Venables *et al.*, 2009). The *cis* regulatory RNA elements, termed exonic splicing enhancers or intronic splicing enhancers are the regions upon which the *trans* RNA splicing proteins assemble and decide which exons are spliced together (Liu *et al.*, 2014a). For example, the synthesis of AR-V7 arises from the use of the alternative 3' splice site that is

located downstream from exon 3 which causes incorporation of CE3 instead of exon 4 and therefore a constitutively active truncated protein (Hu *et al.*, 2009). The mechanisms behind this have been an important area of research as understanding this biological process will give rise to means of therapeutically targeting AR-V production. It is not fully understood what splicing factors and/or regulators are required for AR-V splicing. Splicing factors ASF/SF2 are reported to be recruited to AR CE3, with knockdown of these reducing expression of AR-Vs (Liu *et al.*, 2014a). Other splicing factors that have been reported to have a role in AR-Vs splicing include Sam68, SFPQ and JMJD1A. Sam68, an RNA binding protein, when overexpressed, was shown to increase the incorporation of CE3 in AR-V7 mRNA in a minigene reporter system as well as endogenous levels in the CWR22Rv1 cell line. Sam68 that was also shown to interact with AR-V7 protein in the nucleus and drive AR-V7 transcriptional activity and expression of the AR-V7 regulated gene *UBE2C* (Stockley *et al.*, 2015). JMJD1, a histone demethylase, was shown to regulate AR-V7 splicing through heterogeneous nuclear ribonucleoprotein F (HNRNPF). Both proteins were associated with AR-V7 expression and immunoprecipitated at AR-V7 splice sites of mRNA (Fan *et al.*, 2018). Finally, SFPQ, an RNA binding protein was shown to control expression of several spliceosome genes, that together, co-ordinate to form a splicing complex that is involved in AR alternative splicing (Takayama *et al.*, 2017). Together, there have been some advancements in our understanding around AR alternative splicing but there is still a lot that is unknown about the mechanism. Given that AR and AR-Vs drive PCa progression, identification of other splicing factors involved in AR-V generation could reveal tractable therapy targets.



**Figure 1.8** The AR gene and sequence of exons and cryptic exons in several androgen receptor variants arising from alternative splicing.

A. A diagrammatic representation of the AR gene with exons represented by black vertical bars and cryptic exons represented by red vertical bars. B. Representation of AR-V transcripts and their alternative names. Image adapted from (Ware *et al.*, 2014).

### 1.8.2 Regulation of AR-Vs

AR-Vs have been identified in normal and cancerous prostate tissue, suggesting a role in normal functions of AR signalling. However, many reports show an increased expression of AR-Vs in CRPC, in comparison to early-grade androgen dependent PCa, suggesting ADT creates a selective pressure for the upregulation of AR-Vs (Guo *et al.*, 2009).

Like FL-AR, AR-Vs are regulated by a number of co-activators and co-repressors. Work conducted in my host laboratory have so far identified FOXA1 (Jones *et al.*, 2015), Aurora kinase A (Jones *et al.*, 2017), GATA2 (Chaytor *et al.*, 2019) and PARP1/2 (Kounatidou *et al.*, 2019) as important co-regulators of AR-V function.

Transcriptional regulators of AR function, such as p300/CBP, have also been identified to be a co-activator of AR-V function (Debes *et al.*, 2003, Welti *et al.*, 2021). Depletion of p300 and CBP reduced expression of several AR target genes in AR-V-driven cell lines, including the AR-V positive LNCaP cell derivative LNCaP95 and CWR22Rv1, while re-activation of p300 and CBP lead to restoration of AR signalling (Welti *et al.*, 2021). A selective inhibitor of the p300/CBP bromodomain (CCS1477) was used to validate this using transcriptomic analysis of CCS1477-treated CWR22Rv1 cells which demonstrated down-regulation of the androgen response hallmark. Importantly, AR-V protein levels were decreased upon treatment with the inhibitor in several AR-V expressing cell lines as well as in a CWR22Rv1 mouse xenograft model. Transcriptomic analysis of the xenograft tumours after 28-day treatment showed significant de-enrichment of AR and AR-V driven gene signature (Welti *et al.*, 2021).

Recently, a genome-scale CRISPR/Cas9 screen in CWR22Rv1 cells, using AR/AR-V-GFP expression as a readout, identified and validated PRMT1 as a regulator of AR and AR-V7 expression in CWR22Rv1 and VCaP cells. Differential alternative splicing analysis showed PRMT1 depletion causes splicing changes in 728 genes and inhibition of PRMT1 synergises with enzalutamide to suppress proliferation in AR-V positive cell lines (Tang *et al.*, 2022).

Furthermore, gene expression analysis from the cell lines LNCaP and the AR-V positive LNCaP95 derivative was performed to determine spliceosome-related genes that are upregulated and therefore potentially correlated with AR-V7 generation or regulation. This revealed a gene-list that was cross-referenced with the same spliceosome-related genes that are down-regulated in LNCaP95 in response to BET inhibition, as well as genes from an siRNA screen that are associated with a more than 50% reduction in AR-V protein levels. This revealed the 2OG-dependent dioxygenase, JMJD6, as the only gene that was in all three of the datasets. JMJD6 expression correlated with a worse overall survival and depletion caused a reduction in AR-V7 protein levels and significantly impacted PCa cell growth in LNCaP, LNCaP95, VCaP and CWR22Rv1 cell lines. JMJD6 regulation of AR-V expression was shown to

be regulated through recruitment of the splicing factor U2AF65, as knockdown of JMJD6 caused a significant reduction in U2AF65 recruitment to three separate regions of AR-V7 transcripts (Paschalis *et al.*, 2021).

### 1.8.3 Biological and Clinical relevance

When AR-Vs were first being fully characterised, it was important to determine if they could exert similar transcriptional effects as FL-AR. To do this, AR-Vs were depleted in steroid-free conditions in CWR22Rv1 cells which led to a significant reduction in ligand-independent AR target gene expression. This was not observed upon depletion of FL-AR using an exon 7 targeting siRNA, nor did the addition of androgens cause upregulation of the AR target gene *PSA*, further validating AR/AR-V transcriptional activity in this model was ligand-independent (Dehm *et al.*, 2008). The same study also demonstrated that AR-Vs are able to promote androgen-independent proliferation in CWR22Rv1 cells, as depletion of AR-Vs in steroid-free conditions inhibited growth but selective depletion of only FL-AR did not; demonstrating AR-Vs can substitute for FL-AR in the absence of androgens (Dehm *et al.*, 2008). Furthermore, it has been shown that AR-Vs drive resistance to anti-androgens in CWR22Rv1 cells as depletion of AR-Vs, but not FL-AR, restored androgen and anti-androgen responsiveness (Li *et al.*, 2013). Although AR-Vs have a considerable overlap of transcriptional targets to FL-AR, it has been reported that they have distinct transcriptomes. Recent studies, using LNCaP and VCaP cell line models with doxycycline inducible AR-V7 expression, showed in LNCaP-AR-V7 cells, 54% and 57% of AR and AR-V regulated genes were unique to each isoform, respectively (Basil *et al.*, 2022). This finding that AR and AR-V have individual transcriptomes was also shown in the VCaP-AR-V7 cell line with 30% of the AR-V transcriptome being distinct from the FL-AR (Basil *et al.*, 2022). This was also shown in a study by Cato *et al.*, that utilised the LNCaP95 cell line, that FL-AR and AR-V7 have distinct transcriptional roles as there were significant differences between their transcriptomes (Cato *et al.*, 2019). AR-V expression has been linked to upregulation of cell cycle genes such as *UBE2C*, which, importantly was not dependent on FL-AR and was correlated in clinical samples with AR-V7 levels (Hu *et al.*, 2012). A recent study reported that AR-Vs negatively regulate various genes involved in tumour-suppression and limiting proliferation, suggesting a mechanism of CRPC growth and progression (Cato *et al.*, 2019). Moreover, Kounatidou *et al.*, demonstrated a pro-proliferative role of AR-Vs after investigation into genes that were downregulated in response to AR-V depletion (Kounatidou

*et al.*, 2019). Genes involved in cell cycle regulation and mitosis were also part of the AR-V driven gene signature (Schiewer *et al.*, 2012).

There have been 12 AR-V mRNA species identified in primary PCa and up to 23 in metastatic CRPC. The best characterised species that has been the most extensively researched so far is AR-V7 which arises from splicing of AR exons 1/2/3/CE3 and AR<sup>v567es</sup>, arising from skipping of exons 5-7 and have been detected in clinical samples. AR<sup>v567es</sup>, a constitutively active AR-V, was shown to contribute to cancer progression in mice PCa xenograft models following castration by driving AR target gene expression, increasing the expression of FL-AR and enhancing FL-AR activity. The expression of AR<sup>v567es</sup> in the xenografts correlated with castration resistance. Importantly, AR<sup>v567es</sup> was commonly found in patient tissue of PCa metastases (Sun *et al.*, 2010). AR-V7 expression is also clinically relevant and is correlated with drug resistance; a study investigating this, using AR-V7 expression in circulating tumour cells (CTCs), reported that 75% of patients who tested positive for AR-V7 developed drug resistance (Wang *et al.*, 2018).

As AR-V7 can be detected in clinical samples, and is thought to predict resistance, it has potential to be used as a biomarker to determine what treatment a patient receives and to predict therapy outcomes. According to a meta-analysis, AR-V7 negative patients are more likely to have a better response to ADT in comparison to AR-V7 positive patients (Li *et al.*, 2018). Another paper reported that AR-V7 positive (defined by immunofluorescence detection in circulating tumour cells and nuclear localisation of AR-V7) patients had a worse OS rate when treated with ADT, but had a better OS when treated with chemotherapy (Scher *et al.*, 2018). A study analysing mRNA levels of FL-AR, AR-V7, AR<sup>v567es</sup> and AR-V3 in matched hormone-sensitive PCa and CRPC patients determined AR-V levels were significantly increased in 81.2% CRPC patients, but the authors note that sampling bias cannot be ruled out of the results and due to a small sample size of 19 patients, the clinical implications cannot be determined (Park *et al.*, 2019). A much larger study was conducted using an immunohistochemistry readout to measure AR-V7 expression in 358 primary prostate tissues and 293 metastatic biopsies. This revealed that over 75% of patients were AR-V7 positive following ADT which increased further in patients who received the antiandrogens abiraterone or enzalutamide (Sharp *et al.*, 2019). Furthermore, patients who were negative for AR-V7 expression had significantly better overall survival (Sharp *et al.*, 2019). These results

support a potential prognostic benefit of using AR-V7 as use as a biomarker to guide personalised treatment in advanced PCa.

#### 1.8.4 Therapeutically targeting AR-Vs

Because of the prevalence of AR-V expression in late stage PCa patients and the correlation with diminished treatment response and overall survival, there is a need to develop compounds to target them and abrogate AR signalling.

##### NTD targeting

Currently, all hormone therapies for PCa target the LBD of the AR, which is not present in AR-Vs. Therefore, targeting the N-terminal domain of the AR is logical, due to this region being crucial in transcriptional activity of the AR. However, this has proved very difficult due to the largely unstructured and flexible nature of this region, making drug design problematic (Antonarakis *et al.*, 2016). Despite this, there have been compounds that have made it to clinical trials, including the EPI family of compounds which bind to the Tau-5 region of the NTD and inhibit AR and AR-V transcriptional activity (Myung *et al.*, 2013). However, due to minor PSA changes in patients, the drug's low potency and a short half-life, the first generation of these compounds were unsuccessful with termination of Phase I clinical trials because of excessive pill burden (Trials.gov, 2018).

##### DBD targeting

Targeting the DBD is another strategy to either block AR-DNA interactions or disrupt its homodimerization and inhibit AR/AR-V activity (Dalal *et al.*, 2014, Radaeva *et al.*, 2021). There has been limited progress in the development of DBD targeting compounds, but one such compound is VPC-14449, which pre-clinically, has been shown to have some transcriptional impact on AR-V7, but more so on FL-AR (Dalal *et al.*, 2014). More recently, results have been published using refined derivative compounds VPC-17160 and VPC-17281 which have demonstrated improved inhibition of AR-V7 activity and microsomal stability. These compounds inhibit the dimerization of the AR DBD and were shown to inhibit the proliferation of LNCaP and CWR22Rv1 PCa cell lines. However, one of the compounds also showed activity in the AR-negative PC3 cell line, indicative of off-target activity (Radaeva *et al.*, 2021).

Although this compound series still requires optimisation, it provides strong proof of concept to further develop compounds that selectively target the AR DBD as a means to diminish FL-AR and AR-V activity going forward.

#### Bromodomain and extraterminal (BET) inhibitors

Inhibition of AR-V chromatin binding using bromodomain and extraterminal (BET) inhibitors, as exemplified by pre-clinical studies using JQ1, is another potential treatment strategy for AR-V positive tumours. BET proteins are epigenetic readers that bind to acetylated histones that have been shown to interact with the NTD of the AR and as previously mentioned, blockade with JQ1 has been shown to disrupt this interaction (Asangani *et al.*, 2014). JQ1 has been shown to diminish AR and AR-V transcriptional activity by reducing AR and AR-V protein levels which in turn decreased AR occupancy at AREs on chromatin (Chan *et al.*, 2015). This was shown to reduce growth in cell lines as well as xenograft tumours that express AR<sup>V567es</sup> (Chan *et al.*, 2015). BET inhibitors are currently being tested clinically in numerous clinical trials. For example, the pan BET inhibitor ZEN-3694, in combination with enzalutamide, showed potential efficacy in CRPC patients, with a subset of patients with an aggressive phenotype displaying prolonged progression-free survival (Aggarwal *et al.*, 2020). However, AR-V positivity data was not captured in this trial, so it was not evaluated if this could influence how patients respond to BET inhibition.

#### HSP90 inhibitors

Indirectly targeting AR-V activity through inhibition of the chaperone protein HSP90 has shown some potential in PCa cell line models using the next generation HSP90 inhibitor, onalespib. However, there are some conflicting reports on the sensitivity of AR-Vs to HSP90 inhibition. One investigation using three HSP90 inhibitors showed AR-V7 and AR<sup>V567es</sup> were less sensitive to HSP90 inhibition compared to FL-AR when analysing their transcriptional activity using a luciferase transactivation assay (Gillis *et al.*, 2013). Furthermore, they showed AR-Vs do not require HSP90 for nuclear translocation and HSP90 inhibition induced the expression of AR-V7 transcripts and protein in VCaP cells. However, HSP90 inhibition significantly reduced the growth of the AR-V expressing cell lines VCaP and CWR22Rv1 meaning AR-Vs do not confer resistance (Gillis *et al.*, 2013). Similar findings were reported using another HSP90 inhibitor, ganetespib. Cell death was induced in AR-V positive VCaP and

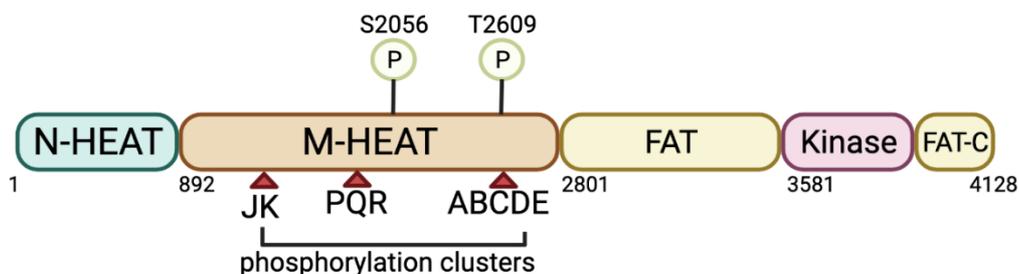
CWR22Rv1 cells, but AR-V transcriptional activity or expression were not altered (He *et al.*, 2013). Conversely, using a second generation HSP90 inhibitor, onalsepib, AR-V protein levels were diminished upon HSP90 inhibition in AR-V positive PCa cell lines CWR22Rv1, VCaP and LNCaP95 (Ferraldeschi *et al.*, 2016). Unlike the expected mechanism of interrupting the chaperone-receptor interaction, it was shown that AR-Vs do not interact with HSP90, and HSP90 inhibition disrupts AR-V splicing and therefore AR-V protein levels. This was validated by determining onalsepib did not decrease pre-mRNA levels of AR-V7 using intronic-spanning primers but did reduce levels of AR-V7 mRNA (Ferraldeschi *et al.*, 2016). This splicing impairment could be due to the activation of the heat shock response which reportedly represses splicing (Shin *et al.*, 2004). RNA-sequencing followed by global splicing analysis did in fact show onalsepib caused splicing changes in 557 genes, including AR-V7 CE3 inclusion (Ferraldeschi *et al.*, 2016). However, because the splicing alterations go beyond AR-V splicing this could be problematic when evaluating the clinical use of these compounds and warrants further investigation.

Other possible strategies to target AR-Vs include protein degradation using PROTACs, synthesis inhibition and the targeting of AR co-activators such as CDK9 (Richters *et al.*, 2021). A greater understanding of the function and regulation of AR-Vs will assist in the development of more effective and selective therapies.

### 1.9 DNA-PKcs

The DNA dependent protein kinase (DNA-PK) is a serine/threonine protein kinase complex that consists of a Ku heterodimer (Ku70/ Ku80) and a catalytic subunit (DNA-PKcs). It is essential in the DDR pathway, particularly in the non-homologous end joining (NHEJ) pathway of double strand break (DSB) repair (Chan *et al.*, 2002). Because of this, it is an attractive cancer target to combine with DNA damaging agents such as radiotherapy. However, it has been well established as a multifunctional protein kinase with involvement in a number of DNA damage independent cellular processes including cell cycle (Douglas *et al.*, 2014, Lee *et al.*, 2011), transcription (Jackson *et al.*, 1990) and telomere maintenance (Espejel *et al.*, 2002, Ruis *et al.*, 2008).

DNA-PKcs is the largest subunit of the DNA-PK complex and is a 469 kDa protein encoded by the *PRKDC* gene and belongs to the family of phosphatidylinositol 3-kinase-related kinases (PI3KKs), which also includes ataxia-telangiectasia mutated (ATM), ATM and RAD3-related (ATR), mammalian target of rapamycin (mTOR), suppressor of morphogenesis in genitalia (SMG1) and transformation/transcription associated protein (TRAP) (Menolfi and Zha, 2020). DNA-PKcs is made up of distinct functional domains (as shown in Figure 1.9); the N-terminal domain contains HEAT (Huntingtin, Elongation Factor 3, PP2A and TOR1) repeats and JK, PQR and ABCDE regulatory phosphorylation clusters. The C-terminus contains a highly conserved PIKK catalytic domain that is surrounded by FAT (nomenclature derived from homologous regions in FRAP, ATM, and transcription domain-associated protein TRRAP), and FATC (FAT at the C-terminus) domains (Rivera-Calzada *et al.*, 2005). Between the FAT and kinase domain sits a FKBP12-rapamycin binding domain and between the kinase and FATC domains there is a PIKK-regulatory domain. DNA-PK has several autophosphorylation sites, but sites in the ABCDE cluster, such as T2609, and sites in the PQR cluster, such as S2056, have been shown to be important in DNA repair. Autophosphorylation at these sites causes a conformational change in DNA-PKcs allowing other NHEJ factors to access and process the DNA to ultimately complete damage repair (Liu *et al.*, 2022, Ding *et al.*, 2003).



**Figure 1.9 Structure of DNA-PKcs.**

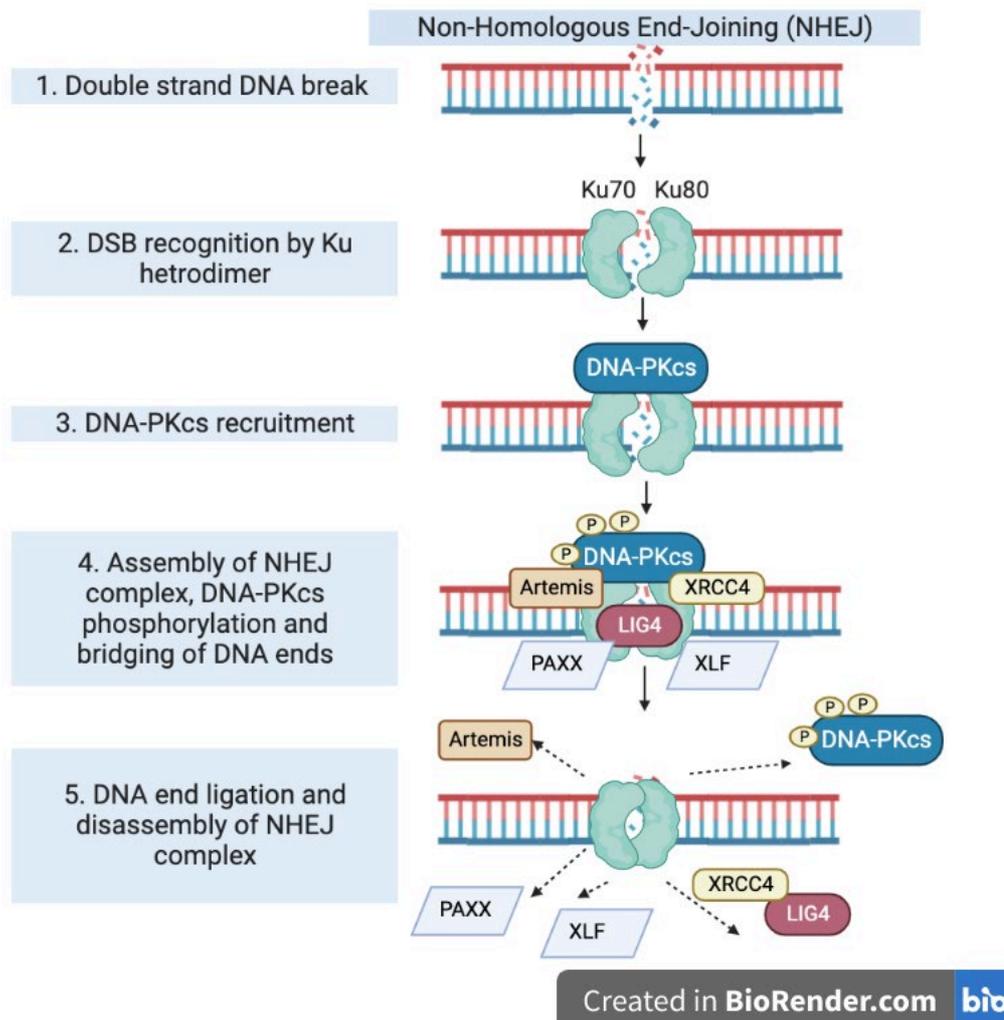
DNA-PKcs is comprised of 4128 amino acids which make up an N-terminal heat domain that harbour HEAT repeat regions and the JK, PQR and ABCDE phosphorylation clusters. It also includes FAT, Kinase and FAT-C domains. The C-terminus is responsible for Ku70/Ku80 heterodimer binding.

### 1.9.1 DNA-PKcs in DNA repair

DSBs can form as a result of endogenous (e.g. reactive oxygen species) and exogenous (ionising radiation) insults and are the most cytotoxic DNA lesions; causing chromosomal

rearrangements or cell death if not properly repaired. The two main pathways of DSB DNA repair are NHEJ and homologous recombination (HR). NHEJ, shown in Figure 1.10, is the primary mechanism of DSB repair as it can occur throughout the cell cycle so can rapidly respond to DNA damage. HR is extremely accurate as it requires a DNA template homologous to the region being repaired and is most active during S/G2 phases of the cell cycle.

Upon DSB induction, the Ku70/Ku80 heterodimer, the regulatory subunit of the DNA-PK heterotrimeric complex, is responsible for recognition of the damaged DNA (Hartley *et al.*, 1995). This recognition and binding can occur within seconds of the DNA break happening due to its high affinity for DNA ends and is responsible for the subsequent recruitment of DNA-PKcs. DNA-PKcs binds to the Ku complex, via its C terminal region, which activates the kinase activity of DNA-PKcs. As a consequence of phosphorylation-dependent recruitment of other core NHEJ factors including Artemis, XRCC4, DNA ligase IV and XLF, which form the NHEJ complex, ligation of the DNA DSB and resolution of DNA damage is achieved (Calsou *et al.*, 1999) (Davis *et al.*, 2014). Once DNA ends have been ligated, the components of the NHEJ complex disassemble from the DNA and the DSB is successfully repaired. Although mainly dominant during G1 or G0 phases of the cell cycle, NHEJ is operational during the whole cell cycle as it does not require a homologous template for repair of the damaged DNA strands (Rothkamm *et al.*, 2003). This lack of a homologous template means NHEJ is a relatively error prone mechanism of DNA repair and can lead to the introduction of insertions and deletions (indels).



**Figure 1.10 Overview of non-homologous end joining pathway.**

1. DNA double strand breaks are introduced. 2. Ku70 and Ku80 detect and bind to the DNA break. 3. DNA-PKcs is recruited to the site of damage. 4. DNA-PKcs is auto-phosphorylated and the other components of the NHEJ complex are recruited including Artemis, XRCC4, Ligase IV, PAXX and XLF. 5. DNA ends are ligated and the NHEJ complex disassembles from the DNA. The DNA break is successfully repaired. Adapted from: (Blackford and Jackson, 2017).

### 1.9.2 DNA-PKcs in transcriptional regulation

Interestingly, before DNA-PK was characterised as a DDR-associated protein, it was shown to be required for SP1 transcriptional activity as it was initially isolated with SP1 transcriptional complexes (Jackson *et al.*, 1990), and as an RNA polymerase II co-regulator (Dvir *et al.*, 1992). Although critical in DNA DSB repair, DNA-PK has been widely investigated outside of the repair process and in cancer, and its critical role in transcriptional regulation has been confirmed for several transcription factors, including FL-AR (Goodwin *et al.*, 2013, Goodwin and Knudsen,

2014). DNA-PK has already been described as a potential therapeutic target in cancers where DNA-PK expression is higher, such as in B-cell chronic lymphocytic leukaemia (Willmore *et al.*, 2008), colorectal cancer (Hosoi *et al.*, 2004), oesophageal cancer (Tonotsuka *et al.*, 2006) and non-small cell lung cancer (Xing *et al.*, 2008). Clinical trials of DNA-PKcs inhibitors are described in section 4.1.

There have been many reports indicating interplay between DNA-PK and hormone receptors, with reports from over 20 years ago describing phosphorylation of the chicken PR by DNA-PK (Weigel *et al.*, 1992) and also the rat GR, ultimately effecting their transcriptional activity. Giffin *et al.*, described that co-localisation of DNA-PK with the GR caused the GR to become phosphorylated and could affect receptor function (Giffin *et al.*, 1997). DNA-PKcs has also been shown to regulate the transcriptional activity of the oestrogen receptor, involved in cell cycle progression and proliferation of breast cells (Medunjanin *et al.*, 2010). Importantly, DNA-PKcs has also been shown to play a role in mediating the activity of the AR (Goodwin *et al.*, 2013, Goodwin *et al.*, 2015).

### 1.9.3 DNA-PKcs and the AR

Goodwin *et al.*, firstly reported that the AR is activated in response to DNA damage, causing transcription of several proteins that are involved in the DDR (Goodwin *et al.*, 2013). Later, the same group reported that DNA-PKcs plays a key role in the progression and metastases of PCa by regulating transcription (Goodwin *et al.*, 2015). They demonstrated that DNA-PKcs can potentiate AR function and described DNA-PKcs as both a transcriptional target of the AR after DNA damage and a regulator of its transcriptional activity (Goodwin *et al.*, 2013). Initial experiments *in vitro* show an interaction between the AR and DNA-PKcs and that DNA-PKcs is recruited to AR target genes such as *PSA* and *TMPRSS2*. The same AR regulated genes were downregulated following DNA-PKcs depletion and genetic analysis of all of the downregulated genes revealed an enrichment in pathways that are involved in cancer progression. Interrogation of DNA-PKcs functions *in vivo* showed that inhibition of DNA-PKcs using NU7441 reduced the incidence of metastases in xenograft AR-negative PC3 and AR-positive CWR22Rv1 mouse models, supporting the suggestion that DNA-PKcs is a key driver of metastasis (Goodwin *et al.*, 2015). Furthermore, PCa tissue, derived from patient biopsies, was cultured *ex vivo* and metastatic factors, such as PREX1, ROCK2, ITGB4, and VAV3 were quantified and

were all downregulated after administration of the DNA-PKcs inhibitor NU7441 to the culture media (Goodwin *et al.*, 2015). Finally, retrospective DNA-PKcs expression analysis was performed on a cohort of 232 patients to determine if DNA-PKcs could be used as a predictor of recurrence and metastases. Results showed higher DNA-PK expression correlated with a significantly worse freedom from metastatic progression and reduced freedom from biochemical recurrence (Goodwin *et al.*, 2015). Collectively, these results strongly support the hypotheses that DNA-PKcs is essential for AR transcription and promotes metastasis of PCa, highlighting it as an important therapeutic target in PCa and are the basis for a Phase I/II clinical trial using CC-115; a dual inhibitor of mTOR and DNA-PK. This trial aims to evaluate the safety of the drug in CRPC patients with metastases, in combination with enzalutamide (Clinicaltrials.gov; NCT02833883). Details of other DNA-PKcs inhibitors utilised in this study are provided below.

#### 1.9.4 DNA-PKcs inhibitors used in this study

##### **NU7441**

NU7441 (2-N-morpholino-8-dibenzothiophenyl-chromen-4-one) is a potent and specific small-molecule inhibitor of DNA-PK with a 100-fold selectivity for DNA-PK compared to other members of the PIKK family of enzymes and has an IC<sub>50</sub> of 14 nM (Leahy *et al.*, 2004). The compound targets the ATP binding-groove of DNA-PKcs. The effects of NU7441 inhibition has been investigated in many human cancer models including breast cancer (Ciszewski *et al.*, 2014), colon cancer (Zhao *et al.*, 2006), liver cancer (Yang *et al.*, 2016) and nasopharyngeal carcinoma (Dong *et al.*, 2018), as well as PCa (Goodwin *et al.*, 2015). In all these studies, NU7441 has been used to investigate potentiating radio- and/or chemotherapy through inhibiting DNA repair. It has also been used in models interrogating the role of AR-Vs in relation to the DDR; supporting a role for AR-Vs in driving a DDR following DNA damage (Yin *et al.*, 2017). Although a relatively selective DNA-PKcs inhibitor suitable for laboratory use, this small-molecule lacks the specificity required to progress into clinical trials.

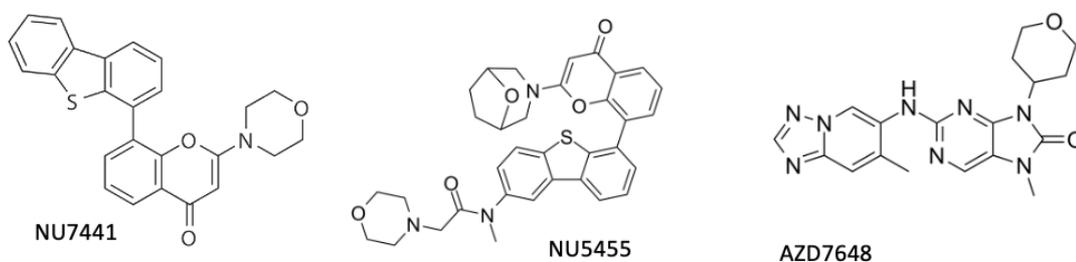
##### **NU5455**

NU5455 (N-(6-(2-(8-oxa-3-azabicyclo [3.2.1]octan-3-yl)-4-oxo-4H-chromen-8-yl)dibenzo[b,d]thiophen-2-yl)-N-methyl-2-morpholinoacetamide) is a newer, more selective DNA-PKcs inhibitor. Like NU7441, it competes for the ATP binding grove of DNA-PKcs and

inhibits DNA-PKcs autophosphorylation at Serine 2056, which is consistent with the earlier generation compound NU7441. This new inhibitor is more specific and has considerably less inhibitory effects on PI3K as determined by phospho-AKT Ser473 immunoblotting after 30 minutes IGF-1 treatment in MCF7 cells (Willoughby *et al.*, 2020). In the same study, activity against DNA-PKcs was quantified after 1  $\mu$ M treatment and DNA-PKcs was inhibited by 98% with an  $IC_{50}$  of  $8.2 \pm 2$  nM, which is increased potency over NU7441 ( $IC_{50}$  of 14 nM). In a kinase screen it was shown to have selectivity for DNA-PKcs over Vps34 (8.7-fold), PI3K $\delta$  (33.7-fold), and ATM and ATR (>1200-fold). NU7441 was shown to have non-specific inhibitory effects on PI3K $\alpha$ , but NU5455 demonstrated a 228-fold selectivity margin for DNA-PKcs kinase activity.

### **AZD7648**

AZD7648 (7-methyl-2-[(7-methyl [1,2,4]triazolo [1,5-a]pyridin-6-yl)amino]-9-(tetrahydro-2H-pyran-4-yl)-7,9-dihydro-8H-purin-8-one) was identified from a compound library screen conducted by AstraZeneca to identify potent DNA-PKcs inhibitors with selectivity over PI3K $\alpha$ . The hit compound was then optimised to improve its potency and pharmacokinetics to generate AZD7648. This compound has a  $IC_{50}$  of 0.6 nM and in a kinase screen DNA-PK, PI3K $\alpha$ , PI3K $\delta$  and PI3K $\gamma$  showed inhibition of >50% at 1  $\mu$ M. AZD7648 was then shown to have selectivity for DNA-PKcs versus PI3K $\alpha$  and PI3K $\delta$  (>100-fold) and PI3K $\gamma$  (>63-fold). The compound was shown to potentiate radiation in A549 cells demonstrated by persistent DNA damage markers  $\gamma$ H2AX and 53BP1 foci and decreased clonogenic survival in combined radiation plus AZD7648 experimental arms (Fok *et al.*, 2019). Clinical trials have progressed with this compound and is discussed further in section 4.1.



**Figure 1.11 Chemical structures of NU7441, NU5455 and AZD7648.**

### 1.9.5 The androgen receptor and the DNA damage response

There have been multiple clinical trials investigating the efficacy of combined radiotherapy and ADT versus radiotherapy alone in PCa. A systematic review with meta-analyses of randomised control trials of hormone therapy used alone or in combination with radiotherapy for PCa reported that the efficacy of the combined therapies was superior to each therapy alone, which increased further in advanced, higher risk patients (Schmidt-Hansen *et al.*, 2014). Although this report indicates that some results reported could be biased, it supports the hypothesis that androgen signalling is involved in DDR following radiation.

The FL-AR is involved in the DDR through directly up-regulating transcription of DNA repair genes that are involved in NHEJ, HR, base excision repair and mismatch repair in response to DNA damage (Polkinghorn *et al.*, 2013b). It was reported that PCa cells treated with anti-androgens downregulated their DNA repair genes, and conversely PCa cells treated with androgens upregulated their DNA repair genes, in response to DNA damage caused by ionising radiation. Other DDR proteins such as PARP1 have also been identified as AR co-regulators (Schiewer *et al.*, 2012).

There are various therapeutic agents available in the clinic or in clinical development that target the DDR, either as single agent treatments to further the genomic instability in cancer cells, or in combination with DNA damaging therapies to enhance DNA damage and sensitise cells, such as PARP1, ATR and ATM inhibitors (Plummer *et al.*, 2008, Prevo *et al.*, 2012, Hickson *et al.*, 2004). Radiotherapy is commonly given to PCa patients with localised disease and after radical prostatectomy as salvage therapy. Importantly, as discussed, radiotherapy is commonly used in combination with ADT as it has been shown to be more effective than either agent alone which is, in part, a consequence of down-regulating AR-driven DDR gene expression (Schmidt-Hansen *et al.*, 2014, Polkinghorn *et al.*, 2013a). However, patients often develop radio-resistance and the mechanism of action for this is largely unknown. The AR has been suggested to be important in this radio-resistance, which could be due to its direct role in DDR or the upregulation of a number of key DDR-related genes. Recently, it has been reported that AR-Vs drive a DDR gene signature by driving transcription of DDR genes that lead to a desensitisation of radiation shown by persistence of  $\gamma$ H2AX foci (Kounatidou *et al.*,

2019). Targeting AR-Vs in these patients could potentiate sensitivity to DNA damaging therapies and lead to a better outcome.

#### 1.9.6 DNA-PKcs as a potential regulator of AR-V activity

Due to the reports of DNA-PKcs as a regulator of FL-AR, DNA-PKcs could also regulate AR-Vs. Unpublished rapid immunoprecipitation mass spectrometry of endogenous protein (RIME) data obtained in the host laboratory, has indicated that DNA-PKcs is part of the AR-V interactome. Another study that used immunoprecipitation-mass spectrometry (IP-MS) identified DNA-PKcs as the most abundant interacting protein of AR<sup>v567es</sup> in R1-D567 cells irrespective of if the cells had been exposed to irradiation (Yin *et al.*, 2017). Both of these findings strongly suggest DNA-PKcs plays a role outside of DNA repair in the regulation of AR-Vs, but little is known around the mechanism or its function in transcriptional capability of AR-Vs.

Interrogation of an in-house AR-V transcriptome derived from RNA-sequencing of the CWR22Rv1 cell line derivative CWR22Rv1-AR-EK, to determine if AR-Vs have a direct role in DNA damage repair, showed that AR-Vs positively regulated 41 genes involved in the DDR (Kounatidou *et al.*, 2019). These genes are involved in HR and NHEJ pathways.

The expression of AR-Vs is a major mechanism for resistance to current PCa treatments, and our understanding of their regulation is limited. As DNA-PKcs has been identified as a co-regulator of FL-AR, and that previous data from the host laboratory has demonstrated an interaction between DNA-PKcs and AR-Vs, we hypothesise there could potentially be a similar co-regulatory relationship to FL-AR between the proteins. Therefore, the co-regulatory relationship between DNA-PKcs and AR-Vs will be investigated in this project using both siRNA mediated depletion of the DNA-PKcs protein and DNA-PKcs inhibition using various compounds that have been developed.

## Chapter 2 Aims and objectives

There are several resistance mechanisms that render prostate tumours insensitive to the current repertoire of AR-targeting agents including *AR* amplification (Taylor *et al.*, 2010), *AR* mutation (Balbas *et al.*, 2013) and the emergence of *AR* splice variants (Dehm *et al.*, 2008). Importantly, although classified as castrate-resistant, *AR* signalling remains active in these tumours and drives PCa growth and progression (Chen *et al.*, 2004). The generation of alternatively spliced *AR*-Vs that are constitutively active and drive transcription of target genes in the absence of androgens represent a major clinical challenge as there remains no approved therapies that compromise their activity. A better understanding of *AR*-V biology and determining how they are generated and regulated will provide new avenues for identifying tractable drug targets. Furthermore, resistance to radiotherapy is another major clinical problem (Bolla *et al.*, 2010). It has been established that FL-*AR* can drive transcription of DDR genes which in turn can contribute to radio-resistance (Polkinghorn *et al.*, 2013a). Therefore, a combination of ADT plus radiotherapy can potentiate sensitivity to DNA damaging therapies (Bolla *et al.*, 2010). However, despite some reports of *AR*-Vs driving transcription of DDR genes (Kounatidou *et al.*, 2019), and an indication of *AR*-Vs interacting with sites of DNA damage (Yin *et al.*, 2017), little is known about the mechanisms in which *AR*-Vs themselves contribute to DNA repair. Determining what role DNA-PKcs plays in the regulation of *AR*-Vs as well as the role *AR*-Vs play in the DDR will provide a rationale and means to target *AR*-Vs alongside radiotherapy in advanced, *AR*-V positive PCa to improve treatment response. Therefore, the objectives of this study are to:

- Investigate DNA-PKcs as a regulator of *AR*-V transcriptional activity
- Investigate wider transcriptional changes in response to DNA-PKcs inhibition and depletion using RNA-sequencing
- Develop an assay to identify the interactome of *AR*-V7 in the presence and absence of DNA damage by ionising radiation using a proximity biotinylation – mass spectrometry approach.

## Chapter 3 Materials and Methods

### 3.1 Mammalian cell culture

CWR22Rv1 (ATCC® CRL-2505™), VCaP (ATCC® CRL-2876™), PC3 (ATCC® CRL-1435™), and HEK293T (ATCC® CRL-3216™) cell lines were purchased from ATCC (Virginia, US). CWR22Rv1-AR-EK is a CRISPR-engineered cell line generated from CWR22Rv1 cells in the host laboratory (Kounatidou *et al.*, 2019). CWR22Rv1-AR-EK has a premature stop codon knocked into the *AR* gene in exon 5 so the cells express all endogenous AR-Vs but does not express FL-AR.

#### 3.1.1 Subculturing

Cell lines (Table 3.1) were cultured at 5% CO<sub>2</sub> and 37 °C in RPMI-1640 (R5886, Sigma Aldrich) (CWR22Rv1, CWR22Rv1-AR-EK, HEK293T and PC3 cell lines) or DMEM (D5030, Sigma Aldrich) (VCaP cell line) supplemented with 10% (v/v) foetal calf serum (FCS) (Sigma Aldrich) and 2 mM L-glutamine (Sigma Aldrich) (referred to as full media). Cells seeded down for experiments to analyse AR-V activity (CWR22Rv1 parental cell line and VCaP) were seeded in RPMI 1640 or DMEM respectively, supplemented with 10% (v/v) dextran-charcoal stripped steroid-depleted FCS (Hyclone) and 2mM L-glutamine (referred to as steroid depleted media (SDM)) and incubated for 48 hours prior to 24-hour drug treatments or seeded in SDM with relevant transfections for 72 hours.

Cell culture was carried out in a BioMat class II Biosafety hood in sterile conditions. Passaging of cells was performed every 2-3 days by discarding old media and washing cells with warm phosphate buffered saline (PBS) prior to detaching the cells in 1 x trypsin (Sigma Aldrich) in PBS at 37 °C. The cell suspension was centrifuged at 400 x g for 5 minutes, supernatant removed, and cells resuspended in media before re-seeding in 175 cm<sup>2</sup> flasks (Corning). Cell lines were routinely tested for mycoplasma in-house.

**Table 3.1 Cell lines**

Cell Line	FL-AR/AR-V expression in FM	FL-AR/AR-V expression in SDM	Media cultured in	Origin
CWR22Rv1	+++/>++	++/>+++	RPMI 1640	Xenograft tumour from primary tumour
CWR22Rv1-AR-EK	-/>+++	-/>+++	RPMI 1640	CWR22Rv1 CRISPR-derivative
VCaP	++++/>+	+++/>++	DMEM	Xenograft tumour from metastasis
PC3	-/>-	-/>-	RPMI 1640	Bone metastasis
HEK293T	-/>-	-/>-	RPMI 1640	Embryonic kidney

### 3.1.1 Cell storage

For freezing and storage of cell lines,  $1 \times 10^6$  cells were resuspended per mL of freezing media (RPMI or DMEM media supplemented with 10% FCS and 10% DMSO) and transferred to cryogenic vials (Thermo Scientific) in 1 mL aliquots. Vials were then frozen in a freezing container at  $-80\text{ }^{\circ}\text{C}$  and kept at  $-80\text{ }^{\circ}\text{C}$  for short-term (1-2 months) or transferred to liquid nitrogen for long term ( $> 2$  months) storage. To thaw for subsequent culturing, cells were thawed, centrifuged (3 min,  $400 \times g$ ) and resuspended in full media before seeding into T25 flasks. Cells were then passaged at least 2 times before use in experiments.

### 3.1.2 Compounds

#### **DNA-PKcs inhibitors**

NU7441 and NU5455 were synthesised at Newcastle University's chemistry department and were resuspended in DMSO at a concentration of 10 mM and 5 mM respectively, and stored at  $-20\text{ }^{\circ}\text{C}$ . See Table 3.2 for the concentrations used in experiments for all compounds.

AZD7648 (Selleckchem) was resuspended in DMSO at a concentration of 10 mM and stored at  $-80\text{ }^{\circ}\text{C}$ .

#### **Anti-androgens**

Enzalutamide (Selleckchem), a second-generation androgen receptor antagonist, was purchased as a powder and resuspended in DMSO at a concentration of 30 mM and stored at  $-80\text{ }^{\circ}\text{C}$

Darolutamide (Selleckchem), a second-generation androgen receptor antagonist, was purchased as a powder and resuspended in DMSO at a concentration of 10 mM and stored at  $-80\text{ }^{\circ}\text{C}$ .

### **Synthetic androgens**

Dihydrotestosterone (DHT) (Sigma), an androgen receptor agonist, was purchased as a powder and resuspended in 20% ethanol at a concentration of 10 mM and stored at  $-80\text{ }^{\circ}\text{C}$ .

R1881 (Perkin Elmer), a synthetic androgen receptor agonist, was purchased as a powder and resuspended at 1 mM in DMSO and stored at  $-80\text{ }^{\circ}\text{C}$ .

Carfilzomib (Selleckchem), a proteasome inhibitor, was purchased as a powder and resuspended in DMSO at a concentration of  $5\text{ }\mu\text{M}$  and stored at  $-20\text{ }^{\circ}\text{C}$ .

### **Antibiotics**

Doxycycline (Merck), a broad-spectrum tetracycline-class antibiotic, was purchased as a powder and resuspended in DMSO at a concentration of 1 mg/mL and stored at  $-20\text{ }^{\circ}\text{C}$ .

Puromycin dihydrochloride (Sigma), a mammalian antibiotic, was purchased as a powder and resuspended in DMSO at a concentration of 1 mg/mL and stored at  $-20\text{ }^{\circ}\text{C}$ .

**Table 3.2 List of compounds used with their respective concentrations**

<b>Compound</b>	<b>Concentration(s) used</b>
NU7441	0.1 – 5 $\mu\text{M}$
NU5455	0.1 – 5 $\mu\text{M}$
AZD7648	0.1 – 5 $\mu\text{M}$
Enzalutamide	10 $\mu\text{M}$
Darolutamide	10 $\mu\text{M}$
DHT	10 nM
R1881	1 nM

Carfilzomib	200 nM
Puromycin	1-5 µg/mL
Doxycycline	1 µg/mL

### 3.2 siRNA and synthetic guide RNA transfection

Small interfering RNA (siRNA) and guide RNA (gRNA), purchased from Dharmacon (DNA-PKcs smartpool) or Sigma were used for gene knockdown/knockout. gRNA was composed of composite crRNA/tracrRNA, called single guide (sgRNA), to enable single molecule transfections of the gRNAs. siRNA and sgRNA were purchased in lyophilised form and resuspended to 25 µM in sterile RNase/DNase free water and stored at -20 °C. Lipofectamine® RNAiMAX transfection reagent (Invitrogen) was used for siRNA and sgRNA delivery to cells. Transfection mixes made up of siRNA: RNAiMAX at a 1:2 ratio were incubated in RPMI1640 without FCS (referred to as basal media) for 30 minutes at room temperature. Mixes were then added to appropriate wells before drop-wise addition to cells to give a final concentration of 25 nM siRNA and 0.1% RNAiMAX. See Table 3.3 for siRNA and sgRNA sequences. Cells were incubated for 72 hours at 37 °C before harvesting.

**Table 3.3 Small interfering RNA and synthetic guide RNA sequences**

Oligonucleotide	Sequence (5' – 3')
siScr	UUCUCCGAACGUGUCACGU
DNA-PKcs si1	GGAAGAAGCUCAUUUGAUU
DNA-PKcs si2	GAGCAUCACUUGCCUUUAA
DNA-PKcs si3	GCAGGACCGUGCAAGGUUA
DNA-PKcs si4	AGAUAGAGCUGCUAAAUGU
siRBMX	AUCAAGAGGAUUAUAGCGA
sgScr	AACCCCTGATTGTATCCGCA
sgAR	ATTCAGCCAAGCTCAAGGA
sgRBMX	AUCAAGAGGAUUAUAGCGA

### 3.3 RNA extraction

RNA was extracted from cells using TRIzol™ according to the manufacturer's handbook (Invitrogen). Briefly, CWR22Rv1 and CWR22Rv1-AR-EK were seeded at a cell density of 250,000 cells per well in 6-well plates (Corning) for 48 hours, followed by 24-hour drug treatment (CWR22Rv1-AR-EK) or 24-hour co-treatment plus/minus enzalutamide (CWR22Rv1). VCaP cells were seeded at a cell density of 300,000 cells per well in 6-well plates in SDM for 48 hours, followed by 24-hour drug treatment, in the presence and absence of enzalutamide. Cells were then washed twice with PBS prior to addition of 500 µL TRIzol™ reagent and incubated for 10 minutes at room temperature with agitation. Lysates were transferred into Eppendorf tubes and thoroughly mixed with 100 µL chloroform before incubation at room temperature for 2 minutes and centrifuged at 12,000 x g for 15 minutes at 4 °C. The RNA-containing aqueous phase was then transferred into a fresh Eppendorf tube, and RNA precipitated by addition of 500 µL isopropanol and incubation at room temperature for 10 minutes before centrifuging samples at 12,000 x g for 10 minutes at 4 °C. Supernatants were removed and the resultant RNA pellet was washed twice in 75% ethanol prior to air-drying and resuspension in 30 µL molecular grade water. RNA was incubated at 55 °C for 5 minutes and then quantified, and quality checked using a spectrophotometer (Nanodrop, Thermo Scientific) before being stored at -80 °C.

### 3.4 Reverse transcription

1 µg of RNA was incubated at 37 °C for 1 hour with reverse transcription reagents (Table 3.4), before termination of the reaction at 100 °C for 10 minutes. The cDNA samples were then diluted in 150 µL molecular grade water for analysis by quantitative real-time polymerase chain reaction (PCR) (qRT-PCR). Samples were stored at -20 °C.

**Table 3.4 Components of Reverse transcription reaction with their respective volumes**

Reagent	Volume/reaction (µL)
1 µg RNA	12.7
M-MLV Reverse Transcription Buffer (5x)	4
4 mM dNTPs	2
Oligo (DT) <sub>18</sub> Primers (100 µg/ml)	1
M-MLV Reverse Transcriptase	0.3

### 3.5 Quantitative RT-PCR

Analysis of AR and AR-V target gene expression was performed using qRT-PCR incorporating SYBR® Green DYE 1 (Life technologies) and custom primers which were purchased from Sigma (Table 3.6). 384-well qPCR plates were used with a 10 µL total reaction volume in each well containing the reagents shown in Table 3.5 and cDNA from the reverse transcription reaction. Each reaction was performed in triplicate alongside a no template control (primer mix containing SYBR® only) to determine if there was any contamination in the primer sets. Plates were sealed with MicroAmp optical adhesive films and centrifuged in a plate centrifuge at 100 x g for 20 seconds. Plates were analysed in a QuantStudio 12k Flex Real-Time PCR system (Thermo-Fisher). The comparative cycle threshold (Ct) method was used to determine the relative gene expression in samples. Ct values were normalised to a *RPL13A* housekeeping gene. The following equation was used to for comparison of gene expression;

$$[\Delta][\Delta]Ct = [\Delta]Ct, \text{ sample} - [\Delta]Ct, \text{ reference}$$

**Table 3.5 Components used per well for qRT-PCR with their respective volumes.**

Reagent	Volume/ well ( $\mu\text{L}$ )
Forward Primer (25 ng/ $\mu\text{L}$ )	0.4
Reverse Primer (25 ng/ $\mu\text{L}$ )	0.4
dH <sub>2</sub> O	2.2
SYBR® Green qPCR SuperMix (5x)	5
cDNA	2

**Table 3.6 Primer sequences used in qPCR.**

Gene	Forward (5'-3')	Reverse (5'-3')
<i>RPL13A</i>	CCTGGAGGAGAAGAGGAAAGAGA	TTGAGGACCTCTGTGTATTTGTCAA
<i>PSA</i>	GCAGCATTGAACCAGAGGAG	AGAAGTGGGGAGGCTTGAGT
<i>KLK2</i>	AGCATCGAACCAGAGGAGTTCT	TGGAGGCTCACACACCTGAAGA
<i>UBE2C</i>	TGCCCTGTATGATGTCAGGA	GGGACTATCAATGTTGGGT
<i>CCNA2</i>	GAAGACGAGACGGGTTGCA	AGGAGGAACGGTGACATGCT
<i>TMPRSS2</i>	CTGCTGGATTTCCGGGTG	TTCTGAGGTCTTCCCTTTCTCCT
<i>PRKDC</i>	GAGAAGGCGGCTTACCTGAG	CGAAGGCCCGCTTTAAGAGA
<i>AR-V1</i>	AACAGAAGTACCTGTGCGCC	TGAGACTCCAAACACCCTCA
<i>AR-V6</i>	AACAGAAGTACCTGTGCGCC	TATGACTCTGCTGCCTTGC
<i>AR-V7</i>	AACAGAAGTACCTGTGCGCC	TCAGGGTCTGGTCATTTTGA
<i>AR-V9</i>	AACAGAAGTACCTGTGCGCC	GCAAATGTCTCCAAAAGCAGC
<i>FL-AR</i>	AACAGAAGTACCTGTGCGCC	TTCAGATTACCAAGTTTCTTCAG
<i>RBMX</i>	TGGAAGAGGAGGAAGTGGAGG	GGTCCCCTGGAAGAACTCAT

### 3.6 Agarose gel electrophoresis

1% agarose gels were made by mixing agarose (Thermo Scientific) in 1x Tris-acetate-EDTA (TAE) buffer with the addition of 10,000x Gel Red nucleic acid stain (Biotium) to a final concentration of 1x to separate DNA resulting from plasmid maxi-preparations or restriction enzyme digestion. 5  $\mu$ L of a 100 bp and/or a 10 kb Hyperladder (NEB) were loaded per gel to estimate DNA size. DNA samples were mixed with a 6x loading buffer (NEB) prior to loading and were ran at 70 Volts for 30-40 minutes. Visualisation of bands was carried out using an ultraviolet transilluminator to view cut out bands or ChemiDoc system to image the gel.

### 3.7 Chromatin Immunoprecipitation

Chromatin immunoprecipitation (ChIP) was performed as described in (Jones *et al.*, 2015). Cells were cultured in full media in 150 mm dishes at a cell density of  $5 \times 10^6$  (CWR22Rv1 cells incubated in full media were changed to SDM media for 24 hours prior to 24-hour enzalutamide treatment), until 70-80% confluent, before 24-hour treatment with 1  $\mu$ M NU7441. Cells were then fixed using 1% formaldehyde, gently mixed and incubated at room temperature for 7-10 minutes. Glycine was added to a final concentration of 1.25 mM to stop fixation, gently mixed and incubated for a further 5 minutes at room temperature. Cells were then washed twice in 12 mL ice cold PBS before adding 2 mL ice cold PBS containing protease inhibitors (Roche). Cells were scraped into the PBS and transferred into cold falcon tubes and then centrifuged at 400 x g for 5 minutes before removing the supernatant and snap freezing the pellets using liquid nitrogen. The pellets were then stored at -80 °C prior to processing.

Pellets were thawed on ice and resuspended in 5 mL Lysis Buffer 1 (LB1) (Recipes for all ChIP reagents in Table 3.7) and agitated for 10 minutes on ice, then centrifuged at 1500 x g for 5 minutes at 4 °C. The supernatant was removed, and pellets were resuspended in 5 mL lysis buffer 2 (LB2) and agitated on ice for 10 minutes. The sample was centrifuged again at 1500 x g for 5 minutes at 4 °C, the supernatant was removed, and the pellet was resuspended in 300  $\mu$ L lysis buffer (LB3) and transferred into Eppendorf tubes. Samples were sonicated using a Biorupter (Diagenode) for 30-45 minutes (30 seconds on/ 30 seconds off cycles) and then centrifuged at 12,000 x g for 10 minutes at 4 °C. The supernatant was transferred into a fresh Eppendorf tube and quantified using a spectrophotometer (Nanodrop). 70  $\mu$ g of chromatin was used per ChIP experiment, which was diluted to a total volume of 700  $\mu$ L using LB3/triton

X-100. 70 µL was taken as an input sample for downstream data analysis and stored at -20 °C until the protein-DNA crosslink reversal step. The remaining 63 µg of chromatin was incubated for 16 hours at 4 °C on rotation with protein-A-conjugated Dynabeads™ (Thermo Scientific) attached to 2 µg appropriate antibodies.

After incubation, Dynabeads™ were washed 5 times with cold RIPA buffer and once with cold Tris-buffered saline (TBS) at 4 °C. Samples were then centrifuged at 12,000 x g for 5 minutes at 4 °C and supernatants removed before Dynabeads™ and input samples taken the previous day were resuspended in 200 µL ChIP elution solution and incubated at 65 °C for 8 hours to reverse protein-DNA crosslinks. Supernatants of immunoprecipitation (IP) and input samples were then transferred into fresh Eppendorf tubes, diluted in 200 µL TE buffer pH 8 and stored at -20 °C overnight. Samples were thawed and incubated with 4 µL of Proteinase K (Life Technologies) at 55 °C for 1 hour to degrade proteins. DNA was then purified using GeneElute™ genomic mammalian mini-prep kit (Sigma Aldrich) according to the manufacturer's handbook. Purified DNA samples were eluted in 130 µL molecular grade water and subject to qRT-PCR (see section 3.5) using primers complementary to *cis*-regulatory elements of AR target genes (Table 3.8) to assess protein enrichment at these loci. Data was analysed using % input calculated using the formula: % Input =  $100 \times 2^{((\text{input}-3.2)-\text{IP Ct})}$ . The values were normalised to the DMSO value to determine relevant enrichment at each AR target loci.

**Table 3.7 ChIP solutions and buffers**

Buffer	Reagents
LB1	50 mM HEPES-KOH, pH7.5, 140 mM NaCl, 1 mM EDTA, 10% glycerol, 0.5% NP-40 and 0.25% Triton X-100
LB2	10 mM Tris-HCl pH8, 200 mM NaCl, 1 mM EDTA and 0.5 mM EGTA
LB3	100 mM Tris-HCl, pH8, 100 mM NaCl, 1mM EDTA, 0.5 mM EGTA, 0.1% Na-deoxycholate and 0.5% N-lauroylsarcosine
RIPA Buffer	50 mM HEPES pH 7.5, 500 mM LiCl, 1 mM EDTA, 1% NP-40, 0.7% Na-deoxycholate
TBS	20 mM Tris pH 7.6, 150 mM NaCl
ChIP Elution solution	50 mM Tris-HCl pH8, 10 mM EDTA, 1% SDS

TE Buffer	10 mM Tris pH 8, 1 mM EDTA
-----------	----------------------------

**Table 3.8 Primer Sequences for CHIP analysis**

Gene	Forward (5'-3')	Reverse (5'-3')
<i>PSA AREIII</i>	TGGGACAACCTTGCAAACCTG	CCAGAGTAGGTCTGTTTTCAATCCA
<i>KLK2</i>	GGTTGAAAGCAGACCTACTCTGG	AGATCTAGGTTTGCTTACTGCCTTAG
<i>UBE2C</i>	TGCCTCTGAGTAGGAACAGGTAAGT	TGCTTTTTCCATCATGGCAG
<i>CCNA2</i>	TTAGTGAGCTGTCCAGTGACTION	CCCATGTATTAAGTAGCTTCTGTAAACA
<i>TMPRSS2</i>	TGGTCCTGGATGATAAAAAAAGTT	GACATACGCCCCACAACAGA
<i>AR intron 2A</i>	CCATCATGTGCATTATGTGC	CCAAACAGCACTCCATGTGT
<i>AR intron 2B</i>	CACATGGAGTGCTGTTTGGT	GTAAACATCAGTGAGGATGGTG
<i>AR intron 2C</i>	CACCATCCTCACTGATGTTTAC	TGAGGGTTCACCTGCATTTTC
<i>RBMX (-0)</i>	CAACGAGCTCGGCGATAGG	GTAGTGCTAGCGGCTTCGC
<i>RBMX (-500)</i>	TTCAACCCAGAACCACCGAC	GGCTTCGTATTCATTGGCGG
<i>RBMX (-1000)</i>	GCAACAGCTGCTTAACATTTGA	ACCATCGTTAGGAAGGGTGTG
<i>RBMX (-4000)</i>	ACTGCTACTGCGAACTGGTC	ATTTTTGATGCAGATGACGGTG

### 3.8 RNA Immunoprecipitation

Cells were cultured in full media in 150 mm dishes at a cell density of  $5 \times 10^6$  (CWR22Rv1 cells seeded in SDM media for 5 days with 10  $\mu$ g plasmid transfections at day 0 and day 3), until 70-80% confluent. Cells were then fixed using 0.2% formaldehyde, gently mixed and incubated at room temperature for 15 minutes. Glycine was added to a final concentration of 1.25 mM to stop fixation, gently mixed and incubated for a further 5 minutes at room temperature. Cells were then washed twice in 10 mL ice cold PBS before adding 2 mL ice cold PBS containing protease inhibitors (Roche). Cells were scraped into the PBS and transferred into cold falcon tubes and then centrifuged at 400 x g for 5 minutes before removing the supernatant and

snap freezing the pellets in their falcon tubes using liquid nitrogen. The pellets were then stored at -80 °C prior to processing.

Pellets were thawed on ice and resuspended in 200 µL RIPA buffer (Table 3.9) supplemented with 100U/ mL RNase OUT (Thermo Scientific), 1 x protease inhibitors (Roche), 400 µM VRC and 1 mM DTT (Sigma), and incubated for 20 minutes on ice. Samples were then sonicated using a Biorupter (Diagenode) on low for 3 x 30 seconds on/off cycles and then centrifuged at 15,000 x g for 15 minutes at 4 °C. The supernatant was transferred into a fresh Eppendorf tube and a 10% (20 µL) input sample was taken and stored at -80 °C until required.

DynaBeads (Thermo Scientific) were washed and blocked 2 times with 1 mL 0.5% bovine serum albumin (BSA) in PBS and then conjugated with 5 µg appropriate antibody by rotating at 4 °C for 8 hours. Antibody conjugated DynaBeads were then washed 4 times in 1 mL NT2 buffer (Table 3.9), resuspended in 300 µL NT2 and before adding the 200 µL cell lysate supplemented with 100U RNase OUT, 400 µM VRC 1 mM DTT and 15 mM EDTA. The bead/lysate solution was then incubated at 4 °C overnight on a rotator to bind the antibody to the protein of interest. The following day, beads were washed 7 times with 1 mL NT2 and after the final wash, beads were briefly centrifuged to collect beads at the bottom of the tube and the supernatant was removed. Beads were then resuspended in 100 µL NT2, 4.8 µL 5M NaCl and 10 µL proteinase K. Input samples were thawed and 95 µL NT2 4.8 µL 5M NaCl and 2 µL proteinase K was added. Input and IP samples were vortexed and incubated at 42 °C for 1 hour then 55 °C for 1 hour with frequent mixing to resuspend the beads. 1 mL of TRIzol was then added to the beads/lysate and the RNA was extracted as above (section 3.2) with 3 µL GlycoBlue™ (Invitrogen) added to the isopropanol precipitation step which was extended to overnight to improve RNA yield and an additional 75% ethanol wash. RNA pellets were resuspended in 15 µL nuclease free water. Up to 1 µg of RNA was reverse transcribed and subject to qRT-PCR for analysis of RNA enrichment with proteins of interest.

Data was analysed using the following equations:

(1) Inputs ->  $IgG1 = Ct(\text{Gene Of interest(GOI)}) - Ct(\text{Housekeeping gene(HG)})$

$$GFP1 = Ct(\text{GOI}) - Ct(\text{HG})$$

(2) IPs ->  $IgG2 = Ct(\text{GOI}) - Ct(\text{HG})$

$$GFP2 = Ct(\text{GOI}) - Ct(\text{HG})$$

(3)  $[\Delta][\Delta]Ct1 = IgG2 - IgG1$  &  $[\Delta][\Delta]Ct2 = GFP2 - GFP1$

(4)  $IgG = 2^{Ct1}$   $GFP = 2^{Ct2}$

(5) Fold Enrichment =  $GFP/IgG$

**Table 3.9 RNA immunoprecipitation buffers**

RNA immunoprecipitation buffer	Reagents
RIPA (Ting lab)	50 mM Tris, 150 mM NaCl, 0.1% (w/v) SDS, 0.5% (w/v) sodium deoxycholate, 1% (v/v) Triton X-100 in Nuclease free water, pH 7.5 (HCl)
NT2	50 mM Tris, 150 mM NaCl, 1 mM MgCl <sub>2</sub> , 0.05% IGEPAL

### 3.9 Western Blot

Cells were seeded in 6-well plates at a cell density of 250,000 cells per well for CWR22Rv1 and CWR22Rv1-AR-EK and 500,000 cells per well for VCaP cells in full media (CWR22Rv1-AR-EK) or SDM media (CWR22Rv1 and VCaP cells) for 48 hours. Cells were then treated for 24 hours with specific doses of NU7441 only or enzalutamide plus NU7441 before being harvested for western blotting. Cell lysates were acquired by adding 130  $\mu$ L of SDS sample buffer (Table 3.10) to each well before boiling at 100 °C for 10 minutes in Eppendorf tubes before resolution on 10% acrylamide gels by SDS-polyacrylamide-gel electrophoresis (SDS-PAGE) alongside a Spectra™ multicolor broad range protein ladder (Thermo Fisher)(Composition of gels in Table 3.10). Protein gels were routinely run at 140 volts in running buffer (Table 3.10) for approximately 60 minutes and then transferred onto a nitro-cellulose membrane (GE Healthcare) at 100 volts for 1 hour in transfer buffer (Table 3.10). Membranes were then blocked in 5% (w/v) milk powder (Marvel) dissolved in TBS for 1 hour before being washed twice in TBST (Table 3.10) and once in TBS before incubation at 4 °C overnight in primary antibody solution (1:1000 in 1% (w/v) milk in TBS) with rotation. Membranes were then washed twice in TBST and once in TBS before incubation with secondary antibody (1:1000 in 1% (w/v) milk in TBS) at room temperature for 1 hour with rotation. Antibodies and appropriate concentrations used are detailed in Table 3.11. Membranes were then developed

using enhanced chemiluminescence western blotting substrates (BioRad) using a ChemiDoc imaging system (Life Sciences, BioRad).

**Table 3.10 Western blot solutions and reagents**

<b>Solution</b>	<b>Reagents</b>
Buffer A 2X	750 mM Tris-HCl, pH 8.8, 0.2% SDS
Buffer B 2X	250 mM Tris-HCl, pH 6.8, 0.2% SDS
Running Gel 10%	Acrylamide (30% (w/w))(Sigma Aldrich) (3.33 mL), water (distilled) (1.67 mL), 2 X Buffer A (5 mL), N,N,N',N'-tetramethylethane1,2-diamine (TEMED) (20 µL), Ammonium persulphate (APS) (10% (w/v)) (100 µL)
Stacking Gel	Acrylamide (30%) (0.84 mL), Water (distilled) (1.67 mL), 2 X Buffer A (2.5 mL), TEMED (18 µL), APS (10%) (100 µl)
Running Buffer	25 mM Tris, 190 mM glycine, 0.1% SDS
Transfer Buffer	25 mM Tris-HCl, pH8.3, 150 mM glycine, 10% methanol
SDS Sample Buffer	125 mM Tris-HCl, pH6.8, 5% SDS, 10% glycerol, 10% β-mercaptoethanol and 0.01% bromophenol blue
TBS	500 mM NaCl, 200 mM Tris-HCl, pH 7.5
TBST	500mM NaCl, 200 mM Tris-HCl, pH 7.5, 0.1% Tween-20

**Table 3.11 Antibodies**

<b>Antibody</b>	<b>Species</b>	<b>Supplier</b>	<b>Application</b>	<b>Dilution</b>	<b>Catalogue number</b>
AR	Rabbit	Cell signaling	IF/ChIP	1:400/ 2 µg	5153
AR	Mouse	BD	WB	1:1000	554225
AR-V7	Rabbit	Abcam	WB	1:500	ab198394
AR-V7	Rabbit	Precision	WB	1:500	AG10008
DNA-PKcs	Rabbit	Abcam	WB/ChIP	1:500	Ab70250
DNA-PKcs	Rabbit	BioRad	WB	1:500	AHP318
p-DNA-PKcs	Rabbit	Abcam	WB	1:500	Ab18192
γH2AX	Mouse	Millipore	IF	1:400	NB100-74435

Biotin	Mouse	Santa Cruz	WB	1:1000	Sc-101339
Streptavidin-HRP	N/A	Abcam	WB	1:2000	Ab7403
FLAG M2	Mouse	Sigma	WB/ChIP	1:1000/ 5µg	F1804
Cas9	Mouse	Abcam	WB	1:1000	ab191468
RBMX	Rabbit	Cell signaling	WB	1:500	14794
GFP	Rabbit	Abcam	WB/RIP	1:1000/ 5 µg	Ab290
a-tubulin	Mouse	Sigma	WB	1:4000	T9026

### 3.10 Immunofluorescence

Cells were seeded at a cell density of  $1 \times 10^5$  onto sterile 22 mm x 22 mm glass cover slips (Thermo Scientific) in 6-well plates. After appropriate treatments, cells were fixed with 1 mL 4% (v/v) paraformaldehyde in PBS for 20 minutes at room temperature. Paraformaldehyde was removed and cells were washed twice in 1 mL PBS for 5 minutes before permeabilisation in 1 mL 0.1% triton X-100 (v/v) (Merck) in PBS for 10 minutes at room temperature. Cells were then washed twice in PBS for 5 minutes and blocked in 1 mL 4% (w/v) BSA (Merck) in PBS for 30 minutes at room temperature to prevent non-specific binding of antibody. Coverslips were then placed onto appropriate antibodies (Table 3.11) diluted in 4% BSA in PBS and incubated overnight at 4 °C. Control cells used for secondary antibody only were placed in 50 µL 4% BSA in PBS without primary antibody. The following day, cells were washed three times in PBS for 5 minutes and incubated with 50 µL of an appropriate AlexaFluor<sup>®</sup> (Invitrogen) secondary antibody at a 1:400 dilution for 1 hour at room temperature protected from light. Finally, following 3 final 1 mL PBS washes, coverslips were mounted onto slides using mounting media with DAPI DNA stain (Abcam). Slides were imaged using the Leica DM6 B widefield fluorescent microscope (Leica microsystems).

### 3.11 Cell proliferation assays

#### 3.11.1 Cell counts

CWR22Rv1-AR-EK and CWR22Rv1 cells were seeded at a cell density of 100,000 cells per well of a 6-well plate for 24 hours before being treated with 0.1, 0.5, 1, 2 and 5 µM NU7441 or

DMSO for 96 or 120 hours. CWR22Rv1 cells were also co-treated with 10  $\mu$ M enzalutamide. VCaP cells were seeded at a cell density of 300,000 cells per well for 24 hours before 96 hours treatment with NU7441 or DMSO plus or minus 10  $\mu$ M enzalutamide at the same concentrations as CWR22Rv1 cells. Cells were then washed twice with PBS, trypsinised, pelleted by centrifugation at 400 x g for 5 minutes and resuspended in full media and then counted using a haemocytometer. Experiments were repeated three times and each count repeated in quadruplicate.

### 3.11.2 Live cell imaging

CWR22v1-AR-EK and CWR22Rv1 were seeded at a cell density of 10,000 cells per well of a 12-well plate in full media or SDM respectively. Wells were seeded in triplicate. Appropriate compounds were added 24 hours after seeding and the plate was inserted into an Incucyte<sup>®</sup> ZOOM (Essen BioScience) live-cell analysis system. Images were taken every 4 hours for 5 days and the confluence normalised to the DMSO (NT) or siScr control.

### 3.11.3 Sulforhodamine B (SRB) growth assay

CWR22Rv1-AR-EK and CWR22Rv1 cells were seeded to a density of  $3 \times 10^3$  on a 96-well plate with the outermost wells filled with PBS to mitigate edge effect. Cells were seeded in triplicate wells for each experimental arm and allowed to adhere for 24 hours. Cells were then treated with the appropriate compounds suspended in 10  $\mu$ L of full media ( $V_f = 100 \mu$ L). At the point of treatment, time point zero cells were harvested and fixed with ice cold 50% trichloroacetic acid (TCA) (w/v) to a final concentration of 10% (v/v) and were stored at 4 °C until processing. Cells were grown for 120 hours before being fixed as above and were stored overnight at 4 °C. Plates were then washed with water and allowed to air dry. To dye the cells, 0.4% (w/v) SRB dissolved in 1% glacial acetic acid (v/v) was then added and incubated for 30 minutes at room temperature. The plates were then washed in 1% glacial acetic acid (v/v) 5 times and were allowed to air dry. Cell bound dye was then resuspended by adding 10 mM Tris-HCl pH 10.8 for 15 minutes at room temperature. Absorbance was measured at 570 nm excitation using a microplate reader (Bio-Rad). GI<sub>50</sub> values were determined using the log(inhibitor) vs normalised response analyses on Prism 8.

### 3.12 Cell cycle analysis

Cells were seeded in 6-well plates at a cell density of  $2 \times 10^5$  and allowed to adhere for 24 hours before appropriate drug treatments were added. If cells were being transfected with siRNA this was performed at the time of cell seeding. Cells were incubated for 24-96 hours before harvesting. To harvest cells, media was transferred to falcon tubes, cells were washed once in PBS and trypsinised to detach cells from the wells and then added to the appropriate falcon tube containing media. Cells were then centrifuged at 400 x g for 5 minutes and then washed in PBS before a second round of centrifugation at 400 x g for 5 minutes. Cells pellets were then resuspended in 100  $\mu$ L citrate buffer (250 mM sucrose, 40 mM sodium citrate, pH7.6) before adding 400  $\mu$ L of DNA staining buffer (20  $\mu$ g/mL propidium iodide (PI), 0.5 mM EDTA, 0.5% NP40, 10  $\mu$ g/ $\mu$ L RNase A) and then incubated at 4 °C for at least 1 hour protected from light.

Samples were briefly mixed and loaded into the Attune™ NxT flow cytometer (Invitrogen) to acquire data from at least 10,000 events. Cell debris and aggregates were removed from analysis by gating single cell populations. Histogram plots were generated using FCS Express (DeNovo Software) and percentage of cells in each cell cycle phase were quantified and compared to the control.

### 3.13 Plasmids

The AR-V7-encoding plasmid, lenti-AR-V7, was a kind gift from Scott Dehm, University of Minnesota, Minnesota, United States. The streptavidin-tagged Cas9-encoding plasmid lentiCRISPR V2-mSA was subcloned from the Cas9-encoding lentiCRISPR V2 (LCV2) plasmid (plasmid #52961, purchased from AddGene). The streptavidin-tagged tetracycline-inducible Cas9-encoding plasmid tetracycline-lentiCRISPR V2-mSA was subcloned from tetracycline-inducible Cas9-encoding plasmid tetracycline-lentiCRISPR V2 (TLCV2) (plasmid #87360, purchased from AddGene). Minigene AR-V7 reporter plasmid was a kind gift from X. Dong, University of Vancouver, Canada. RBMX-GFP-encoding plasmid was a kind gift from Professor David Elliot, Newcastle University.

### 3.10.2 Bacterial transformation

Chemically competent 5 $\alpha$  *E.coli* cells (NEB) were thawed on ice for 10 minutes. 50  $\mu$ L of the competent cell suspension was added to a chilled Eppendorf tube on ice and 50-100 ng of plasmid DNA was added. The tube was gently mixed and incubated on ice for 30 minutes. Cells were heat shocked at 42 °C on a heat block for 45 seconds and placed back on ice for 2 minutes to recover. 500  $\mu$ L SOC outgrowth medium (NEB) was then added and the cells were incubated at 37 °C with shaking at 220 rpm for 1 hour and finally spread onto pre-warmed LB agar plates (1.5% (w/v) agar, 1% (w/v) NaCl, 1% (w/v) tryptone, 0.5% (w/v) yeast extract) containing 50 mg/mL ampicillin and incubated at 37 °C overnight to allow colonies to grow. The following day, single colonies were transferred to universals containing 5 mL LB (1% (w/v) NaCl, 1% (w/v) tryptone, 0.5% (w/v) yeast extract) that contained the appropriate antibiotic and were incubated at 37 °C with shaking at 220 rpm for 6-8 hours (for mini-prep cultures bacteria was centrifuged at this step). For maxiprep, cultures were then transferred to conical flasks containing 200 mL LB containing the appropriate antibiotic and incubated at 37 °C with shaking at 220 rpm overnight. The following day, bacteria were transferred to falcon tubes and pelleted by centrifugation at 2000 x g for 5 minutes before plasmid DNA extraction. Lentiviral vectors were transformed using One Shot<sup>®</sup> Stbl3<sup>™</sup> chemically competent *E. coli* cells (Invitrogen) using the same protocol as above.

### 3.10.3 Plasmid extraction

Endotoxin-free plasmid DNA was extracted from cultured bacteria using the PureLink<sup>™</sup> HiPure Plasmid Maxiprep Kit (Thermo Fisher) according to manufacturer's instructions. Briefly, cell pellets were re-suspended in 10 mL chilled R3 buffer supplemented with RNase A and then lysed in 10 mL L7 buffer for 5 minutes at room temperature. 10 mL N3 precipitation buffer was then added before lysates were loaded onto a HiPure column that had been equilibrated by passing through 15 mL EQ1 equilibration buffer. Plasmid DNA captured in the column was then washed with 30 mL W8 wash buffer and eluted in 15 mL E4 buffer. Plasmid DNA was precipitated by adding 11.5 mL isopropanol and centrifuged at 12,000 x g for 30 minutes at 4 °C to pellet the DNA. The pellet was washed with 5 mL 70% ethanol and centrifuged at 12,000 x g for 5 minutes at 4 °C. The supernatant was removed, and the cell pellet was allowed to air dry for 10 minutes and re-suspended in 250-500  $\mu$ L nuclease free water. DNA concentration

and purity was measured using Nanodrop. Plasmid miniprep was performed using the PureLink™ HiPure Plasmid Miniprep Kit (Thermo Fisher) according to manufacturer's instructions, using a similar protocol as described above.

#### 3.10.4 Plasmid transfection

Expression plasmids were transfected into cells using TransitIT®-LT1 transfection reagent (Mirus) according to the manufacturer's instructions. Transfection mixes were made in 100 µL basal media per well of a 6-well plate containing LT1 and plasmid at a ratio of 3 µL: 1 µg. Transfection mixes were incubated for 30 minutes at room temperature and then added dropwise to the well prior to overlaying cells (reverse transfection) or directly to pre-seeded cells (forward transfection).

### 3.14 Virus production and transductions

#### 3.14.1 Virus generation

Lentivirus was generated using a 2<sup>nd</sup> generation lentiviral system incorporating separate packaging plasmids. Firstly, HEK293T cells were seeded at a cell density of  $3 \times 10^6$  cells in a 90 mm dish and allowed to adhere for 24 hours. Cells were then transfected with 2.25 µg pMD2.G (VSV-G envelope expressing plasmid) (AddGene plasmid #12259), 6.75 µg psPAX2 (viral packaging vector) (AddGene plasmid #12260) and 9 µg of the plasmid to package. 24 hours later, the culture media was replaced with 6 mL high-serum growth media (DMEM supplemented with 30% (v/v) FCS and 2 mM L-glutamate) and cells were incubated for a further 24 hours. The culture media was then harvested every 24 hours and replaced with 6 mL fresh high-serum media for 3 days. The three batches of harvested media were combined, centrifuged at 400 x g for 5 minutes and the supernatant was filtered using a sterile 0.45 µm syringe filter (Sigma Aldrich). The media was aliquoted into 1 mL cryogenic tubes and stored at -80 °C.

#### 3.14.2 Mammalian cell transduction

Cells were seeded into appropriate cell culture vessels and allowed to adhere for 24 hours. Cells were transduced with virus by adding 500 µL thawed virus-containing media per 2 mL media and cells were incubated for 24 hours before the media was replaced. Cells were then incubated for at least 72 hours to allow integration of the transgene into the genome. For

selection, 1-5 mg/mL puromycin was added to culture medium for ~2-4 weeks to allow clonal populations of plasmid integrated cells to grow before picking colonies.

### 3.15 Statistical analysis

Unless otherwise stated, all graphical data represents the mean of three individual experiments and error bars indicate  $\pm$  SEM. For analysis of DNA-PKcs inhibition on AR-mediated gene expression, ChIP, and cell viability in CWR22Rv1-AR-EK cells, one-way ANOVA was conducted, and for cell viability in CWR22Rv1 and VCaP cells a two-way ANOVA was conducted using Prism 8 software and \* =  $p < 0.05$ , \*\* =  $p < 0.01$ , \*\*\* =  $p < 0.001$  and \*\*\*\* =  $p < 0.0001$  were classified as statistically significant.

## Chapter 4 Assessing DNA-PKcs as an AR-V coregulator

### 4.1 Introduction

There are several approved therapies for early-stage PCa patients, most of which centralise around inhibiting the AR signalling axis. ADT plus surgery and/or radiotherapy are commonly used treatment options for early-stage disease. ADT inhibits AR signalling as a consequence of starving the AR of ligand by inhibiting the production of androgens and is currently the gold standard therapy for PCa. However, response to this therapy is relatively short-lived and patients go on to relapse, with tumours becoming resistant to castrate levels of circulating androgens and progressing to a more aggressive stage of disease termed CRPC. As well as chemotherapy, there are newer, second-generation hormonal therapy options available for CRPC patients, such as abiraterone, which effectively reduces the generation of adrenal and intra-prostate cell androgens to further starve the AR of ligand. Furthermore, enzalutamide is a second-generation AR antagonist that directly binds and inhibits the AR via the LBD of the full-length receptor. Although the response to AR based therapies is initially good and they significantly increase overall survival, inevitably patients become resistant after around 18 months and their cancer progresses to the more aggressive and lethal CRPC stage (Scher *et al.*, 2012). Importantly, this stage of disease is androgen independent but not AR independent as the AR remains active, driving survival and progression of the tumour. Therefore, the AR remains a critical target. There are several resistance mechanisms that lead to the development of CRPC such as AR gene amplification, AR overexpression, gain-of-function AR mutations and the generation of alternatively spliced forms of the AR, termed AR-Vs. AR-Vs are expressed in up to 70% of patients treated with ADT and are not currently targeted by any approved therapies. Inhibitors of AR-Vs are therefore a key unmet clinical need.

AR-Vs lack the LBD of the AR and are constitutively active in the absence of ligands such as DHT (Dehm and Tindall, 2011). AR-Vs retain the transcriptionally active NTD, so they remain active in castrate levels of androgens. However, the NTD is largely unstructured and therefore difficult to target directly with drugs. AR-Vs also contain the DBD but because this region is highly conserved between hormone receptors, selectivity of targeted compounds is a challenge.

Therefore, developing a better understanding of AR-V biology, in particular with relation to how they are controlled by coregulatory proteins, may reveal tractable therapeutic targets for CRPC patients who are AR-V positive. Determining the mechanisms involved in how these coregulatory proteins regulate AR-Vs will provide key avenues for the development of new treatments for CRPC patients.

DNA-PKcs is most well studied for its key role in the DDR, particularly in NHEJ. The DNA-PK complex is comprised of three proteins: Ku70, Ku80 and DNA-PKcs. The Ku proteins are responsible for sensing DNA DSBs and the recruitment of DNA-PKcs to the damaged DNA locus. This DNA-PK complex assembled at the DNA break site is then responsible for ligation of the DNA strands, with the help from other DDR proteins (Davis and Chen, 2013). Aside from its role in DDR, DNA-PKcs has been shown to have a number of other roles in various cellular pathways including cell cycle progression and transcriptional regulation (Goodwin *et al.*, 2015, Lee *et al.*, 2011). Interestingly, DNA-PKcs was originally characterised as a key regulator of the transcription factor SP1 in 1990. It was shown to phosphorylate SP1 upon its binding to *cis*-regulatory elements of target genes (Jackson *et al.*, 1990). Shortly after, it was then shown to also regulate and phosphorylate RNA polymerase II during transcription (Dvir *et al.*, 1992). Since those initial discoveries, DNA-PKcs has been shown to be involved in the regulation of a number of transcription factors including the GR (Giffin *et al.*, 1997) and c-Myc (An *et al.*, 2008), and importantly, the AR (Goodwin *et al.*, 2015). The Ku proteins of the DNA-PK holoenzyme complex were also identified as coregulators of the AR, through direct interaction with the LBD of the FL-AR; determined using tandem mass spectrometry (MS) (Mayeur *et al.*, 2005). DNA-PKcs has been shown to co-activate the AR, as well as reciprocally, AR activity increases DNA-PKcs phosphorylation and function (Goodwin *et al.*, 2013). More recent work by Dylgjeri *et al.*, using RNA-sequencing in two models of CRPC, C4-2 and CWR22Rv1 cells, confirmed DNA-PK involvement in numerous pathways, including the androgen response, DNA repair and cell cycle (Dylgjeri *et al.*, 2019). The group also identified DNA-PKcs involvement in additional pathways, such as epithelial–mesenchymal transition and oxidative phosphorylation; with the study concluding that the transcriptional events caused by DNA-PKcs inhibition enriches oncogenic pathways that could be therapeutically exploited.

DNA-PKcs expression and its association with long-term clinical outcomes in PCa patient samples was investigated by Kothari *et al.* DNA-PKcs was identified as the most significant kinase associated with PCa progression and metastasis and that it associated with a decreased overall survival (Kothari *et al.*, 2019). They also reported that Wnt signalling was negatively enriched upon DNA-PKcs siRNA depletion in four PCa cell lines and multiple clinical samples, as determined using microarray. It was suggested that this phenomenon could potentially prevent the development of ADT resistance and PCa progression, as Wnt signalling was found to be associated with the more aggressive PCa models.

DNA-PKcs has also been shown to play a role in other hormone-driven cancers, such as breast cancer, through its interaction with the oestrogen receptor, and inhibition of DNA-PKcs lead to increased oestrogen receptor ubiquitination and subsequent proteasomal degradation (Medunjanin *et al.*, 2010). Furthermore, DNA-PKcs inhibition in three breast cancer cell lines increased the sensitivity to irradiation and doxorubicin treatments, including the triple negative breast cancer cell line MDA-MB-231 that has fewer treatment options (Ciszewski *et al.*, 2014).

Due to its multifunctional nature and implication in several oncogenic pathways, DNA-PKcs blockade represents a tractable therapeutic strategy in cancer patients, both as a single agent and in combination with radio- and chemo-therapy.

There have been numerous DNA-PKcs inhibitors developed including NU7441, NU5455, AZD7648 and M3814. (Willoughby *et al.*, 2020, Leahy *et al.*, 2004, Fok *et al.*, 2019, Zenke *et al.*, 2020). More recently, some of these inhibitors have made it to clinical trials for investigation into their efficacy as anticancer treatments. M3814 is being trialled in combination with capecitabine and radiotherapy in patients with rectal cancer (NCT03770689); and in combination with radiotherapy in patients with localised pancreatic cancer (NCT04172532) and head and neck cancer (NCT04533750). Recently, published phase 1 clinical trial data indicated M3814 was well-tolerated and demonstrated modest efficacy (van Bussel *et al.*, 2021). The compound AZD7648 has also made it to early-stage clinical trials. Pre-clinically, it has been used as a combination treatment with either IR or doxorubicin and the PARP inhibitor Olaparib where it was shown to sensitise xenograft and patient derived xenografts to radio- and chemo-therapy with sustained regressions (Fok *et al.*, 2019).

AZD7648 has progressed to phase I clinical trials in advanced malignancies and has been administered as a monotherapy or in combination with Pegylated Liposomal Doxorubicin (NCT03907969). This trial is ongoing and due to be completed in 2024. There are also ongoing trials assessing the efficacy of dual DNA-PKcs/mTOR inhibitors. The dual inhibitor CC-115 has been shown *in vitro* to be more effective than targeting DNA-PKcs alone with NU7441 in multiple PCa and CRPC cell lines. CC-115 was also tested in combination with enzalutamide, to mitigate an upregulation of androgen response due to mTOR inhibition, both *in vitro* and *in vivo*. Hormone sensitive and castration-resistant cell line models showed synergistic effects between CC-115 and enzalutamide both *in vitro* and *in vivo*; with reduced tumour doubling time with the combination treatment (Dylgjeri *et al.*, 2019). This compound was well tolerated in phase 1 clinical trials (NCT01353625) in patients with advanced solid or hematologic malignancies (Munster *et al.*, 2019).

Despite FL-AR-DNA-PKcs interplay being relatively well defined, less is known about the AR-V-DNA-PKcs relationship. There have been two reports of an interaction between AR-Vs and DNA-PKcs. Firstly, an AR co-immunoprecipitation (co-IP) showed an interaction between AR-V7 and DNA-PKcs in the CWR22Rv1 cell line (Goodwin *et al.*, 2015). More recently, IP-MS analysis identified DNA-PKcs as the most abundant binder of the AR-V, ARV<sup>567es</sup>, in the cell line R1-D567, which was validated using co-IP and proximity ligation assays (Yin *et al.*, 2017). This interaction was enhanced in the presence of IR-induced DNA damage and could be reduced by the DNA-PKcs inhibitor NU7441. However, there has been no reports describing a specific role for DNA-PKcs in AR-V transcriptional regulation. Dr Dominic Jones from the host laboratory, using an AR-V7 RIME experiment in CWR22Rv1 cells, identified DNA-PKcs as an AR-V7 interacting protein. Briefly, Flag-tagged AR-V7 was overexpressed for 48 hours in CWR22Rv1 cells grown in steroid-depleted conditions, prior to formaldehyde crosslinking, chromatin preparation and FLAG antibody IP. The resultant AR-V7 interactome was analysed by MS, and DNA-PKcs was identified as an AR-V7 interacting protein. Interestingly, DNA-PKcs was also identified in a Cas9-directed-RIME (CRIME) experiment carried out by Dr Evangelia Kounatidou (previous PhD student in the host laboratory), in the CWR22Rv1-AR-EK cell line, as a protein associated with the AR CE3 genomic locus. These studies have therefore validated the AR-V-DNA-PKcs interaction and may suggest DNA-PKcs involvement in AR-V biology is

multifaceted. Further exploration into the interplay between DNA-PKcs and AR-Vs is necessary to help understand if DNA-PKcs is a key transcriptional regulator of AR-Vs, and as such represents a suitable target to diminish the activity of these therapy-resistant AR isoforms. As DNA-PK is also involved heavily in the DDR, targeting it in PCa could have dual anti-tumour effects, both directly targeting AR/AR-V signalling as well as increasing the effect of IR and other DNA damaging agents.

#### 4.1.1 Aims

In order to develop treatments that can target AR-Vs, we must better understand their regulatory mechanisms. DNA-PKcs is a well characterised FL-AR co-regulator; facilitating AR transcriptional activity by enabling recruitment of the transcriptional machinery to canonical AR-target genes. Moreover, an interaction between DNA-PKcs and AR-Vs has been reported (Yin *et al.*, 2017) and validated in-house, but there remains major knowledge gaps in whether DNA-PKcs is a key co-regulator of AR-Vs. This project, therefore, is to determine if DNA-PKcs is involved in the regulation of AR-Vs.

Using several AR-V expressing PCa cell lines, this will be done by:

- i. Pharmacological DNA-PKcs inhibition, incorporating three DNA-PKcs inhibitors, and DNA-PKcs depletion to determine the impact on AR-V transcriptional activity
- ii. Determining if DNA-PKcs is recruited to the *cis*-regulatory elements of AR-V target genes and if this is impacted upon DNA-PKcs inhibition
- iii. Examining the effects of compromised DNA-PKcs activity on AR-V-driven cell proliferation using DNA-PKcs inhibitors and depletion

## 4.2 Results

### 4.2.1 DNA-PKcs expression is significantly upregulated in PCa and CRPC

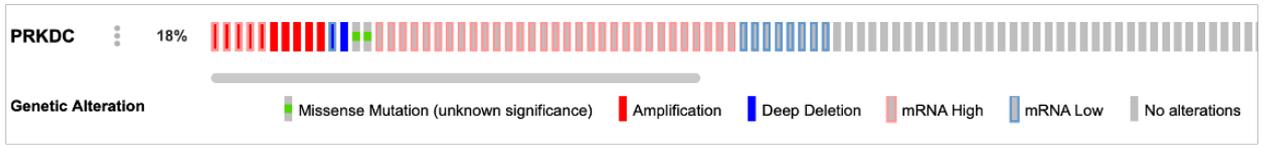
Previous reports have indicated the DNA-PKcs-encoding gene *PRKDC* is significantly upregulated in advanced PCa and expression is correlated with metastatic progression and decreased overall survival (Kothari *et al.*, 2019). To validate this, data from the online database, cBioPortal, was analysed to determine the frequency of *PRKDC* alterations in prostate adenocarcinoma (The Cancer Genome Atlas (TCGA) dataset) and metastatic prostate adenocarcinoma (Stand-up to cancer/ Prostate Cancer Foundation (SU2C/PCF dataset, PNAS, 2019). The DNA-PKcs-encoding *PRKDC* gene is mutated, amplified, or differentially expressed at the mRNA level (1.5-fold change (FC) cut-off) in a large portion of PCa patients; both in the TCGA and SU2C/PCF datasets. As shown in Figure 4.1A, in the TCGA data set, 18% of patients had a genetic aberration in the *PRKDC* gene or demonstrated differential gene expression, by either a missense mutation, amplification, deep deletion (homozygous loss), mRNA high or mRNA low, with the largest proportion being 'mRNA high'. Furthermore, in this cohort of patients (altered group), *PRKDC* alterations were statistically significantly correlated with AR mRNA (Figure 4.1B,  $p = 2.45 \times 10^{-7}$ ) using a Wilcoxon test. The altered group of patients with one or more genetic alteration in *PRKDC* also correlated with a higher Gleason category (Figure 4.1C,  $p = 3.155 \times 10^{-5}$ ). Co-expression of *PRKDC* mRNA and AR mRNA is also positively correlated with a Spearman correlation coefficient of 0.67 and is highly statistically significant (Figure 4.1D,  $p = 1.98 \times 10^{-39}$ ). Taken together, this suggests DNA-PKcs could be a driver of PCa progression and could be used as a prognostic biomarker.

In the SU2C dataset that includes patients with metastatic PCa (Figure 4.1E), a greater proportion of patients (27%) had a genetic alteration in the *PRKDC* gene, or 'high mRNA', with the majority of these patients having high DNA-PKcs mRNA. This could suggest as PCa progresses, there is a greater dependency on DNA-PKcs, hence metastatic patients demonstrate a higher frequency of *PRKDC* alterations. This is consistent with what was shown in the study by Goodwin *et al.*, where they identified DNA-PKcs as a driver of pro-metastatic signalling *in vitro* and that inhibition of DNA-PKcs kinase activity with the inhibitor NU7441 delays the development of metastases *in vivo* (Goodwin *et al.*, 2015).

### A. Prostate Adenocarcinoma (TCGA, Cell 2015)

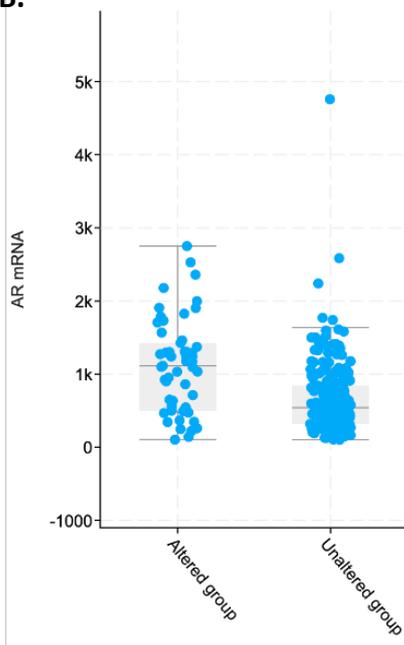
Complete samples (290 patients/samples)

Queried gene is altered in 53 (18%) of queried patients/samples

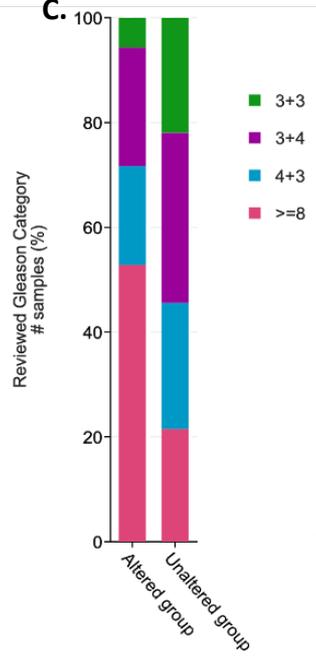


Clinical Attribute	Statistical Test	p-Value	q-Value
AR mRNA	Wilcoxon Test	2.45e-7	2.012e-5
Reviewed Gleason Category	Chi-squared Test	3.155e-5	6.468e-4

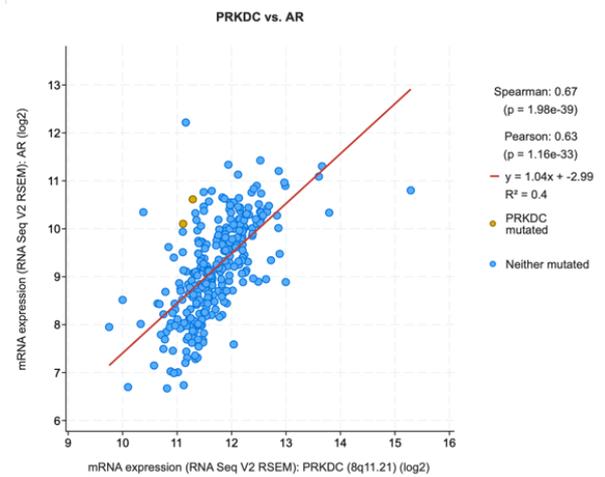
B.



C.



D.



### E. Metastatic Prostate Adenocarcinoma (SU2C/PCF Dream Team, PNAS 2019)

All Tumors (429 patients / 444 samples)

Queried gene is altered in 115 (27%) of queried patients

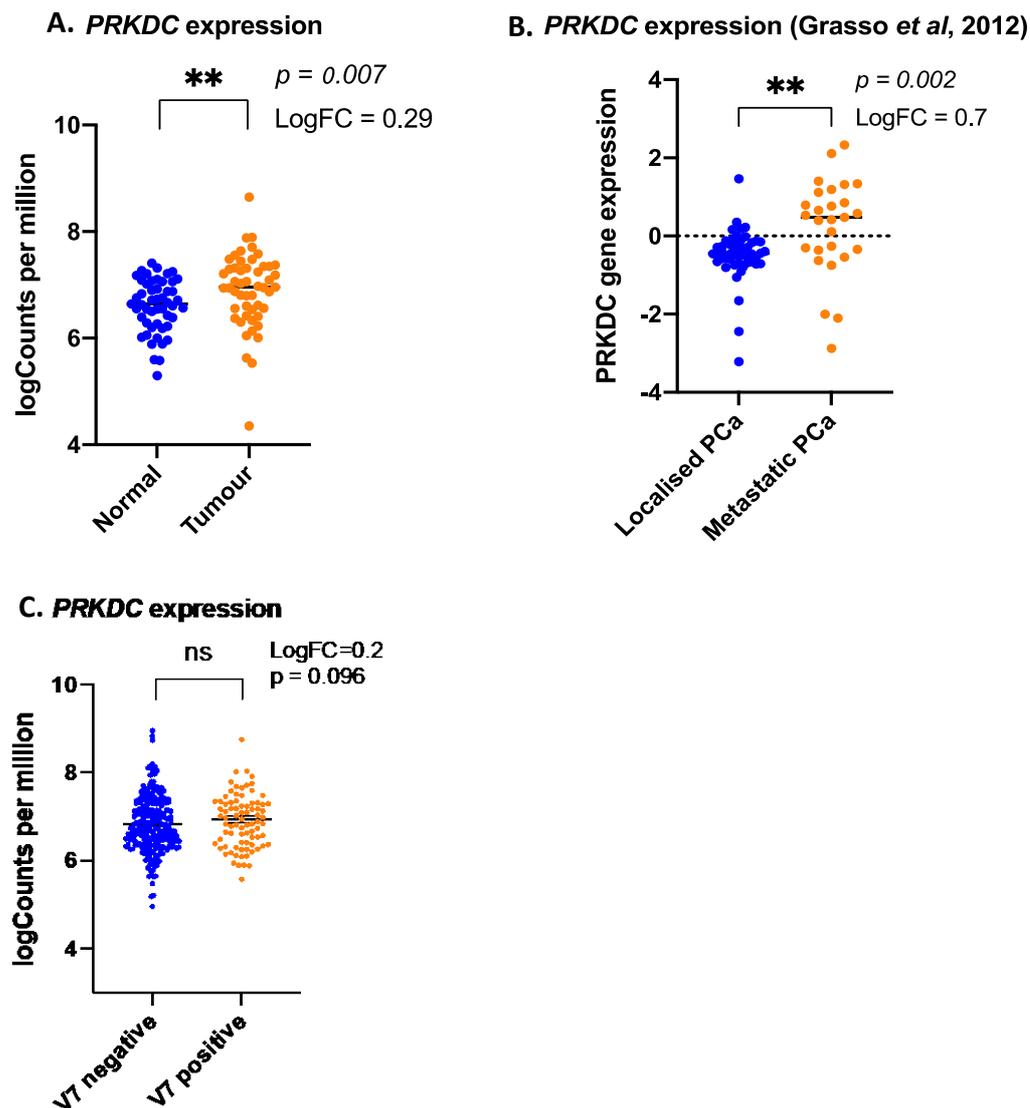


**Figure 4.1** *PRKDC* is altered in a portion of prostate adenocarcinoma and metastatic prostate adenocarcinoma patients.

cBioPortal was used to interrogate the portion of patients that have genetic alterations in the *PRKDC* gene. A-D. Analysis of the TCGA, *Cell*, 2015 (Cancer Genome Atlas Research, 2015) prostate adenocarcinoma dataset showed 18% of patient had alterations in the *PRKDC* gene and/or differential mRNA expression. Clinical attributes that correlated with DNA-PKcs alterations included AR mRNA and Reviewed Gleason category. E. Analysis of the SU2C/PCF, *PNAS*, 2019 (Abida *et al.*, 2019) metastatic

prostate adenocarcinoma dataset showed 27% of patients had alterations in the *PRKDC* gene and/or differential mRNA expression. Figures taken from cBioPortal (Gao et al., 2013, Cerami et al., 2012).

The TCGA dataset was analysed in house by fellow PhD student Nicholas Brittain, to determine if *PRKDC* expression is significantly differentially expressed in tumour versus normal samples, matched from the same patient. As shown in Figure 4.2A, *PRKDC* is significantly ( $p = 0.007$ ) upregulated by 1.22-fold in tumour samples compared to normal. *PRKDC* expression was also compared between localised and metastatic PCa using the Grasso *et al.*, microarray dataset (GSE35988), using Gene Expression Omnibus 2R online tool (Grasso *et al.*, 2012). This showed *PRKDC* expression was significantly ( $p = 0.002$ ) upregulated by 1.6-fold in metastatic PCa versus localised PCa (Figure 4.2B). To investigate if DNA-PKcs is involved in AR-V7 regulation, changes in *PRKDC* expression (log<sub>2</sub> counts per million) was determined in AR-V7 negative versus positive samples. This showed *PRKDC* expression is increased by 1.16-fold in AR-V7 positive samples ( $p = 0.096$ ) compared to patients with no AR-V7 expression (Figure 4.2C). This suggests some correlation between DNA-PKcs expression and AR-V7, although not statistically significant.



**Figure 4.2 PRKDC expression is significantly upregulated in tumour samples compared to normal and in metastatic PCa compared to localised.**

A. The TCGA dataset was analysed to compare *PRKDC* expression in matched normal vs tumour samples,  $n=51$ . (\*\* =  $p < 0.05$ ) B. *PRKDC* expression was measured in localised ( $n=49$ ) and metastatic ( $n=27$ ) PCa from a publicly available Grasso microarray dataset (\*\* =  $p < 0.05$ ) (Grasso *et al.*, 2012). C. *PRKDC* expression was measured in AR-V7 positive ( $n=83$ ) and AR-V7 negative ( $n=237$ ) samples using the TCGA dataset. Data sets A and C were generated by Nicholas Brittain using TCGA data.

#### 4.2.2 DNA-PKcs inhibition reduces proliferation of CRPC cell lines

As previously discussed, DNA-PKcs is a co-regulator of FL-AR activity (Goodwin *et al.*, 2015). Consistent with these findings, we hypothesised that DNA-PKcs would facilitate the activity of AR-Vs. In order to determine if AR-Vs are dependent on DNA-PKcs for transcriptional activity, models that express AR-Vs, and therefore represent CRPC, were used. CWR22Rv1 cells express

both FL-AR and AR-Vs, but when cultured in castrate, steroid-depleted conditions supplemented with enzalutamide, FL-AR activity is largely diminished enabling selective interrogation of AR-Vs. However, some FL-AR activity may remain meaning readouts to interrogate AR-V activity could be compromised by residual FL-AR activity. Therefore, the host laboratory developed a CRISPR engineered cell line termed CWR22Rv1-AR-EK. This cell line has a translational stop codon inserted in exon 5 of the *AR* gene to essentially stop the translation of FL-AR whilst maintaining expression of all other endogenously expressed AR-Vs. This provides a much cleaner readout for investigation of specifically AR-Vs, without the need to overexpress AR-Vs in other cellular backgrounds that are not physiologically relevant. The VCaP cell line was also used in this study as FL-AR is highly expressed, due to *AR* gene amplification, and AR-V expression is up-regulated in steroid-depleted conditions. Again, enzalutamide can also be used in this model as a way of abrogating FL-AR activity to more selectively interrogate AR-V biology. All of the cell lines used in this study expresses a number of AR-Vs that are clinically relevant; having been shown to be expressed in PCa patients, including AR-V7 that is identified in CTCs and is associated with a worse overall survival (Scher *et al.*, 2018, Sharp *et al.*, 2019).

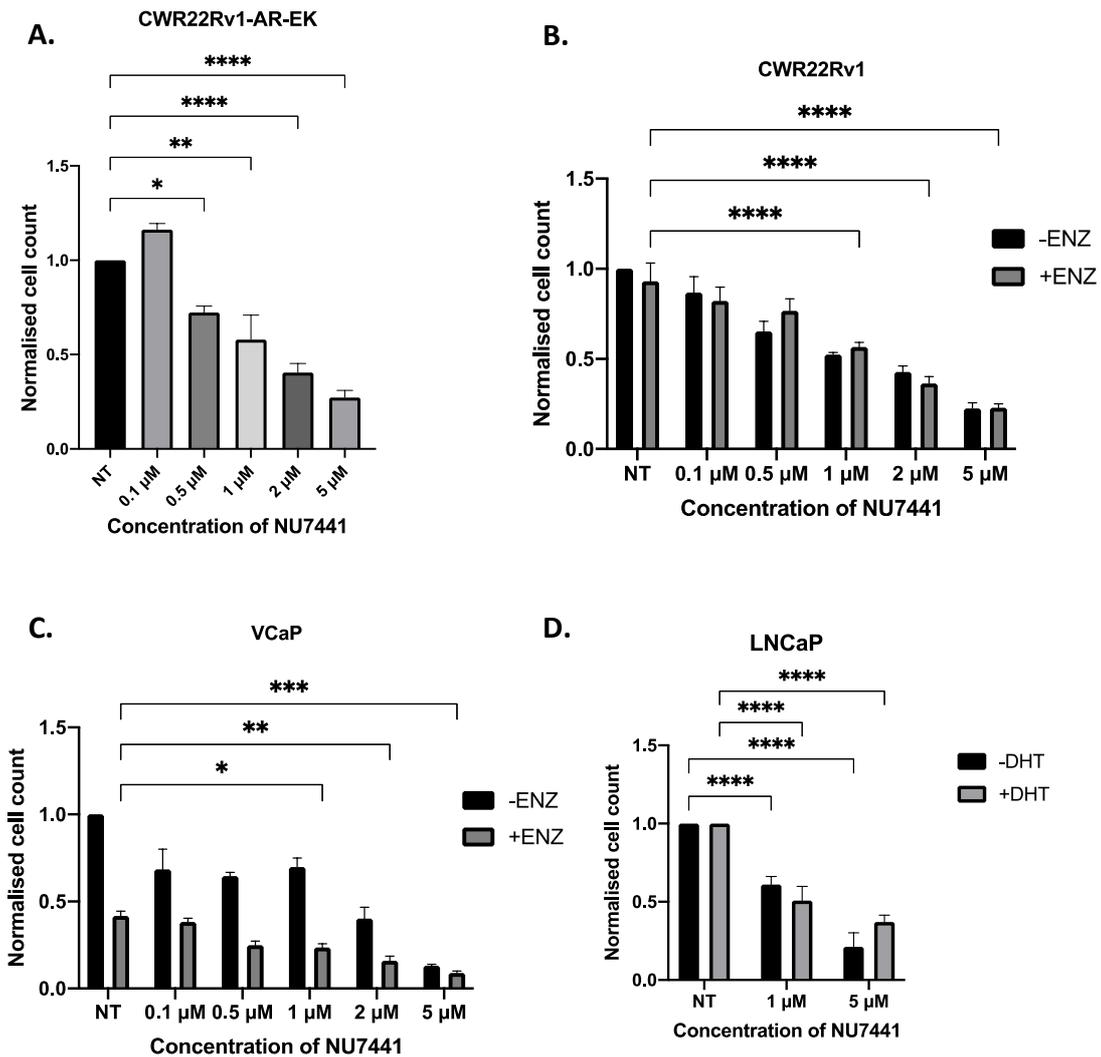
The compound NU7441 was initially used to study cellular responses to DNA-PKcs kinase inhibition in the aforementioned CRPC cell lines because it is a reasonably selective agent that has been used to interrogate DNA-PKcs activity in numerous previously published studies, including those investigating interplay between FL-AR and DNA-PKcs. This provides data that can be cross-referenced to compare the effects in different analyses.

The phenotypic effects of DNA-PKcs inhibition with NU7441 was investigated in CWR22Rv1-AR-EK, CWR22Rv1 and VCaP cell lines, as well as in the FL-AR-only expressing LNCaP cell line (Figure 4.3). Cell counts were carried out after 96-120 hours treatment with increasing concentrations of NU7441 or DMSO vehicle control (NT). A 73% decrease in cell number at 5  $\mu$ M and an estimated  $GI_{50}$  between 1  $\mu$ M and 2  $\mu$ M was observed in the CWR22Rv1-AR-EK cell line (Figure 4.3A). Consistent with this, CWR22Rv1 cells demonstrated a 77% decrease in cell proliferation in response to DNA-PKcs inhibitor treatment in both the presence and absence of enzalutamide, with a similar  $GI_{50}$  between 1  $\mu$ M and 2  $\mu$ M. As expected, there was no significant difference between the plus and minus enzalutamide arms in the absence of

NU7441 in CWR22Rv1 cells due to AR-V-driven proliferation being refractory to enzalutamide (Figure 4.3B).

In VCaP cells, the presence of enzalutamide alone in the DMSO treatment arm caused a 59% reduction in cell proliferation (Figure 4.3C). This was expected due to the greater reliance on FL-AR signalling in the VCaP cell line compared to CWR22Rv1 cells and hence have an increased sensitivity to anti-androgens after a short period of androgen depletion. Critically, there were respective 78.5% and 87% decreases in cell proliferation in the absence and presence of enzalutamide in response to 5  $\mu$ M NU7441 treatment versus control providing further evidence that DNA-PKcs potentially impacts AR-V pro-proliferative signalling and inhibition could provide synergistic effects with anti-androgen treatment.

Finally, the proliferative impact of DNA-PKcs inhibition in LNCaP cells was determined in both androgenic and castrate conditions to compare to published data (Goodwin *et al.*, 2015, Kothari *et al.*, 2019). It was previously reported that in SDM, NU7441 significantly reduced LNCaP proliferation after 6 days. Consistent with this data, Figure 4.3D shows NU7441 significantly reduces proliferation by approximately 40% and 50% in the absence and presence of DHT, respectively at 1  $\mu$ M NU7441; and 80% and 60% minus and plus DHT at 5  $\mu$ M, respectively. These findings are also consistent with DNA-PKcs knockdown data in LNCaP cells which showed that DNA-PKcs depletion caused a significant reduction in proliferation (Kothari *et al.*, 2019).

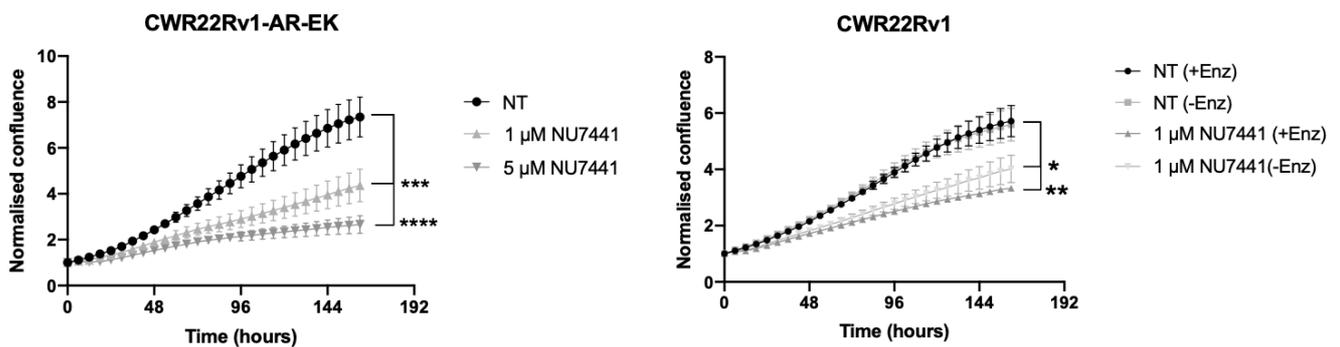


**Figure 4.3** NU7441 significantly effects proliferation of CWR22Rv1-AR-EK, CWR22Rv1, VCaP and LNCaP cells.

A. CWR22Rv1-AR-EK cells were grown in full media and treated with increasing concentrations of NU7441 for 96 hours. B. CWR22Rv1 cells were grown in steroid-depleted conditions in the presence and absence of 10 μM enzalutamide and NU7441 for 96 hours. Data is normalised to the control arm of the -ENZ group. C. VCaP cells were grown in steroid-depleted conditions in the presence and absence of 10 μM enzalutamide and treated with NU7441 for 96 hours. D. LNCaP cells were seeded and allowed to adhere before the media was replaced with steroid- depleted media for 24 hours, then treated with NU7441 ± 10 nM DHT for 5 days. Data is normalised to the control arm of the -ENZ group. Data is representative of three independent repeats ± SEM. One-way ANOVA using Bonferroni Post Hoc analysis was used to determine the statistical significance for CWR22Rv1-AR-EK and 2-way ANOVA was used for CWR22Rv1, LNCaP and VCaP cells (\* =  $p < 0.05$ , \*\* =  $p < 0.01$ , \*\*\* =  $p < 0.001$ ).

The cell count data was validated using live cell imaging of CWR22Rv1-AR-EK and CWR22Rv1 cells, grown in full media and SDM, respectively (Figure 4.4). Cells were seeded down in 12-well plates 24 hours before the addition of the DNA-PKcs inhibitor NU7441 and 10 μM

Enzalutamide (CWR22Rv1 only) to culture media. Cell confluency was then determined every 4 hours over a period of 7 days using an Incucyte® ZOOM live cell imaging system. Data represents mean cell confluency of triplicate wells. The results are consistent with what was shown in the cell count data, with DNA-PKcs inhibition causing a significant reduction in proliferation in at 1  $\mu$ M (CWR22Rv1-AR-EK and CWR22Rv1) and 5  $\mu$ M (CWR22Rv1-AR-EK). The addition of enzalutamide in CWR22Rv1 cells caused a slight additive anti-proliferative effect although this was not significant.



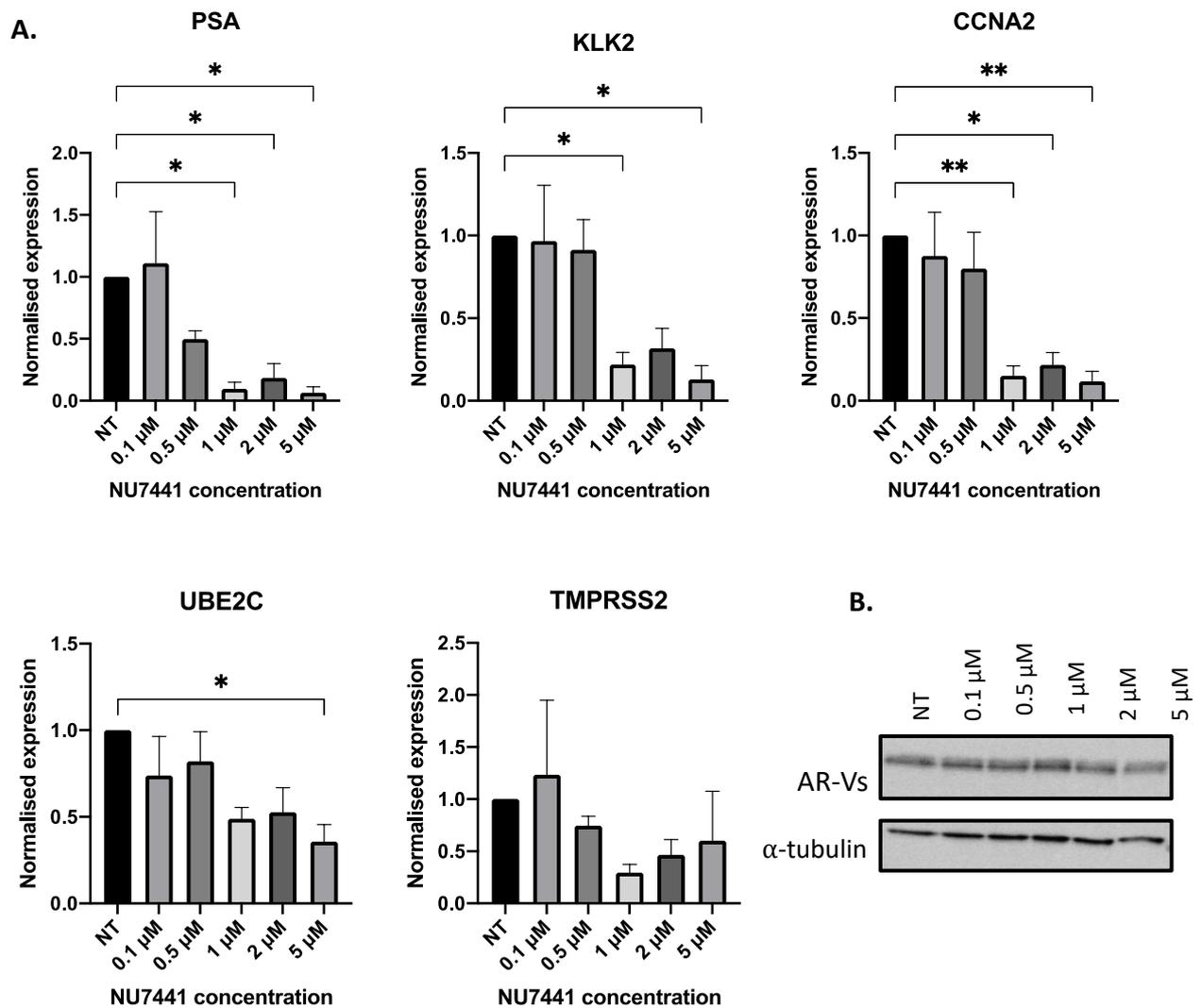
**Figure 4.4** 1  $\mu$ M NU7441 significantly reduces PCa cell proliferation in CWR22Rv1-AR-EK and CWR22Rv1 cells.

CWR22Rv1-AR-EK and CWR22Rv1 confluency over 7 days treatment with NU7441 using live cell imaging. Figures are representative of 3 repeats  $\pm$  SEM and each data point from each inter-experimental repeat are an average of three intra-experimental repeats. One-way ANOVA using Bonferroni Post Hoc analysis was used to determine the statistical significance of CWR22Rv1-AR-EK cells and two-way ANOVA was used for CWR22Rv1 cells (\* =  $p < 0.05$ , \*\* =  $p < 0.01$ , \*\*\* =  $p < 0.001$ ).

#### 4.2.3 DNA-PKcs inhibition reduces AR target gene expression

To determine the effect of DNA-PKcs inhibition on AR-/AR-V-mediated transcription, CWR22Rv1, CWR22Rv1-AR-EK and VCaP cells were cultured in SDM or full media (CWR22Rv1-AR-EK only) and then subject to either co-treatment with an increasing dose-range of NU7441 and 10  $\mu$ M enzalutamide (CWR22Rv1 and VCaP) or NU7441 alone (CWR22Rv1-AR-EK) for 24 hours. Resultant cDNA samples were subjected to qRT-PCR to assess the expression of the canonical AR-regulated genes, *PSA*, *KLK2* and *TMPRSS2* as well as the AR-V only regulated genes *CCNA2* and *UBE2C* in response to drug treatment. To reiterate, enzalutamide was added to CWR22Rv1 and VCaP cells to inactivate FL-AR which enables more robust analysis of the effects of DNA-PKcs inhibition on AR-V transcriptional competency.

DNA-PKcs inhibition caused a significant reduction in *PSA*, *KLK2*, *UBE2C* and *CCNA2* expression at 1-5  $\mu\text{M}$  NU7441 in CWR22Rv1-AR-EK cells (Figure 4.5A). At the 1  $\mu\text{M}$  dose, DNA-PKcs inhibition using NU7441 reduced the expression of several key AR-V regulated genes in this cell line, suggesting DNA-PKcs activity is an important co-regulator of AR-Vs. As seen in the western blot image in Figure 4.5B, AR-V protein expression was not impacted after 24-hour treatment arms at the 0.1 – 2  $\mu\text{M}$  treatment arms but is slightly reduced at the 5  $\mu\text{M}$  NU7441 dose, suggesting that the effect of DNA-PKcs inhibition was largely at the RNA level.



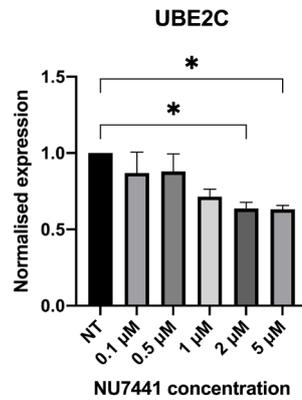
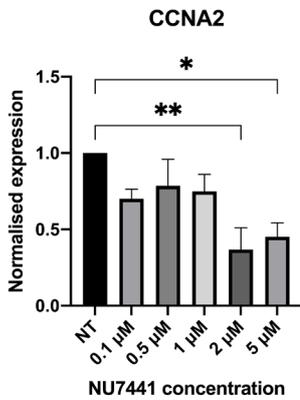
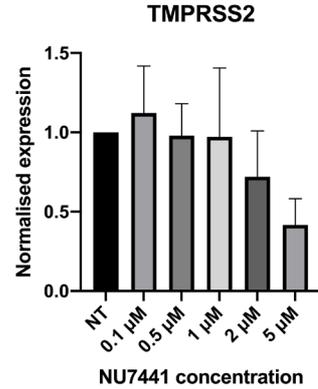
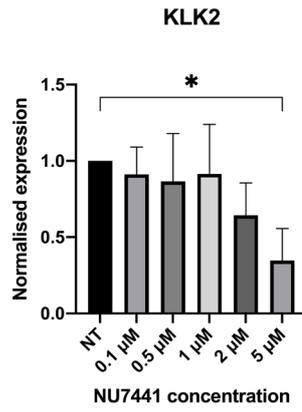
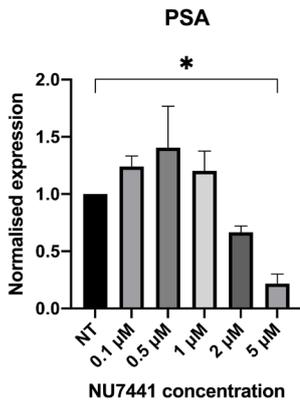
**Figure 4.5 DNA-PKcs inhibition reduces AR-V mediated transcription in CWR22RV1-AR-EK cells.**

CWR22Rv1-AR-EK cells were cultured in full media for 48 hours and then treated with increasing concentrations of NU7441 for 24 hours. A. Cells were then subject to AR target gene expression analysis using qRT-PCR. Data was normalised to the DMSO treatment arm for each target gene. Data is representative of three independent repeats  $\pm$  SEM. One-way ANOVA using Bonferroni Post Hoc analysis was used to determine the statistical significance (\* =  $p < 0.05$ , \*\* =  $p < 0.01$ ) B. In parallel,

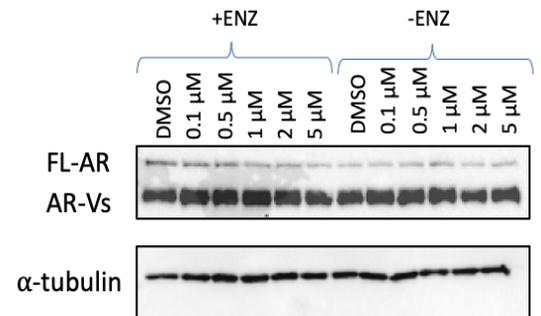
NU7441-treated cells were also subject to western blot analysis of AR-V protein levels using an N-terminal AR-binding antibody and  $\alpha$ -tubulin as loading control.

To determine if the effects observed in CWR22Rv1-AR-EK cells could be replicated in the parental cell line, CWR22Rv1 cells were cultured in steroid-depleted media and then treated with NU7441 plus or minus enzalutamide for 24 hours before downstream analysis of the aforementioned AR target genes. In the plus enzalutamide arm, which, in principle, should allow only AR-V activity to be investigated, *PSA*, *KLK2* and *CCNA2* gene expression was reduced, although it was not statistically significant at the lower concentration of 1  $\mu$ M, that had been sufficient to robustly reduce AR-V targets in CWR22Rv1-AR-EK cells (Figure 4.6A). Although AR-V activity will dominate in these conditions, the effect of DNA-PKcs inhibition on AR target gene expression is not as significant as seen in the FL-AR knockout CWR22Rv1-AR-EK cell line. This may suggest that retention of low levels of FL-AR activity in parental CWR22Rv1 cells treated with enzalutamide or differences in AR-V activity between the two cell lines may give rise to the subtle difference in response to DNA-PKcs inhibition across the two cell line derivatives. The experiment was not conducted in CWR22Rv1-AR-EK cells that were co-treated with enzalutamide, which may be a better comparator to the CWR22Rv1 plus enzalutamide experiment. In the cells treated without enzalutamide (Figure 4.6B), there was also a subtle reduction in AR target gene expression in response to DNA-PKcs inhibition, particularly at the higher concentration of 5  $\mu$ M and the AR-V target gene *UBE2C*. In this treatment arm both FL-AR and AR-V response can be investigated, although there are higher levels of AR-Vs in comparison to FL-AR due to the cells being cultured in SDM.

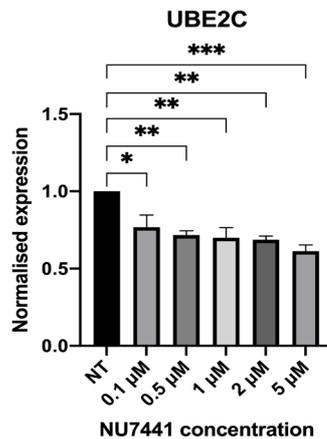
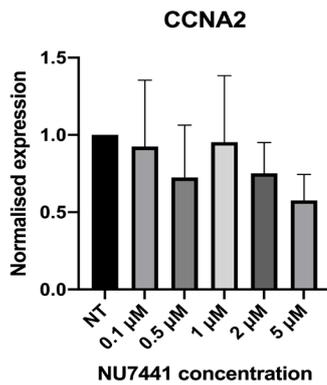
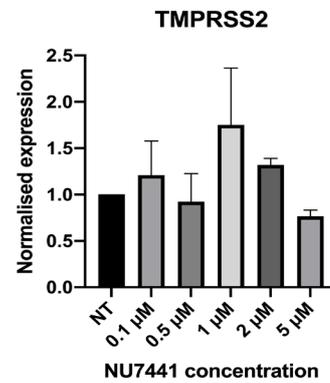
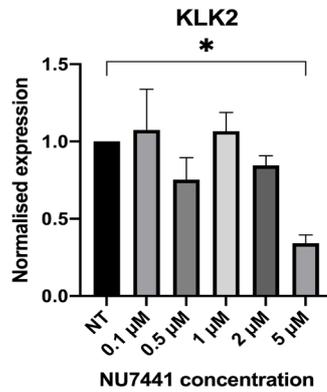
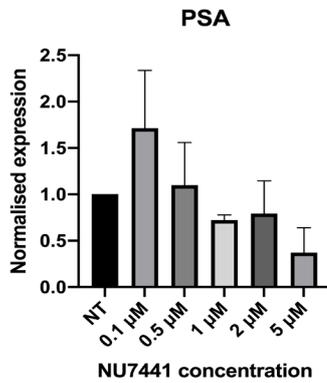
**A.** CWR22Rv1  
+ Enzalutamide



**CWR22Rv1: 24h NU7441 +/- Enz**



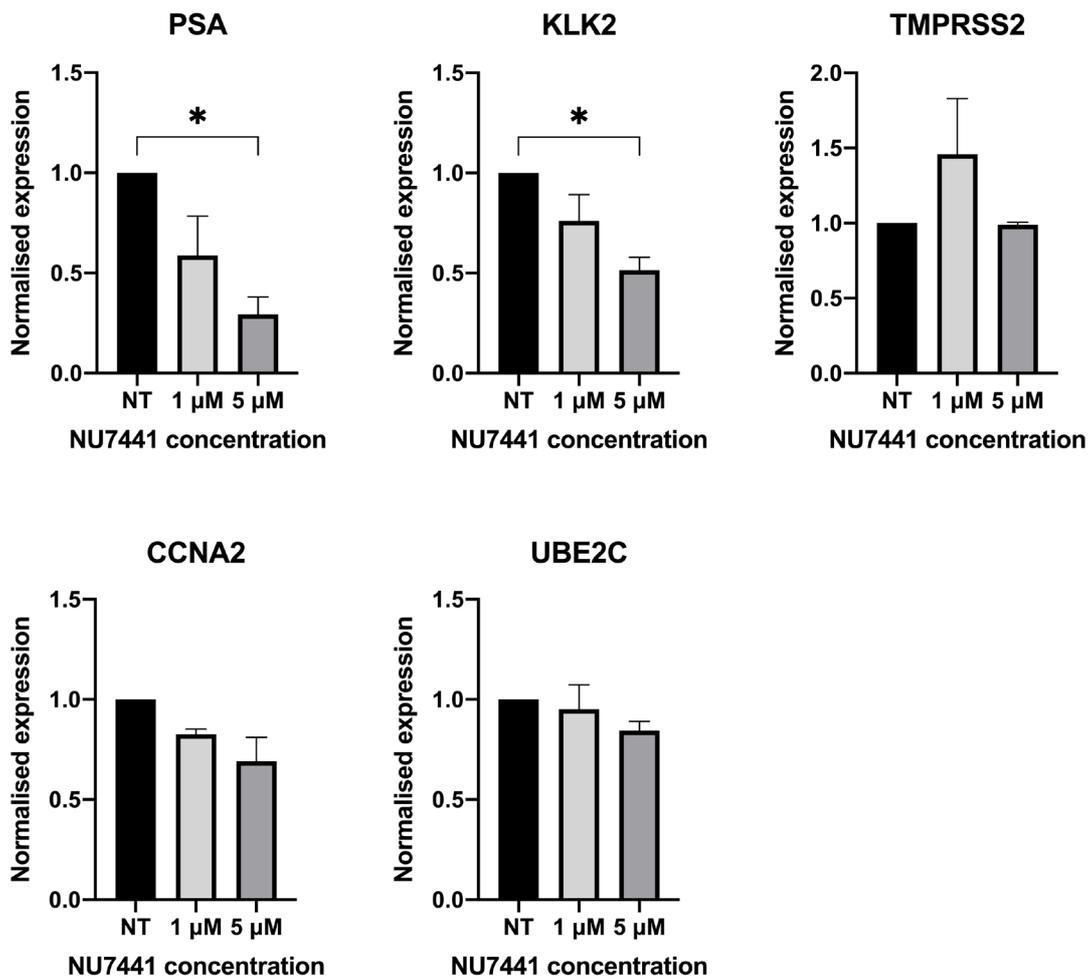
**B.** CWR22Rv1  
- Enzalutamide



**Figure 4.6 DNA-PK inhibition reduces transcription of the AR-V target gene *UBE2C* in CWR22Rv1 cells after 24 hours with enzalutamide treatment.**

CWR22Rv1 cells were cultured in steroid-depleted media for 48 hours and then treated with 10  $\mu$ M enzalutamide for 24 hours. Cells were then subject to AR target gene expression analysis using qRT-PCR. Data was normalised to the NT (DMSO control) arm for each target gene. Data is representative of three independent repeats  $\pm$  SEM. One-way ANOVA using Bonferroni Post Hoc analysis was used to determine statistical significance (\* =  $p < 0.05$ , \*\* =  $p < 0.01$ , \*\*\* =  $p < 0.001$ ). B. Cells were also subject to western blot analysis of AR-V, FL-AR and  $\alpha$ -tubulin protein levels.

Due to a more modest reduction of AR target gene expression in the CWR22Rv1 cell line in comparison to CWR22Rv1-AR-EK cells after DNA-PKcs inhibition, the experiment was repeated with a 24-hour pre-treatment with 10  $\mu$ M enzalutamide, to determine if a longer inactivation of FL-AR before NU7441 treatment effected the AR-V-mediated transcriptional response. This led to a significant reduction of the AR target genes *PSA* and *KLK2* (Figure 4.7), although this was not as pronounced as the previous experiment (Figure 4.6). The AR-V specific target genes *UBE2C* and *CCNA2* did not show any significant change in expression in response to both 1  $\mu$ M and 5  $\mu$ M of the compound and the effect on AR target gene expression was still not as robust as that was seen in the CWR22Rv1-AR-EK cell line.



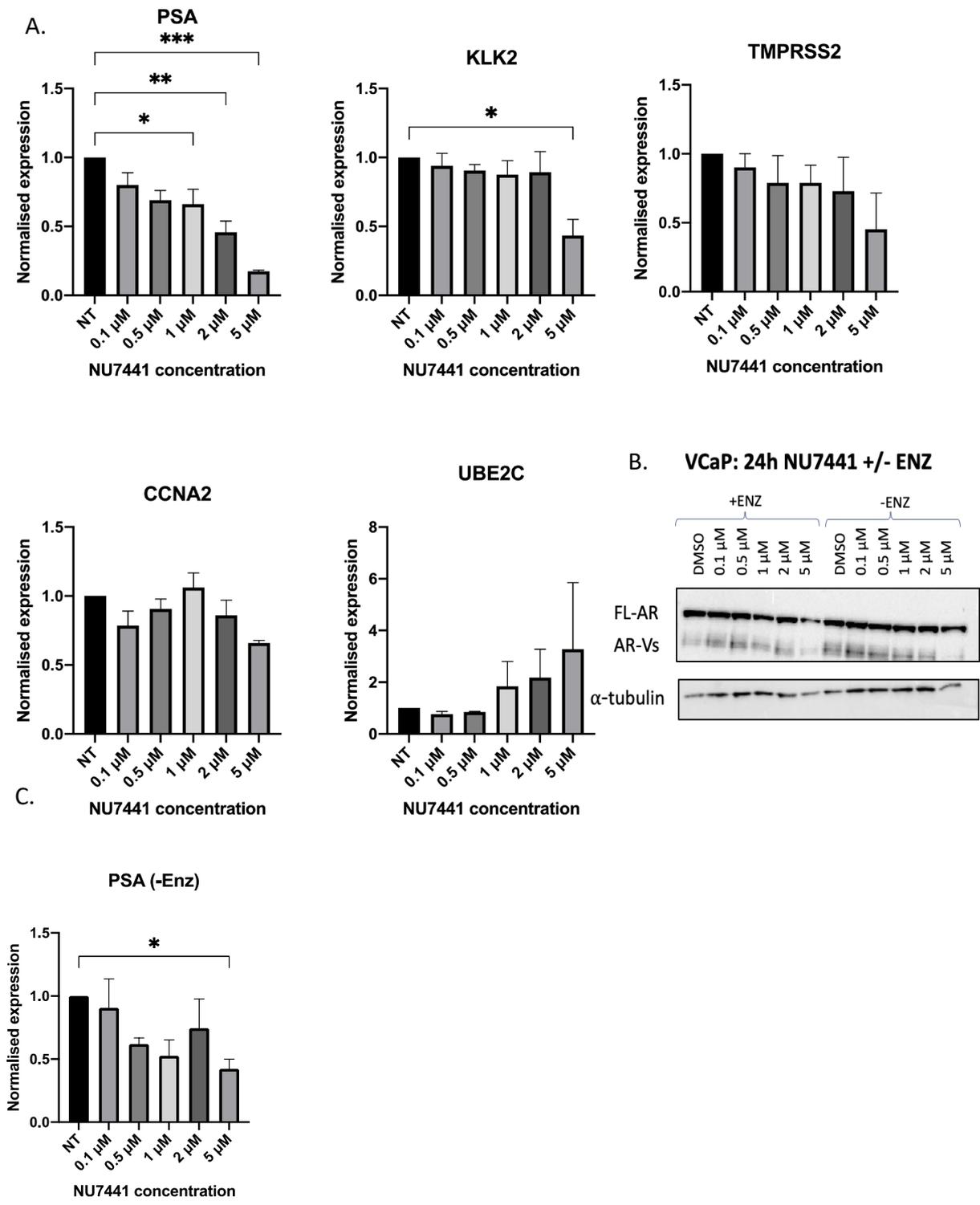
**Figure 4.7 DNA-PKcs inhibition reduces AR mediated transcription post 24-hour pre-enzalutamide treatment in CWR22Rv1 cells.**

CWR22Rv1 cells were cultured in steroid-depleted media for 48 hours, then pre-treated with 10 μM enzalutamide for 24 hours before 24-hour NU7441 treatment at 1 and 5 μM. Cells were then subject to AR target gene expression analysis using qRT-PCR. Data was normalised to the NT (DMSO control) arm for each target gene. Data is representative of three independent repeats ± SEM. One-way ANOVA using Bonferroni Post Hoc analysis was used to determine statistical significance (\* =  $p < 0.05$ , \*\* =  $p < 0.01$ , \*\*\* =  $p < 0.001$ ).

To further validate the effects of DNA-PKcs inhibition on AR-V transcriptional activity, VCaP cells were subject to NU7441 treatments prior to qRT-PCR analysis of AR target gene expression. VCaP cells have an amplified AR gene locus which results in high levels of AR which is a phenomenon commonly observed in CRPC patients and is important in treatment resistance. Considering AR-V expression in VCaP cells is elevated in androgen-depleted conditions, cells were cultured in steroid-depleted conditions to maximise AR-V signalling prior to treatment with NU7441 plus or minus enzalutamide for 24 hours. This experimental

design allowed investigation of the effect of DNA-PKcs inhibition on AR-V-mediated transcription in this cell line. As shown in Figure 4.8, *PSA* and *KLK2* expression were significantly reduced in response to 5  $\mu$ M NU7441 plus enzalutamide, and *PSA* expression was significantly downregulated in the 1  $\mu$ M treatment arm. In contrast, expression of *UBE2C*, an AR-V-regulated gene, increased in the presence of NU7441 and *CCNA2*, another AR-V regulated gene, was not affected by NU7441 treatment in the VCaP cell line. However, as shown in Figure 4.9, the AR agonist DHT diminishes *CCNA2* and *UBE2C* expression, suggesting FL-AR can repress expression, which suggests increased expression of *UBE2C* could be due to compromised FL-AR activity due to enzalutamide.

Cells grown in the absence of enzalutamide appeared to be less effected by DNA-PKcs inhibition; there was some significance in changes to *PSA* mRNA levels although not as robust an effect as observed in the cells grown in enzalutamide, indicating AR-Vs may be more sensitive to DNA-PKcs inhibition (Figure 4.8C). Moreover, the immunoblot in Figure 4.8B shows AR and AR-V protein levels after 24 hours DNA-PKcs inhibition with and without enzalutamide co-treatment. This showed AR-V levels appear to reduce at 1, 2 and 5  $\mu$ M doses. This is further explored in section 4.2.4 and Chapter 5.



**Figure 4.8 DNA-PKcs inhibition reduces transcription of some AR-V target genes in VCaP cells.**

A. VCaP cells were cultured in steroid-depleted media for 48 hours and then treated with 10 μM enzalutamide and increasing concentrations of NU7441 for 24 hours. B. VCaP cells were cultured in steroid-depleted media for 48 hours and then treated with increasing concentrations of NU7441 for 24 hours plus and minus 10 μM enzalutamide. C. VCaP cells were cultured in steroid-depleted media for 48 hours and then treated increasing concentrations of NU7441 for 24 hours. Cells were then subject to target gene expression analysis using qRT-PCR and AR and α-tubulin protein analyses using western blot. Data was normalised to the NT treatment arm for each target gene. Data is

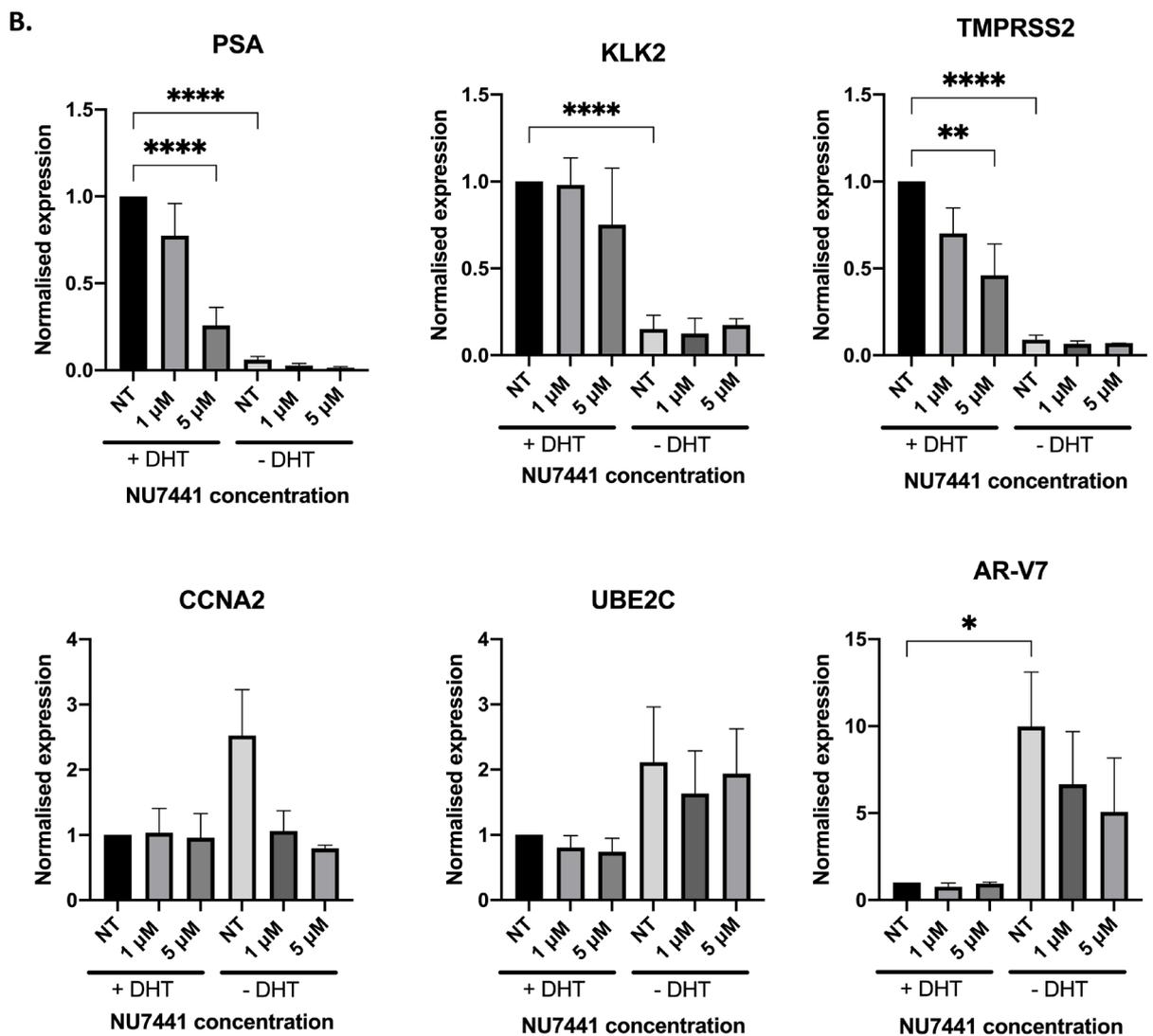
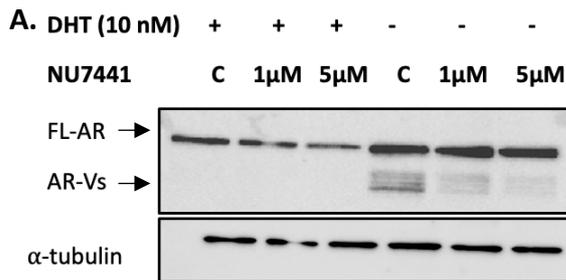
representative of three independent repeats  $\pm$  SEM. One-way ANOVA using Bonferroni Post Hoc analysis was used to determine the statistical significance (\* =  $p < 0.05$ , \*\* =  $p < 0.01$ , \*\*\* =  $p < 0.001$ ).

#### 4.2.4 DNA-PKcs inhibition in VCaP cells reduces AR-V protein levels

After 24 hours NU7441 treatment in VCaP cells grown in steroid depleted conditions with or without AR stimulation with DHT, cells were harvested for protein and mRNA analysis. Figure 4.9A shows AR and AR-V protein levels are downregulated in response to DHT. This is expected as AR-Vs are only expressed in VCaP cells when they are cultured in steroid-depleted conditions due to a repressive element in the *AR* gene which downregulates AR expression in response to AR stimulation (Cai *et al.*, 2011). As expected, DHT treatment increased expression of PSA, KLK2 and TMPRSS2 mRNA, but mRNA levels of AR-V specific target genes *UBE2C* and *CCNA2* are decreased. The downregulation of AR target genes in response to DNA-PKcs inhibition is consistent with what has been shown previously in this study. Furthermore, in the minus DHT arm, DNA-PKcs inhibition robustly downregulates *CCNA2* levels. AR protein levels were reduced as seen in Figure 4.9A, after 1 and 5  $\mu$ M NU7441 in the absence of DHT. This suggests DNA-PKcs is involved in the generation or stabilisation of AR-V protein and could be the mechanism behind the reduction observed in AR-V target gene expression upon DNA-PKcs inhibition.

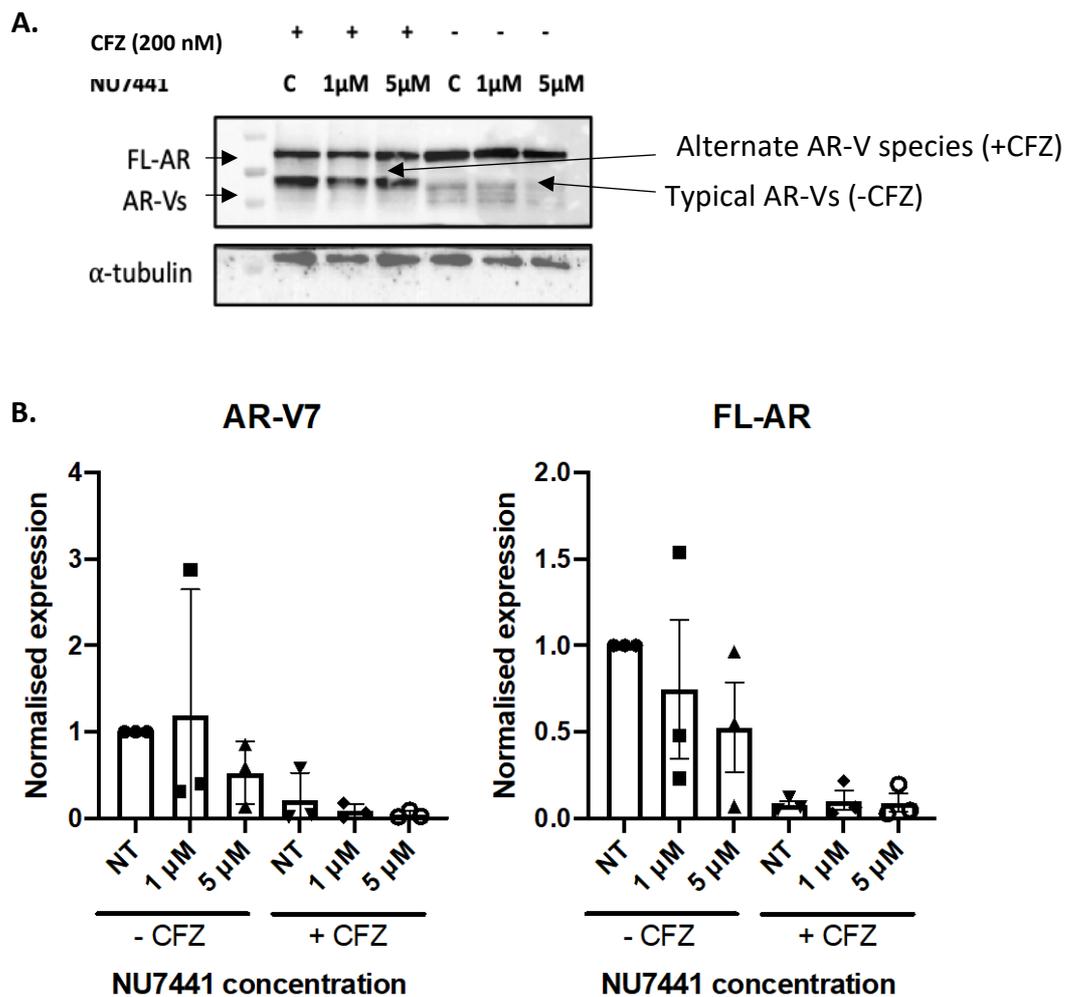
To assess this mechanism further, VCaP cells were subject to NU7441 treatment in the presence and absence of the proteasome inhibitor carfilzomib. If DNA-PKcs inhibition is enhancing destabilisation of AR-Vs, it is expected that blocking the proteasome would prevent downregulation of AR-V levels. Critically, the reduction of AR-V protein levels initially appeared to be blocked with the use carfilzomib without effecting FL-AR protein levels. However, as shown in the protein blot in Figure 4.10, the AR-V protein that is detected is slightly bigger than the 'typical' AR-Vs observed in VCaP cells in the absence of carfilzomib. This suggests that carfilzomib may either cause a build-up of mono-ubiquitylated forms of AR-V that are not detected without proteasome inhibition due to rapid turnover and degradation or it stabilises *de novo* an unstable AR-V that is not usually observed in steady-state. These findings could suggest that DNA-PKcs inhibition impacts AR-V metabolism at two levels; causing AR-V degradation by the proteasome in VCaP cells, and by compromising AR-V mRNA production in response to 1 and 5  $\mu$ M NU7441. However, because of the robust downregulation of AR-V and FL-AR transcript levels in response to carfilzomib alone, this is

difficult to interpret in this particular set up. Individual data points have been plotted to show the datapoint that appears to be an outlier to demonstrate the degree of AR-V transcript level reduction across the other two repeats. Another experiment that could be performed is to incorporate cyclohexamide to inhibit protein translation to then analyse AR-V protein degradation over time using western blotting to determine if DNA-PKcs inhibition reduces the half-life of AR-Vs.



**Figure 4.9 DNA-PKcs inhibition reduces AR-V protein levels in VCaP cells.**

VCaP cells were cultured in steroid-depleted media for 48 hours and then treated with or without 10 nM DHT and increasing concentrations of NU7441 for 24 hours. Cells were then subject to AR protein and  $\alpha$ -tubulin analyses using western blot (A) and target gene expression analysis using qRT-PCR (B). Data was normalised to the DMSO (NT) treatment arm for each target gene. Data is representative of three independent repeats  $\pm$  SEM. One-way ANOVA using Bonferroni Post Hoc analysis was used to determine the statistical significance (\* =  $p < 0.05$ , \*\* =  $p < 0.01$ , \*\*\* =  $p < 0.001$ ).

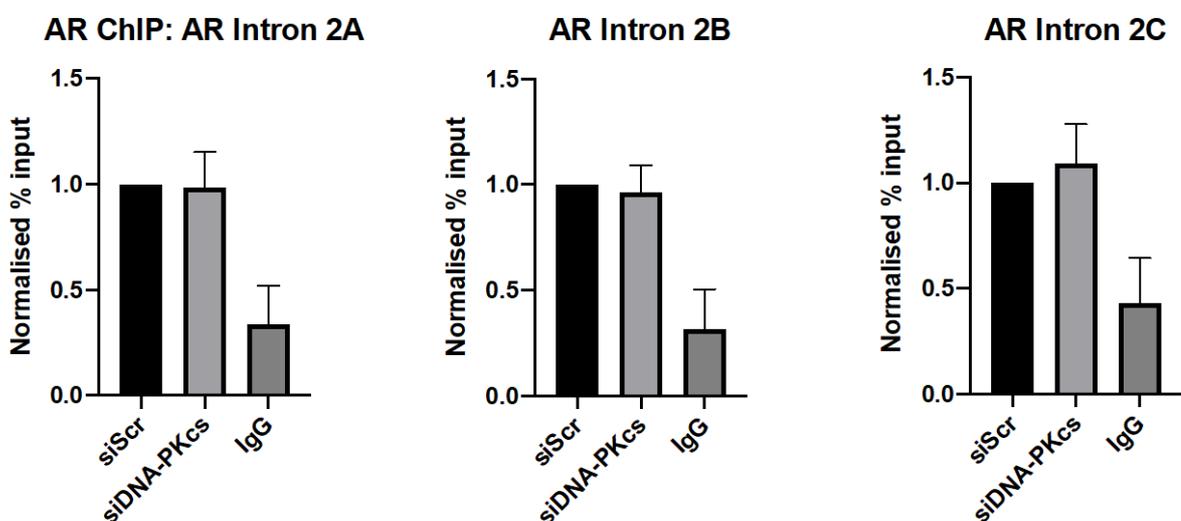


**Figure 4.10: DNA-PKcs inhibition reduces AR-V protein levels in VCaP cells, that can be partially blocked with the addition of a proteasome inhibitor.**

VCaP cells were cultured in steroid-depleted media for 48 hours and then treated with or without 200 nM carfilzomib (CFZ) and increasing concentrations of NU7441 for 24 hours. Cells were then subject to AR and  $\alpha$ -tubulin protein level analysis using western blot (A) and target gene expression analysis using qRT-PCR (B). Data was normalised to the DMSO (NT) treatment arm for each target gene. Data

is representative of three independent repeats  $\pm$  SEM. One-way ANOVA using Bonferroni Post Hoc analysis was used to determine the statistical significance (\* =  $p < 0.05$ , \*\* =  $p < 0.01$ , \*\*\* =  $p < 0.001$ ).

To further interrogate the mechanism by which AR-V protein levels were being reduced upon DNA-PKcs inhibition, the VCaP model was used to interrogate if DNA-PKcs was involved in controlling the AR negative feedback loop active in this cell line. This feedback loop is a mechanism by which the AR controls its own expression. Upon androgenic stimulation, the AR binds to downstream repressive elements within AR intron 2 that in turn reduces AR expression (Cai *et al.*, 2011). Given the effect of DNA-PKcs blockade on AR-V transcript levels, it was hypothesised that DNA-PKcs could be involved in controlling the recruitment of AR to these repressive elements. Therefore, ChIP experiments were carried out to determine if recruitment of AR at 3 downstream repressive elements in intron 2 of the AR gene was impacted by compromising DNA-PKcs activity, which could be causing the reduction of AR-V protein levels. DNA-PKcs siRNA mediated knockdown was used in this experiment to fully deplete DNA-PKcs activity and expression, which will provide a cleaner readout of AR dependency on DNA-PKcs at these specific regions. This revealed robust enrichment of the AR at the intronic regulatory regions over the IgG isotype control, however, no change in AR recruitment was detected upon DNA-PKcs depletion at each of the regulatory regions, suggesting other mechanisms that control AR-V transcript generation are impacted upon DNA-PKcs inhibition (Figure 4.11). The mechanism behind AR-V mRNA and protein reduction is interrogated further in Chapter 5.



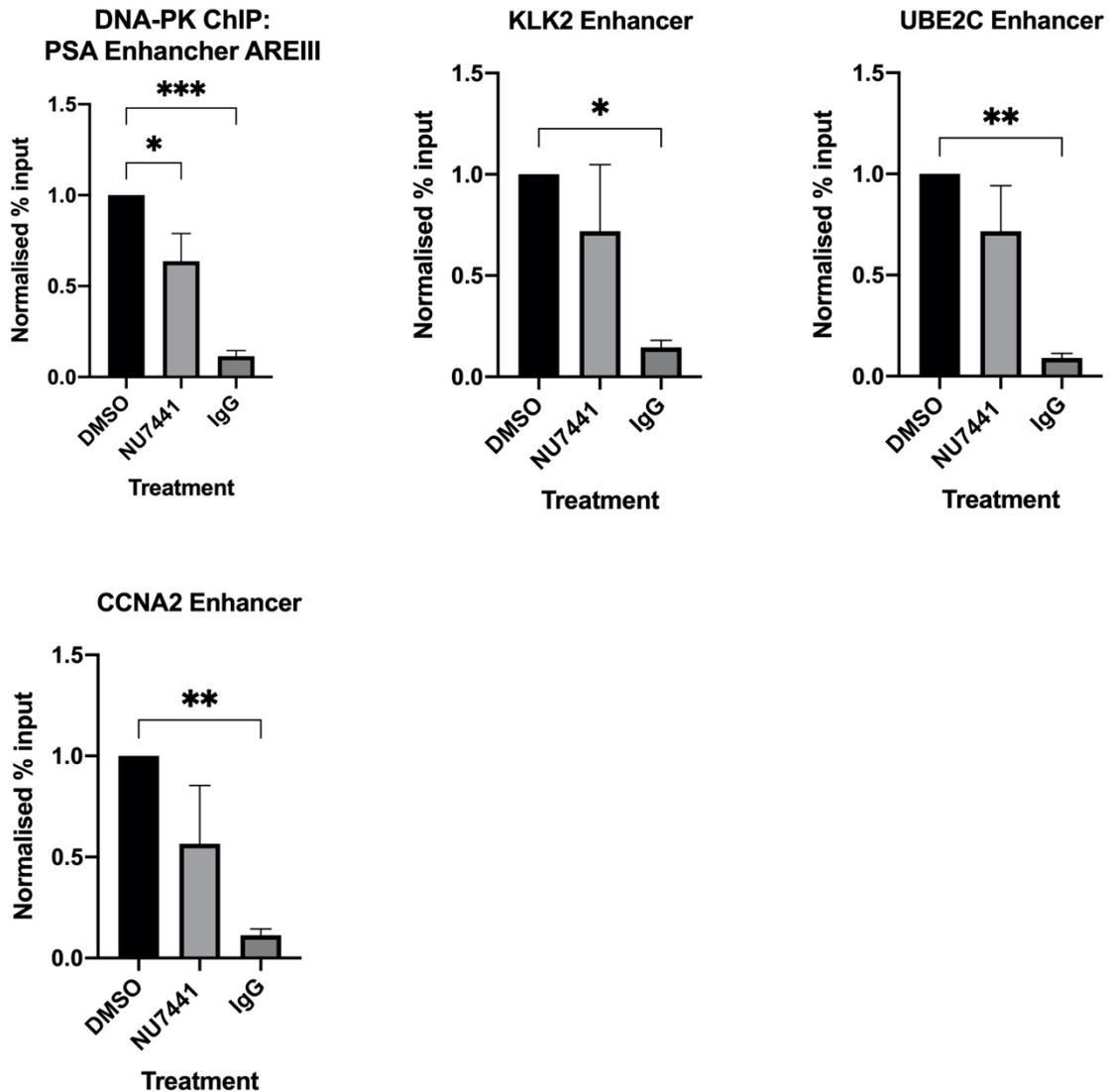
**Figure 4.11 AR CHIP showed no change in recruitment to regulatory regions in AR introns in VCaP cells.**

VCaP cells were grown in steroid-depleted media for 72 hours post siRNA-mediated knockdown of DNA-PKcs (siDNA-PKcs) or control scrambled siRNA (siScr) transfection after which AR CHIP was performed. CHIP-qPCR readouts show recruitment of AR to repressive elements present in *AR* intron 2 over the IgG control. Data shown represents the normalised percentage input to the control and represents 3 independent repeats  $\pm$  SEM.

4.2.5 DNA-PKcs is recruited to AR target genes in CWR22Rv1-AR-EK cells

Previous data has shown that DNA-PKcs is recruited to the regulatory elements of AR target genes *PSA* and *TMPRSS2*, in response to DHT stimulation in FL-AR-expressing C4-2 PCa cells (Goodwin *et al.*, 2015). To determine if DNA-PKcs was also recruited to AR-V regulatory elements in CWR22Rv1-AR-EK cells, CHIP experiments, incorporating anti-DNA-PKcs or isotype control (IgG) antibodies, were performed in full media supplemented with and without 1  $\mu$ M NU7441 for 24-hours. As shown in Figure 4.12, significant DNA-PKcs recruitment to *cis*-regulatory elements of the AR-V target genes *PSA*, *KLK2*, *UBE2C* and *CCNA2* was observed in CWR22Rv1-AR-EK cells compared to IgG controls, that is only marginally, but non-significantly reduced upon inhibition with NU7441 with the exception of the *PSA* enhancer.

## CWR22Rv1-AR-EK



**Figure 4.12: DNA-PKcs is recruited to AR regulatory elements in CWR22Rv1-AR-EK cells, that is largely refractory to inhibition with 1  $\mu$ M NU7441.**

CWR22Rv1-AR-EK cells were seeded out in full media and allowed to adhere to the dish and grow to 70-80% confluency before 24h treatment with 1  $\mu$ M NU7441 after which DNA-PKcs ChIP was performed. ChIP-qPCR readouts show recruitment of DNA-PKcs to AR target genes over the IgG control. Data shown represents the normalised percentage input to the control of 3 independent experiments  $\pm$  SEM. One-way ANOVA using Bonferroni Post Hoc analysis was used to determine the statistical significance (\* =  $p < 0.05$ , \*\* =  $p < 0.01$ )

### 4.2.6 DNA-PKcs inhibition does not affect AR recruitment to target genes

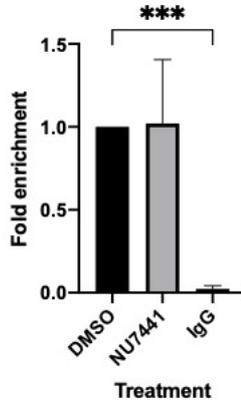
As DNA-PKcs inhibition reduced *PSA*, *KLK2*, *UBE2C* and *CCNA2* mRNA levels in some of the tested cell lines, we next sought to determine if the effect of NU7441 on AR target gene

expression was due to reduced recruitment of AR-Vs to *cis*-regulatory elements of these genes. CWR22Rv1 cells were cultured as previously described, with a 24-hour pre-treatment with 10  $\mu$ M enzalutamide and 24-hour treatment with 1  $\mu$ M NU7441. The cells were then subject to ChIP using either anti-AR or IgG control antibodies to investigate if DNA-PKcs inhibition affected AR/AR-V recruitment to these target genes. The AR antibody used was an N-terminal specific antibody enabling the interrogation of both FL-AR and AR-V recruitment, but because FL-AR should be inactivated by enzalutamide, and hence will be cytoplasmic, enrichment of AR at target loci will likely reflect AR-V abundance. We investigated the recruitment of AR-Vs to the *PSA* enhancer and enhancer elements of *KLK2* and *UBE2C*. As expected, in the control arm, robust enrichment of AR-Vs was detected at all three *cis*-regulatory elements. In response to DNA-PKcs inhibition, there was no significant impact on AR-V recruitment, with the exception of the *KLK2* enhancer which showed a modest increase in AR recruitment, suggesting that DNA-PKcs may alter AR-V activity in manner distinct from regulating chromatin deposition in the investigated cell lines (Figure 4.13). This experiment was repeated in CWR22Rv1-AR-EK cells and VCaP cells. Consistent with what was seen in the parental CWR22Rv1 cell line, AR recruitment to key target gene enhancer elements was not statistically significantly impacted, again indicating it is not AR chromatin binding that is causing a reduction of AR-V transcriptional activity upon DNA-PKcs manipulation. In VCaP cells, ChIP was performed after DNA-PKcs knockdown for 72 hours. Consistent with what was shown in CWR22Rv1 and CWR22Rv1-AR-EK cells, no significant change in AR recruitment was seen upon DNA-PKcs depletion.

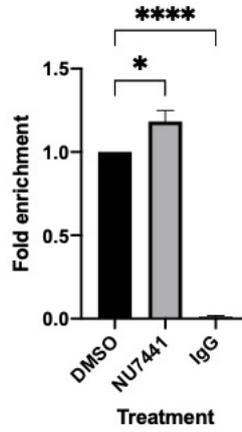
To further interrogate the chromatin biology surrounding DNA-PKcs-AR-V interplay, a DNA-PKcs ChIP following AR siRNA-mediated depletion was carried out. It has been previously shown that FL-AR inactivation by enzalutamide treatment in the CRPC cell line C4-2, attenuates DNA-PKcs recruitment to regulatory elements in the *AR* gene. CWR22Rv1-AR-EK cells depleted of AR using an *AR* exon 1 siRNA (siARex1) were subjected to DNA-PKcs ChIP along with a scrambled siRNA (siScr) and IgG control arms. DNA-PKcs was recruited to *PSA*, *KLK2* and *CCNA2* and this was reduced upon AR depletion, although this was not statistically significant (Figure 4.14). This indicates there is not a requirement for AR chromatin binding to enable DNA-PKcs recruitment at these regions in this cell line.

**CWR22Rv1**

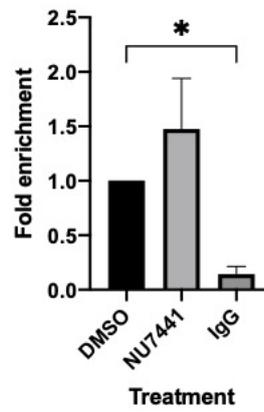
**AR ChIP: PSA Enhancer AREIII**



**KLK2 Enhancer**

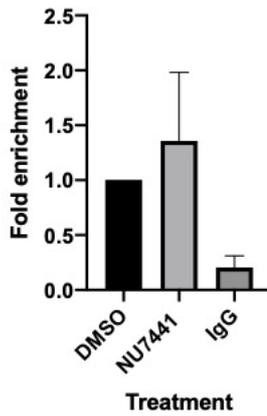


**UBE2C Enhancer**

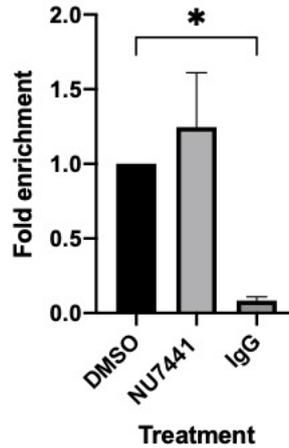


**CWR22Rv1-AR-EK**

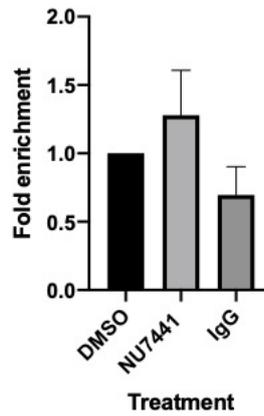
**AR ChIP: PSA Enhancer AREIII**



**KLK2 Enhancer**

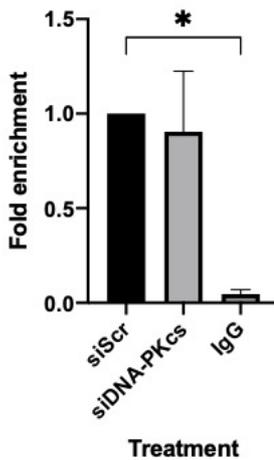


**UBE2C Enhancer**

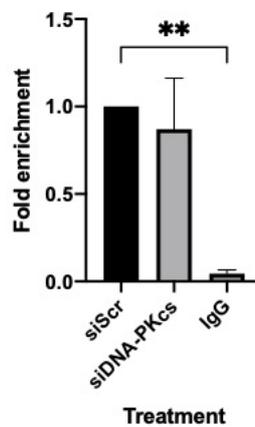


**VCaP**

**AR ChIP: PSA Enhancer AREIII**



**KLK2 Enhancer**



**UBE2C Enhancer**

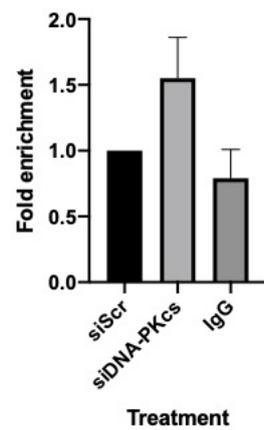
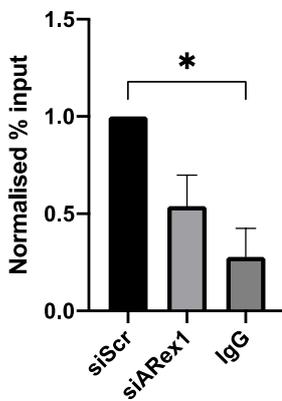


Figure 4.13 NU7441 does not affect AR recruitment to cis-regulatory elements of some key AR-regulated genes.

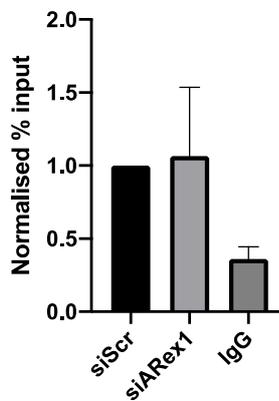
CWR22Rv1 cells were cultured in steroid-depleted media for 48 hours before the addition of 10  $\mu$ M Enzalutamide for 24 hours then 1  $\mu$ M NU7441 for 24 hours before being harvested and subjected to ChIP analysis using either AR or IgG control antibodies. CWR22Rv1-AR-EK cells were cultured in full media until 70-80% confluent and subject to ChIP analysis. VCaP cells were seeded in steroid-depleted media and reverse transfected with DNA-PKcs (siDNA-PKcs) and scrambled (siScr) siRNA and before being harvested for ChIP analysis. AR enrichment at *PSA*, *KLK2* and *UBE2C* cis-regulatory elements over an IgG isotype control was determined by qRT-PCR. Data is displayed as fold enrichment normalised to the DMSO control. Data is representative of three independent repeats  $\pm$  SEM. One-way ANOVA using Bonferroni Post Hoc analysis was used to determine the statistical significance (\* =  $p < 0.05$ , \*\* =  $p < 0.01$ , \*\*\* =  $p < 0.001$ ).

### **CWR22Rv1-AR-EK**

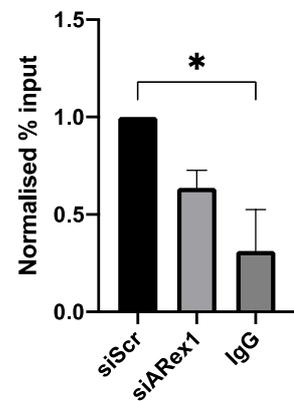
#### **DNA-PKcs ChIP: PSA AREI**



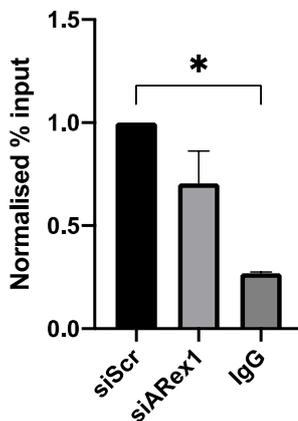
#### **PSA Enhancer AREIII**



#### **KLK2 Enhancer**



#### **CCNA2 Enhancer**

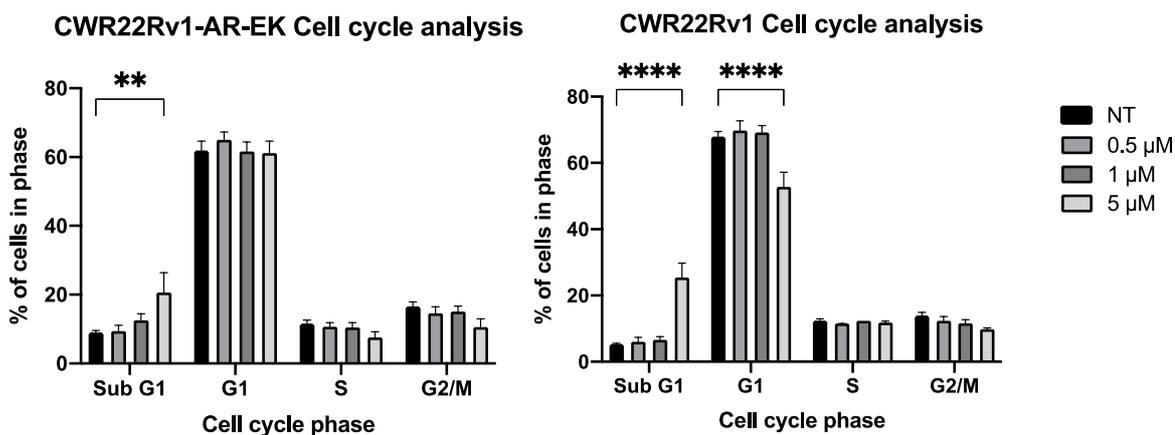


**Figure 4.14 DNA-PKcs is not dependent on AR presence at AR target genes in CWR22Rv1-AR-EK.**

CWR22Rv1-AR-EK cells were seeded out in full media and reverse transfected with either scrambled control (siScr) or AR exon 1-targeting (siARexon1) siRNAs and incubated for 72 hours after which DNA-PKcs ChIP was performed. ChIP-qRT-PCR readouts show recruitment of DNA-PKcs to AR target genes over the IgG control. Data shown represents the normalised percentage input to the control and represents 2 independent repeats. One-way ANOVA was used to determine the statistical significance (\* =  $p < 0.05$ , \*\* =  $p < 0.01$ )

#### 4.2.7 DNA-PKcs inhibition impacts the cell cycle profile of CWR22Rv1 and CWR22Rv1-AR-EK cells

Having determined that DNA-PKcs inhibition has a robust impact on PCa cell line proliferation at 1 and 5  $\mu\text{M}$  (Figure 4.3) and established a reduction of AR-V regulated cell cycle genes, such as *CCNA2*, the cell cycle profiles of CWR22Rv1 and CWR22Rv1-AR-EK cells were next determined following 96-hours 0.5, 1 and 5  $\mu\text{M}$  NU7441 treatment. CWR22Rv1-AR-EK cells were seeded out in full media and CWR22Rv1 in SDM for 24 hours before drug treatments. Cells were then harvested 96 hours later and stained with PI and analysed using a Thermo Attune flow cytometer. Unexpectedly, as shown in Figure 4.15, 96 hours drug treatment at the lower doses of 0.5 and 1  $\mu\text{M}$  NU7441 failed to cause significant changes in any of the cell cycle phases across the two cell lines. At 5  $\mu\text{M}$ , there is a significant increase in the percentage of cells in the sub-G1 phase in CWR22Rv1 and CWR22Rv1-AR-EKs, suggesting the cells have undergone apoptosis. The average values of cells in the sub-G1 phase in CWR22Rv1-AR-EK cells were 9%, 9.4%, 12.6% and 20.7% and in CWR22Rv1 cells were 5.4 %, 6%, 6.7% and 25.4% for NT, 0.5  $\mu\text{M}$ , 1  $\mu\text{M}$  and 5  $\mu\text{M}$  doses, respectively. Moreover, in CWR22Rv1 cells, 5  $\mu\text{M}$  NU7441 caused a significant decrease in the percentage of cells in the G1 phase, likely due to the high rate of apoptosis. This suggests cells are passing through the S and G2/M phases of the cell cycle but some have major problems, resulting in an apoptotic phenotype. In CWR22Rv1-AR-EK cells, there is a drop in the percentage of cells in S and G2/M phases at the 5  $\mu\text{M}$  NU7441 dose which would account for why the decrease of cells in G1 as seen in CWR22Rv1 cells, is not seen in this cell line. However, due to the robust effects seen on proliferation, this was not as predicted. This could potentially be explained by a relatively fast cell death and the time points used in this experiment have missed the time point where cell cycle arrest could be seen.



**Figure 4.15 DNA-PKcs inhibition does not significantly affect the cell cycle profile of CWR22Rv1-AR-EK and CWR22Rv1 cells after 96 hours treatment at 1 μM.**

CWR22Rv1-AR-EK and CWR22Rv1 cells subject to propidium iodide staining showed that 0.5 and 1 μM doses of NU7441 did not significantly affect the cell cycle profile, while 5 μM caused elevation of a sub-G1 population indicative of cell apoptosis. Data was analysed using FCS express 7 and the data represents 3 repeats for CWR22Rv1-AR-EK cells and 2 repeats for CWR22Rv1.

#### 4.2.8 Validating DNA-PKcs-mediated regulation of AR-V activity using DNA-PKcs knockdown

Next, DNA-PKcs depletion using siRNA was investigated to determine if the effects observed using the kinase inhibitor NU7441 were replicated with knockdown of the protein. Firstly, an siRNA SMARTpool (Dharmacon) was reverse transfected into CWR22Rv1-AR-EK, CWR22Rv1 and VCaP cells, for 48 and 72 hours, before analysing AR-target genes and *PRKDC* (DNA-PKcs) expression, and DNA-PKcs protein levels to confirm knockdown, using qRT-PCR and western blot analysis, respectively.

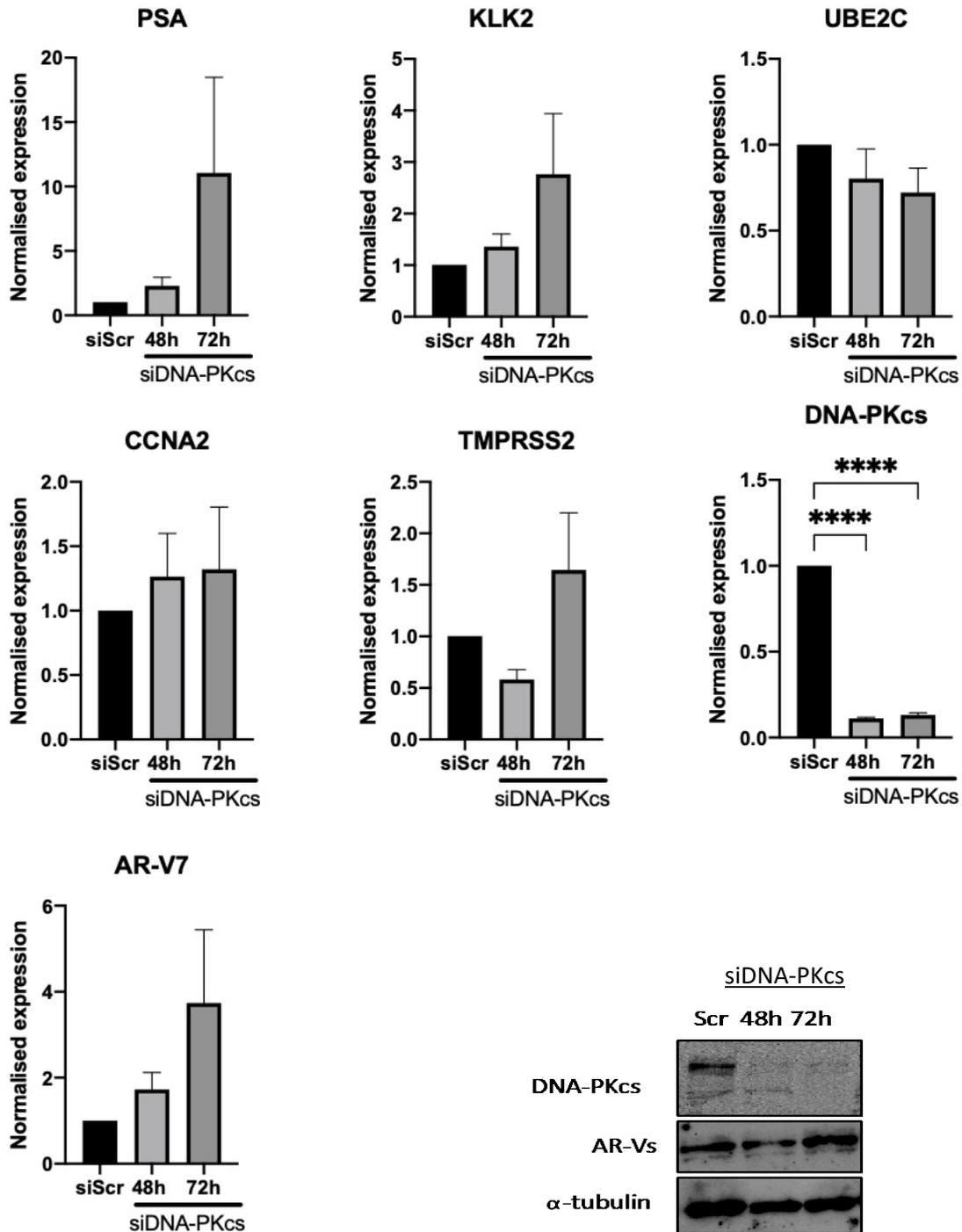
Firstly, DNA-PKcs depletion in CWR22Rv1-AR-EK cells did not mirror what was previously seen with DNA-PKcs inhibition. The target genes *PSA*, *KLK2*, *CCNA2* and *TMPRSS2* all increased 72 hours post knockdown. *UBE2C* showed a slight decrease in expression but this was not as pronounced as with DNA-PKcs inhibition and was not statistically significant. AR-V7 transcript levels were also examined, and this showed that DNA-PKcs depletion increased AR-V7 transcript abundance, which could also be seen at the protein level in Figure 4.18.

In CWR22Rv1 cells grown in SDM and subject to DNA-PKcs knockdown, similarly to CWR22Rv1-AR-EK cells, *PSA*, *KLK2*, *CCNA2* and *TMPRSS2* expression increased, while *UBE2C* was significantly reduced. AR-V7 and FL-AR mRNA levels also increased but to a lesser extent than

what was observed in CWR22Rv1-AR-EK cells (Figure 4.17). Critically, this was not consistent with what was seen with DNA-PKcs inhibition (Figure 4.6).

In VCaP cells grown in SDM, DNA-PKcs depletion caused all of the AR target genes to significantly decrease, indicating a dependency of AR-Vs on DNA-PKcs for maximal activity. Interestingly, AR-V7 and FL-AR transcript levels also significantly decreased after 48 and 72 hours, consistent with what was shown with DNA-PKcs inhibition (Figure 4.9), further validating that DNA-PKcs may have a role in the generation of AR-Vs in this cell line (Figure 4.18).

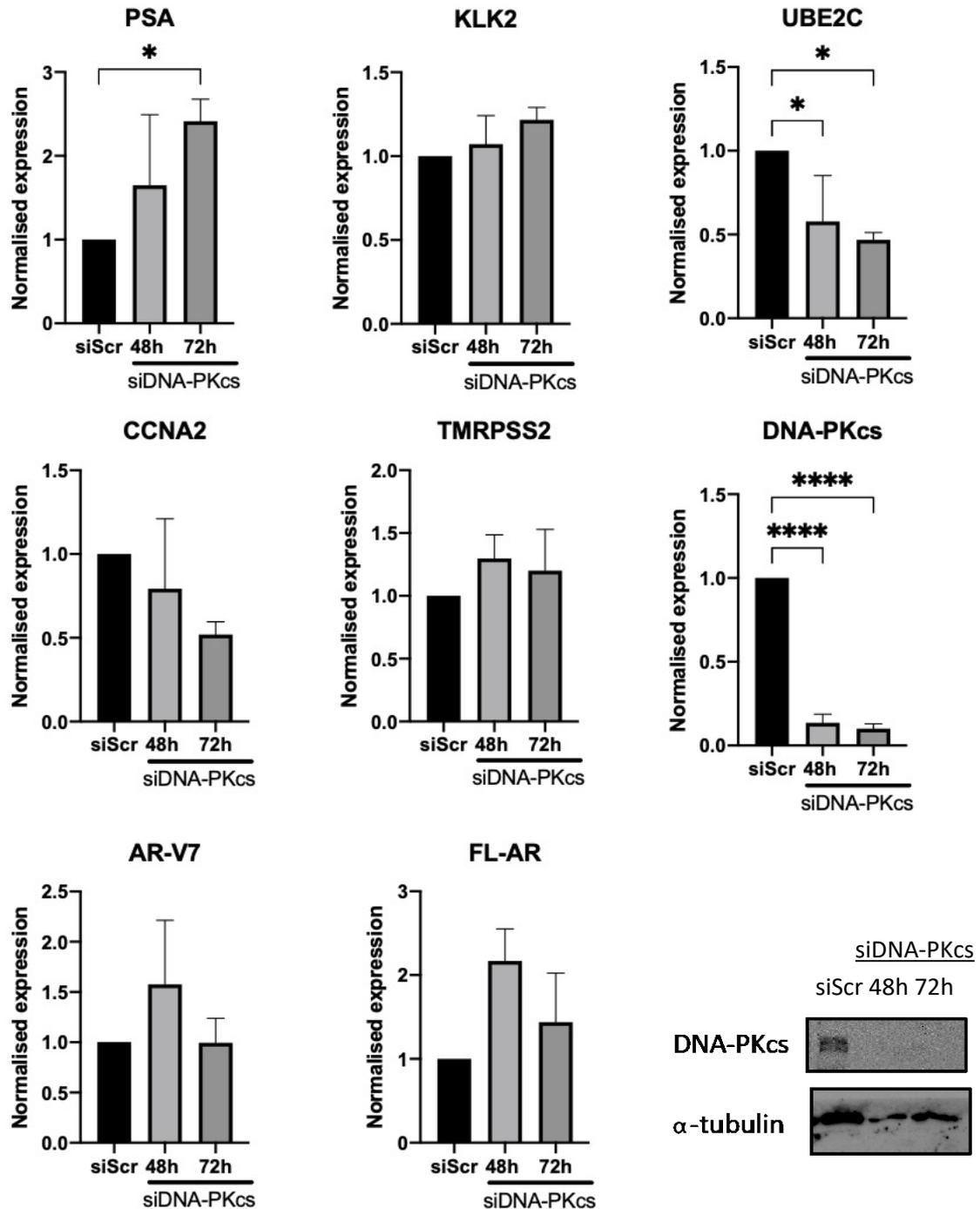
**CWR22Rv1-AR-EK**



**Figure 4.16 DNA-PKcs knockdown did not mirror pharmacological inhibition of DNA-PKcs inhibition on AR gene expression analysis.**

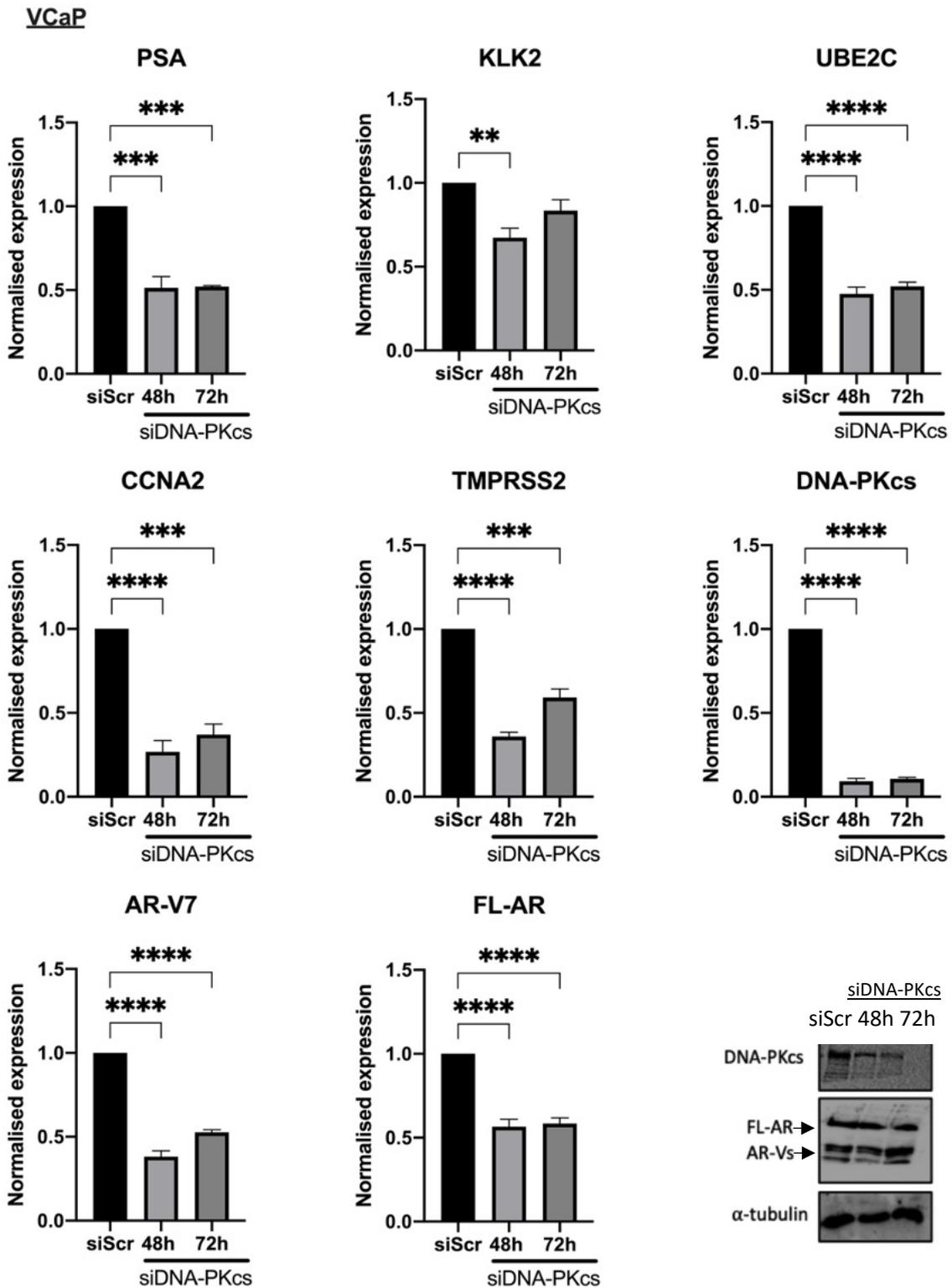
CWR22Rv1-AR-EK cells were reverse transfected with a DNA-PKcs siRNA pool and cells were harvested 48 and 72 hours later for qRT-PCR analysis of AR target genes and western blot analysis of DNA-PKcs, AR and  $\alpha$ -tubulin levels. qRT-PCR data is normalised to the siScr arm. Data is representative of three independent repeats  $\pm$  SEM. One-way ANOVA using Bonferroni Post Hoc analysis was used to determine the statistical significance (\*\*\*\* =  $p < 0.0001$ )

**CWR22Rv1**



**Figure 4.17 DNA-PKcs knockdown did not mirror pharmacological inhibition of DNA-PKcs inhibition on AR gene expression analysis.**

CWR22Rv1 cells were seeded in SDM and reverse transfected with a DNA-PKcs siRNA pool and cells were harvested 48 and 72 hours later for qRT-PCR analysis of AR target genes and western blot analysis of DNA-PKcs, AR and  $\alpha$ -tubulin levels. qRT-PCR data is normalised to the siScr arm. Data is representative of three independent repeats  $\pm$  SEM. One-way ANOVA using Bonferroni Post Hoc analysis was used to determine the statistical significance (\* =  $p < 0.05$ , \*\*\*\* =  $p < 0.0001$ )



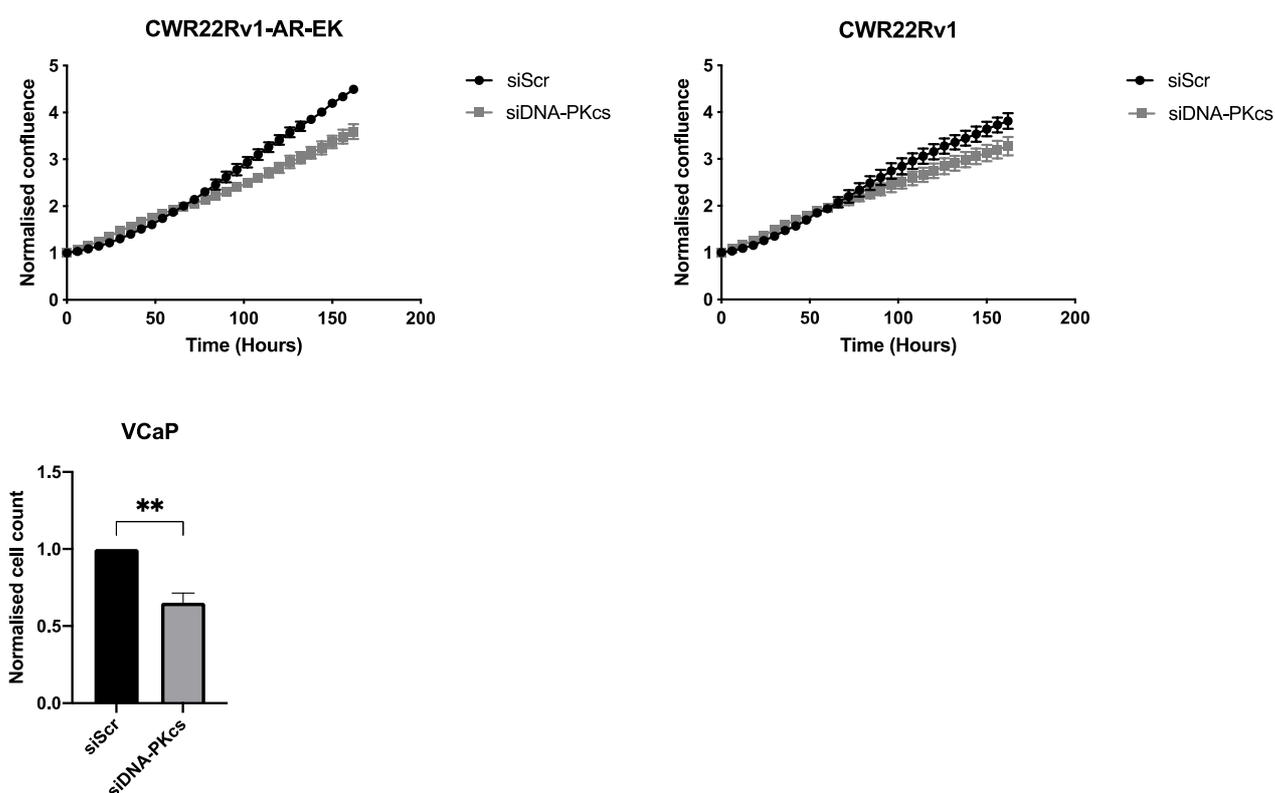
**Figure 4.18: DNA-PKcs knockdown did not mirror pharmacological inhibition of DNA-PKcs inhibition on AR gene expression analysis.**

VCaP cells were seeded in SDM and reverse transfected with a DNA-PKcs siRNA pool and cells were harvested 48 and 72 hours later for qRT-PCR analysis of AR target genes and western blot analysis of DNA-PKcs, AR and  $\alpha$ -tubulin levels. qRT-PCR data is normalised to the siScr arm. Data is representative

of three independent repeats  $\pm$  SEM. One-way ANOVA using Bonferroni Post Hoc analysis was used to determine the statistical significance (\* =  $p < 0.05$ , \*\* =  $p < 0.01$ , \*\*\* =  $p < 0.001$ , \*\*\*\* =  $p < 0.0001$ ).

### Validating DNA-PKcs-mediated regulation of proliferation in PCa cell lines using DNA-PKcs knockdown

Next, we wanted to determine if DNA-PKcs depletion caused similar effects on proliferation as was previously observed with DNA-PKcs inhibition. Proliferation was assessed after DNA-PKcs depletion using a manual cell count for VCaP cells and Incucyte® live cell imaging for CWR22Rv1-AR-EK and CWR22Rv1 cells grown in full and steroid-depleted conditions, respectively. VCaP cells were manually counted as the cells can swell when stressed so the confluence mask calculated using the Incucyte® does not accurately measure cell number or proliferation. As shown in Figure 4.19, DNA-PKcs knockdown caused a significant reduction in VCaP cell growth which is consistent with published reports indicating that DNA-PKcs depletion significantly diminished proliferation of LNCaP and C4-2 PCa cell lines (Dylgjeri *et al.*, 2019). In contrast, only a small, insignificant reduction in proliferation of CWR22Rv1-AR-EK and CWR22Rv1 cells was observed upon DNA-PKcs depletion which was not predicted given the anti-proliferative effects I had shown using DNA-PKcs inhibitors.

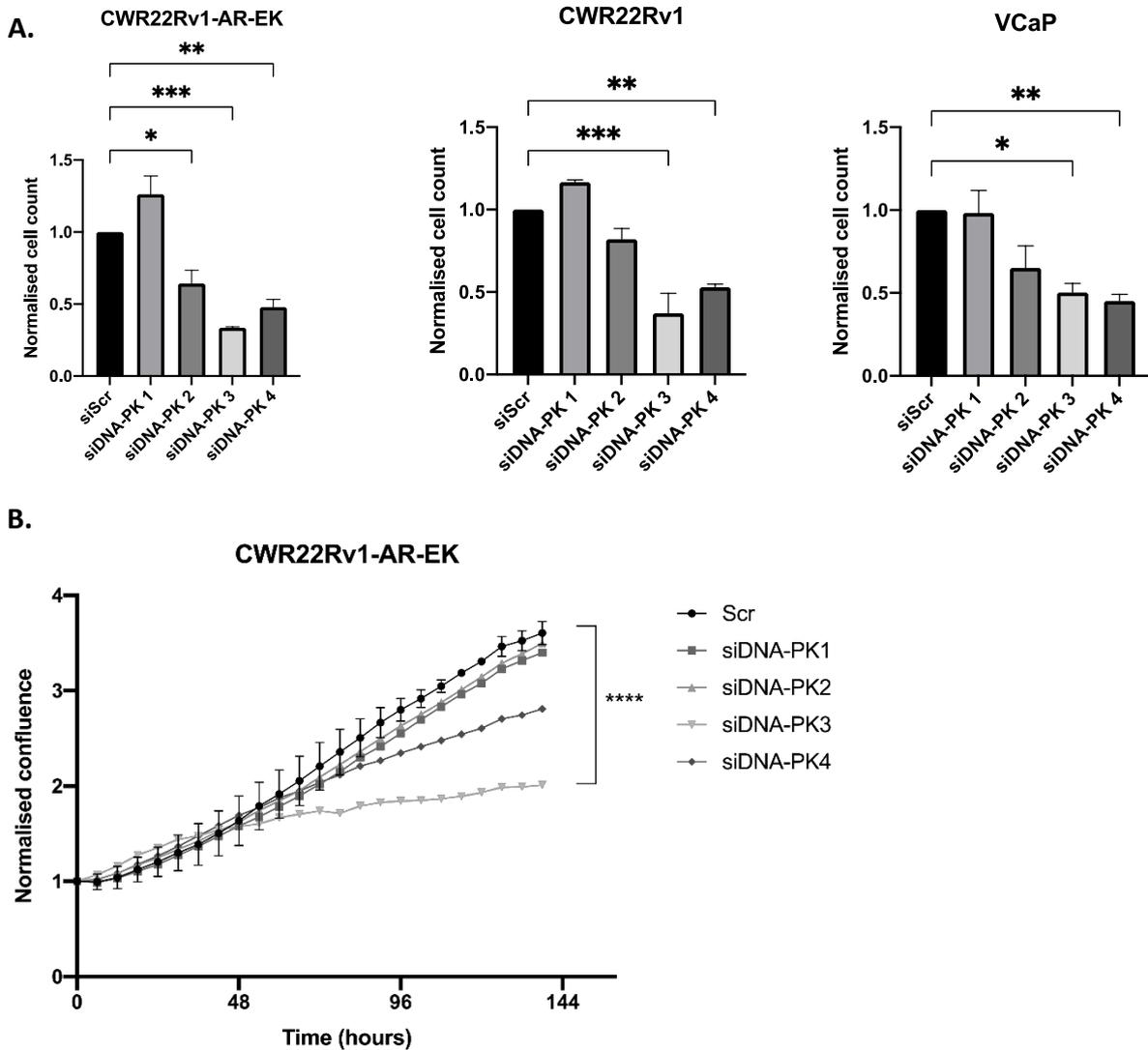


**Figure 4.19 DNA-PKcs knockdown showed varying results across the PCa cell lines CWR22Rv1-AR-EK, CWR22Rv1 and VCaP.**

CWR22Rv1-AR-EK and CWR22Rv1 cells were reverse transfected 24 hours before live-cell imaging using an Incucyte® ZOOM over 7 days to determine confluency. VCaPs were reverse transfected then manually counted 5 days later. Data shows average relative confluence and represent 3 repeats  $\pm$  SEM (\*\* =  $p < 0.01$ ).

As the DNA-PKcs knockdown data was not reflecting what has been shown with the kinase inhibitor, the individual siRNAs from the Dharmacon Smartpool were generated by Sigma; termed siDNA-PK1-4 and tested individually. This was to identify if irregularities observed between the NU7441 treatments and DNA-PKcs knockdown could be due to aberrant activities of an individual oligonucleotide. Cell count data from the deconvoluted library showed that each individual siRNA had different effects on proliferation after 96 hours (Figure 4.20A). This was validated using incucyte live cell imaging, where data showed a similar pattern to cell counts in CWR22Rv1-AR-EK cells (Figure 4.20B). siDNA-PK1 showed little effect

on cell proliferation, whereas siDNA-PK2-4 displayed varying levels of reduction, with siDNA-PK3 causing the greatest reduction in proliferation.



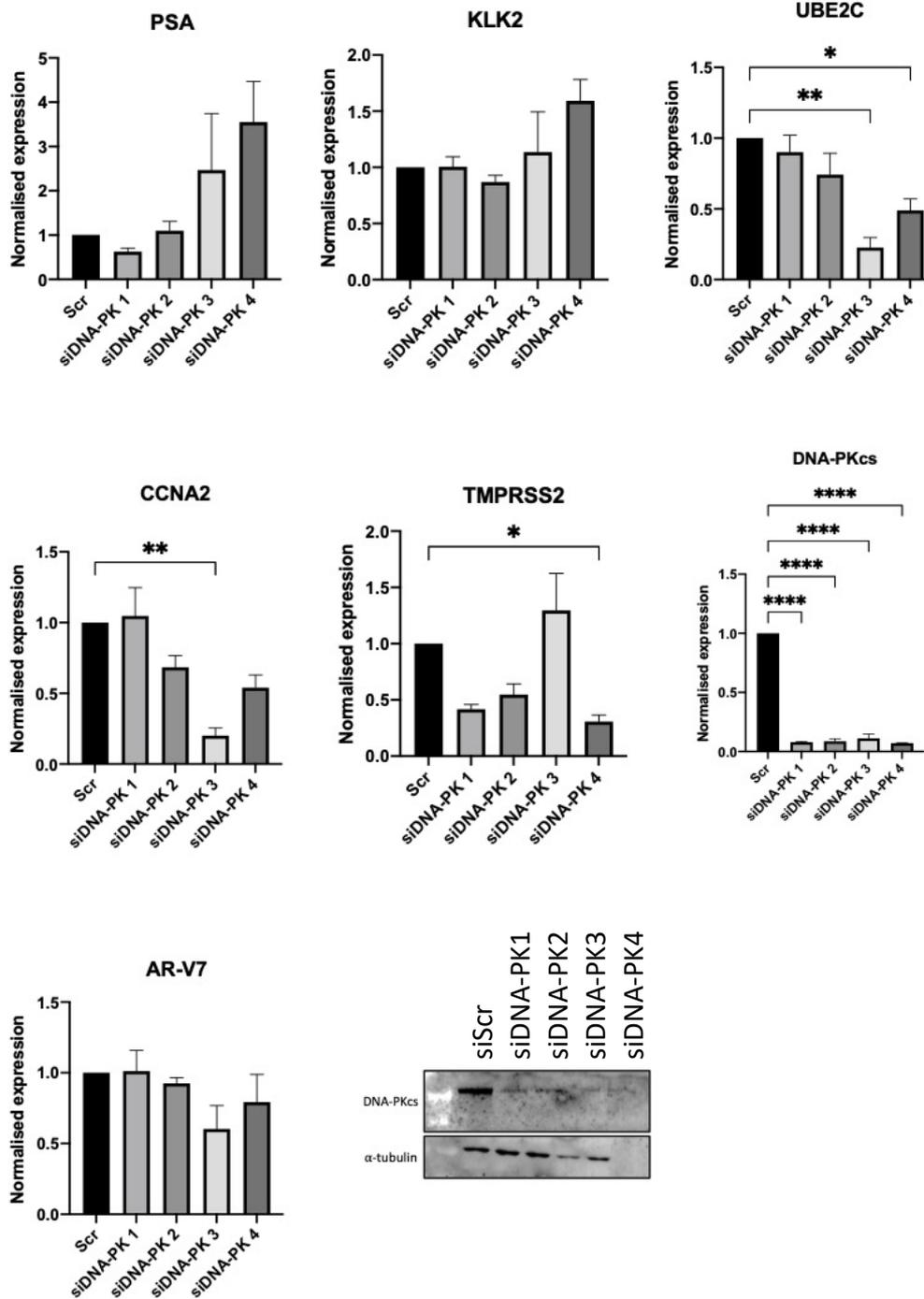
**Figure 4.20 DNA-PKcs knockdown using the deconvoluted Dharmacon SmartPool showed differing effects on proliferation of CWR22Rv1-AR-EK, CWR22Rv1 and VCaP cells.**

CWR22Rv1-AR-EK, CWR22Rv1 and VCaP cells were reverse transfected with each individual siRNA oligonucleotide targeting DNA-PKcs (siDNA-PK1-4) or a scrambled (siScr) control and incubated for 96 hours for cell counts (A) or analysed in an IncuCyte for 7 days (B). Cell counts are representative of 3 repeats and data is normalised to the siScr control and represents mean  $\pm$  SEM (n=3). Incucyte graph represents the mean normalised confluence  $\pm$  SEM (n=2) (\* =  $p < 0.05$ , \*\* =  $p < 0.01$ , \*\*\* =  $p < 0.001$ , \*\*\*\* =  $p < 0.0001$ ).

Additionally, DNA-PKcs depletion and AR target gene expression was analysed after 72-hour knockdown with each individual siRNA oligonucleotide. Intriguingly, although each oligonucleotide depleted DNA-PKcs to comparable levels, their effect on AR target genes was not consistent and this phenomenon was common across the three PCa cell lines. CWR22Rv1-AR-EK, grown in full media (Figure 4.23), and CWR22Rv1, grown in SDM (Figure 4.22), showed similar changes in AR target genes with each oligo. siDNA-PK3 and siDNA-PK4 robustly downregulates *UBE2C* and *CCNA2* expression in both CWR22v1-AR-EK and CWR22Rv1 cells. Furthermore, *TMPRSS2* transcript levels are robustly downregulated in response to siDNA-PK1, 2 and 4 in CWR22Rv1-AR-EK and CWR22Rv1 cells and none of the individual oligonucleotides statistically significantly impacted *PSA*, *KLK2* or *AR-V7* transcript levels.

Whereas VCaP cells, grown in SDM, are more sensitive to DNA-PKcs depletion with siDNA-PK1, 2 and 4 all showing significant reductions in the AR target genes *PSA*, *KLK2*, *UBE2C*, *CCNA2* and *TMPRSS2* as well as significantly reducing *AR-V* transcript levels (Figure 4.23). However, siDNA-PK3 selectively downregulates *UBE2C* and *CCNA2*, consistent with what was shown in CWR22Rv1-AR-EK and CWR22Rv1 cells (Figure 4.5 and Figure 4.6). Depletion of both DNA-PKcs transcript and protein levels were confirmed at 72 hours by each of the four siRNAs.

**CWR22Rv1-AR-EK**

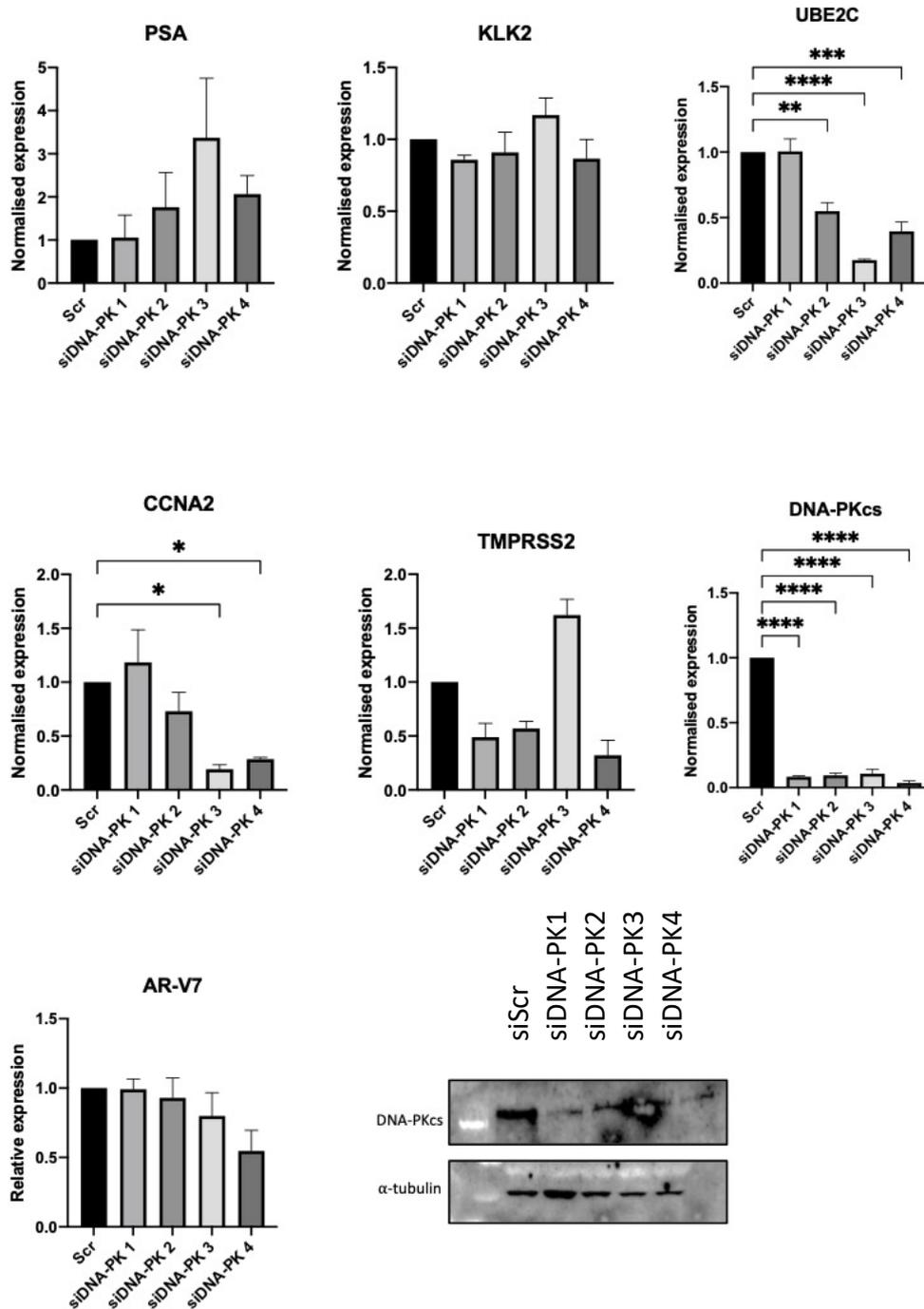


**Figure 4.21 DNA-PKcs knockdown using the deconvoluted Dharmacon SmartPool showed differing effects on AR target gene expression in CWR22Rv1-AR-EK cells.**

CWR22Rv1-AR-EK cells were seeded into full media and reverse transfected with oligonucleotides targeting DNA-PKcs (siDNA-PK1-4) or a scrambled (siScr) control for 72 hours. Cells were then subject to AR target gene expression analysis using qRT-PCR and western blot analysis of DNA-PKcs and  $\alpha$ -tubulin levels. Data was normalised to the siScr treatment arm for each target gene. Data is

representative of three independent repeats  $\pm$  SEM. One-way ANOVA using Bonferroni Post Hoc analysis was used to determine the statistical significance (\* =  $p < 0.05$ , \*\* =  $p < 0.01$ , \*\*\*\* =  $p < 0.0001$ ).

**CWR22Rv1**



**Figure 4.22 DNA-PKcs knockdown using the deconvoluted Dharmacon SmartPool showed differing effects on AR target gene expression in CWR22Rv1 cells.**

CWR22Rv1 cells were seeding in steroid-depleted media and reverse transfected with oligonucleotides targeting DNA-PKcs (siDNA-PK1-4) or a scrambled (siScr) control for 72 hours. Cells were then subject to AR target gene expression analysis using qRT-PCR and western blot analysis of DNA-PKcs and  $\alpha$ -tubulin levels. Data was normalised to the siScr treatment arm for each target gene. Data is representative of three independent repeats  $\pm$  SEM. One-way ANOVA using Bonferroni Post Hoc

analysis was used to determine the statistical significance (\* =  $p < 0.05$ , \*\* =  $p < 0.01$ , \*\*\* =  $p < 0.001$ , \*\*\*\* =  $p < 0.0001$ )

### VCaP

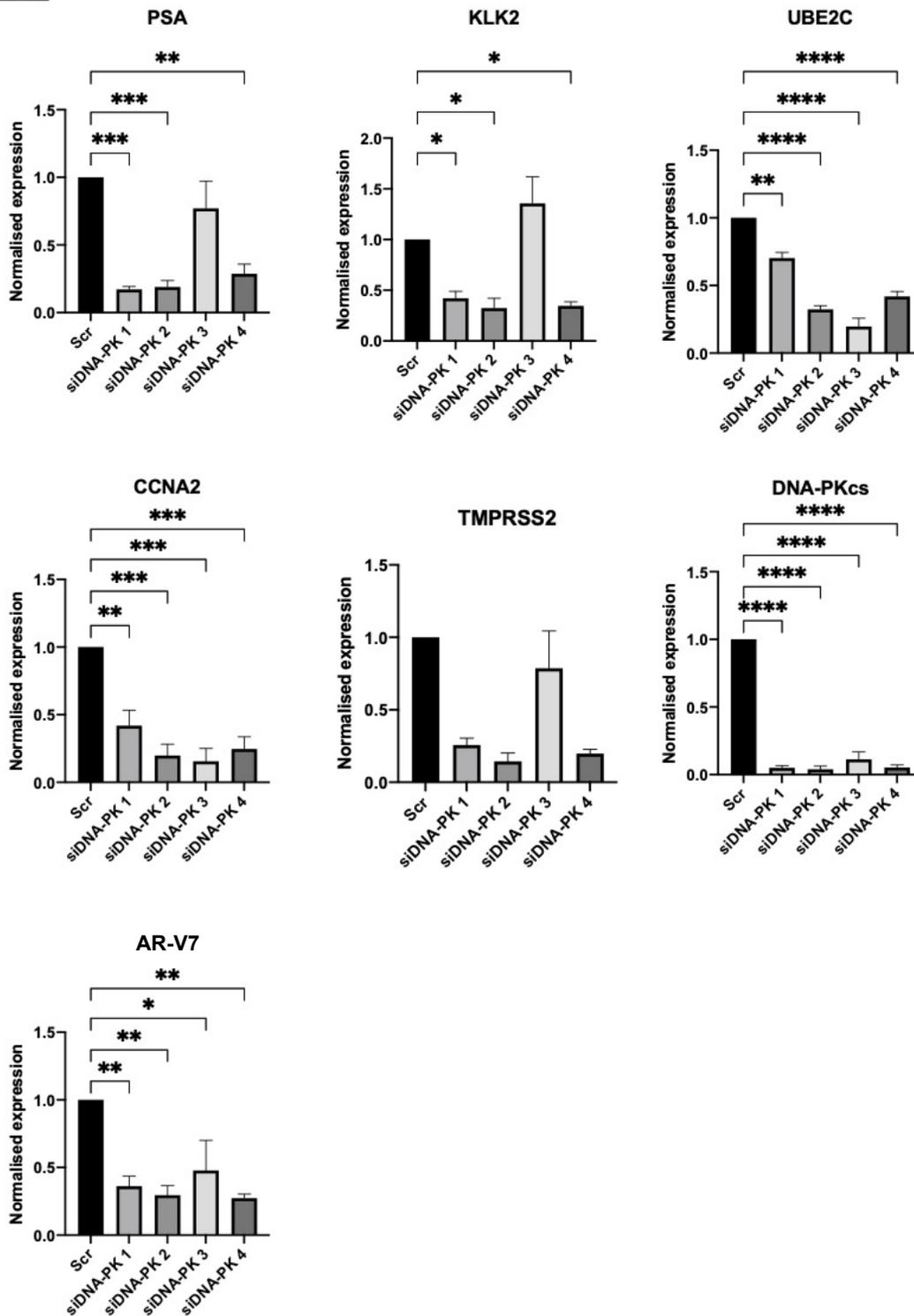
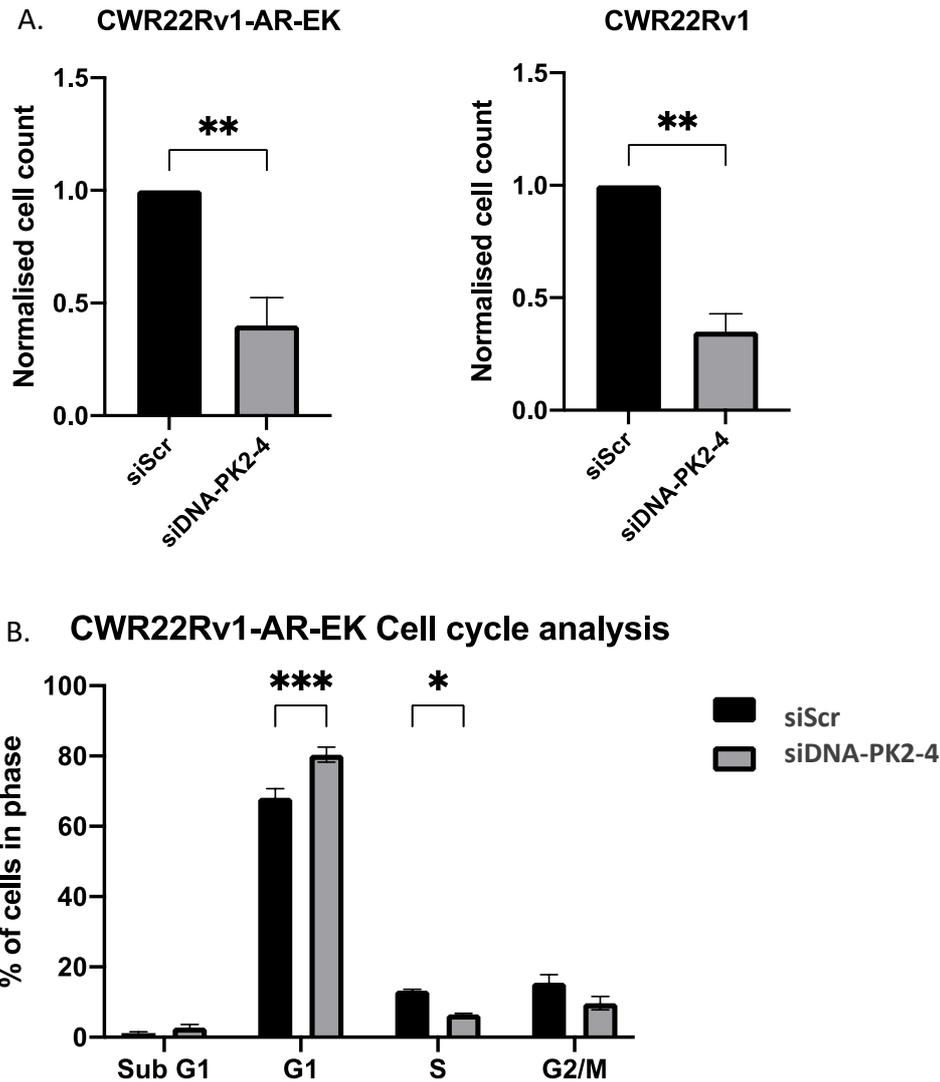


Figure 4.23: DNA-PKcs knockdown using the deconvoluted Dharmacon SmartPool showed differing effects on AR target gene expression in VCaP cells.

VCaP cells were seeded in steroid-depleted media and reverse transfected with oligonucleotides targeting DNA-PKcs (siDNA-PK1-4) or a scrambled (siScr) control for 72 hours. Cells were then subject

to AR target gene expression analysis using qRT-PCR. Data was normalised to the siScr treatment arm for each target gene. Data is representative of three independent repeats  $\pm$  SEM. One-way ANOVA using Bonferroni Post Hoc analysis was used to determine the statistical significance (\* =  $p < 0.05$ , \*\* =  $p < 0.01$ )

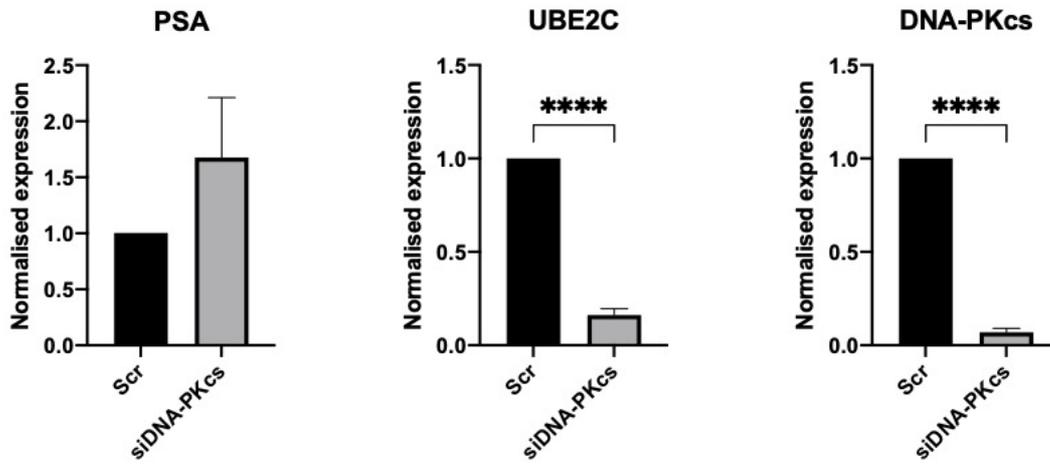
As each siRNA was having differing effects on growth and AR-V target gene expression, and because a pool of siRNA has advantages over the use of one single oligonucleotide, a pool of siDNA-PK2-4 was used in the subsequent experiments based on the consistent anti-proliferative effects of these reagents. The effect on proliferation and the cell cycle with the use of the siDNA-PK2-4 pool is shown in Figure 4.24, and Figure 4.25 shows the effect on AR target gene expression. siDNA-PK2-4 transfection caused a statistically significant reduction in proliferation and significantly impacted the cell cycle profile by impacting G1 and S phase. Importantly, the effect of DNA-PKcs depletion using siDNA-PK2-4 mimics, if not outperforms, the anti-proliferative effects of NU7441 treatment. My findings also indicate the potential hazard of utilising commercially available siRNA pools without deconvolution analyses. The modified oligonucleotide pool also consistently and significantly reduced expression of the AR-V-only regulated gene *UBE2C*, but not *PSA*, suggesting DNA-PKcs presence is particularly important in selective regulation of AR-V target genes. This siRNA pool will be used in an siRNA treatment arm for RNA-sequencing (experimental set up shown in Chapter 5).



**Figure 4.24 DNA-PKcs knockdown using siDNA-PK2-4 caused a significant decrease in proliferation of CWR22Rv1-AR-EK and CWR22Rv1 cells.**

A. CWR22Rv1-AR-EK cells were cultured in full media and CWR22Rv1 cells were cultured in steroid-depleted conditions after being transfected with siScr or siDNA-PK2-4. Cell counts were performed after 5 days. Data is representative of three independent repeats  $\pm$  SEM. An unpaired t-test was used to determine the statistical significance ( $*=p<0.05$ ,  $**=p<0.01$ ). B. CWR22Rv1-AR-EK cells were cultured in full media for 72 hours post transfection with siScr or siDNA-PK2-4. Cell cycle analysis was performed using PI flow cytometry. Data represents mean number of cells in each cell cycle phase  $\pm$  SEM. Two-way ANOVA with Sidak's correction was used to determine statistical significance ( $**=p<0.01$ ).

### CWR22Rv1-AR-EK



### CWR22Rv1

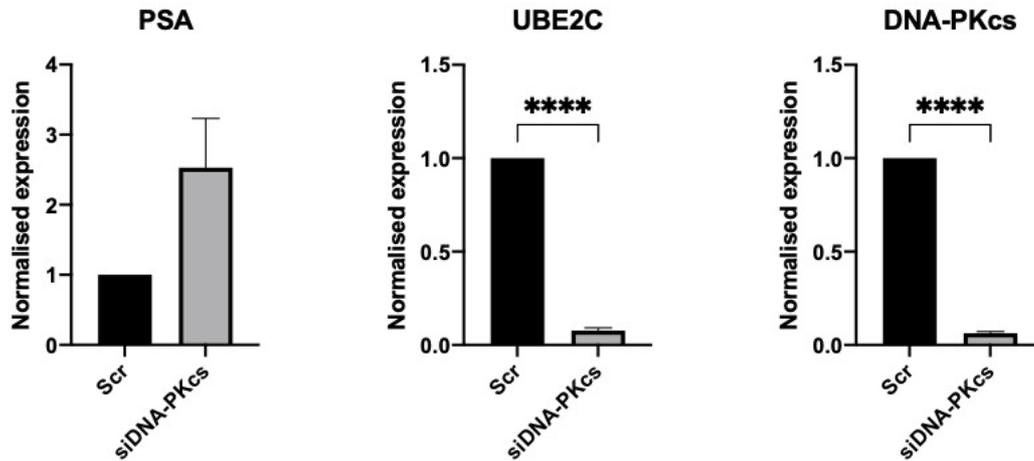


Figure 4.25 DNA-PKcs knockdown using siDNA-PK2-4 downregulated expression of *UBE2C*, but not *PSA*, in CWR22Rv1-AR-EK and CWR22Rv1 cell lines.

CWR22Rv1-AR-EK and CWR22Rv1 cells were reverse transfected with either siScr or siDNA-PK2-4 and incubated for 72 hours. AR target gene expression was analysed using qRT-PCR. Data represents the mean of three repeats  $\pm$  SEM. An unpaired t-test was used to determine the statistical significance (\*\*\*) =  $p < 0.0001$ ).

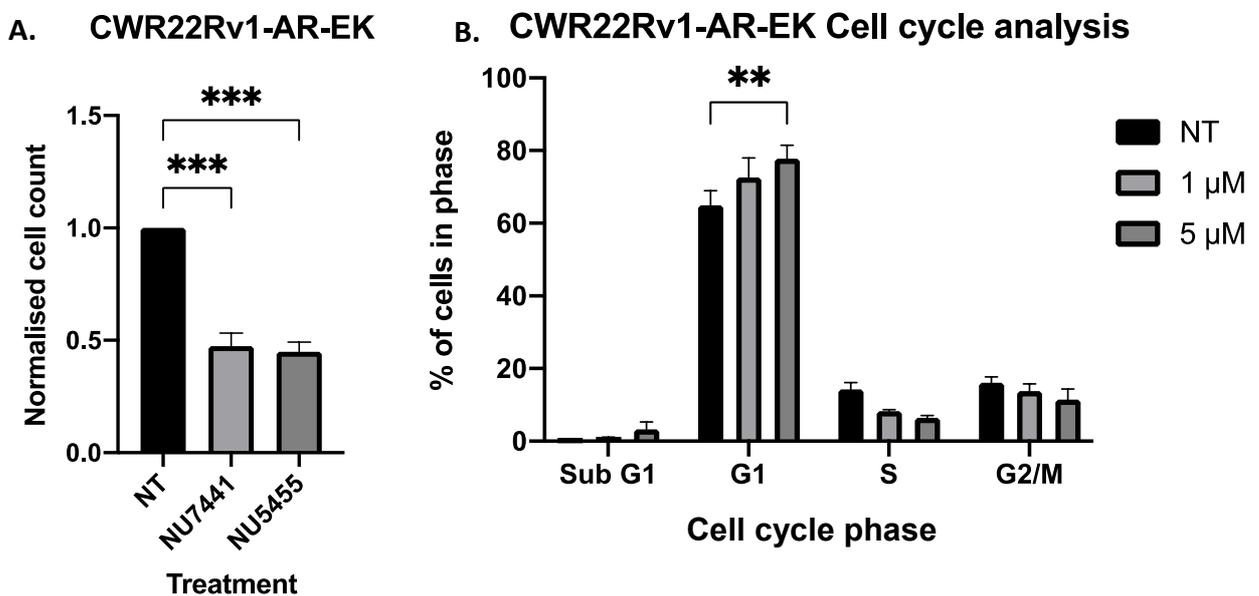
4.2.9 DNA-PKcs inhibition with a more selective inhibitor showed similar effects on proliferation and AR-V target gene expression in CWR22RV1-AR-EK cells

After demonstrating DNA-PKcs manipulation impacts AR-V transcriptional activity and proliferation of several PCa cell lines, it was key to determine if the impact of NU7441 was mirrored with the newer, more specific DNA-PKcs inhibitor, NU5455. NU5455 was used to exclude the possibility that the effects on cell proliferation and AR-V transcriptional activity of

the earlier generation compound were a consequence of off-target effects, as NU5455 has less impact on PI3K members than NU7441 (Willoughby *et al.*, 2020).

The anti-proliferative effects of 1  $\mu\text{M}$  NU5455 on CWR22Rv1-AR-EK cells were comparable to same dose NU7441 using a normalised cell count after 5 days. The cell cycle profile was also significantly impacted with an increase in the percentage of cells in G1 phase after 48 hours 5  $\mu\text{M}$  NU5455 (Figure 4.26). This was consistent with what was seen in response to DNA-PKcs knockdown (Figure 4.24). However, in response to 5  $\mu\text{M}$  NU7441, there was sub-G1 elevation rather than G1 arrest, though this is likely due to the off-target effects of high concentrations of NU7441.

AR target gene expression analysis showed significant reductions in *PSA*, *KLK2* and *UBE2C* in CWR22Rv1-AR-EK cells and significant reduction in *PSA* in VCaP cells when grown in SDM (Figure 4.27), which together with my findings using NU7441 strongly implicate DNA-PKcs as a co-regulator of AR-Vs.

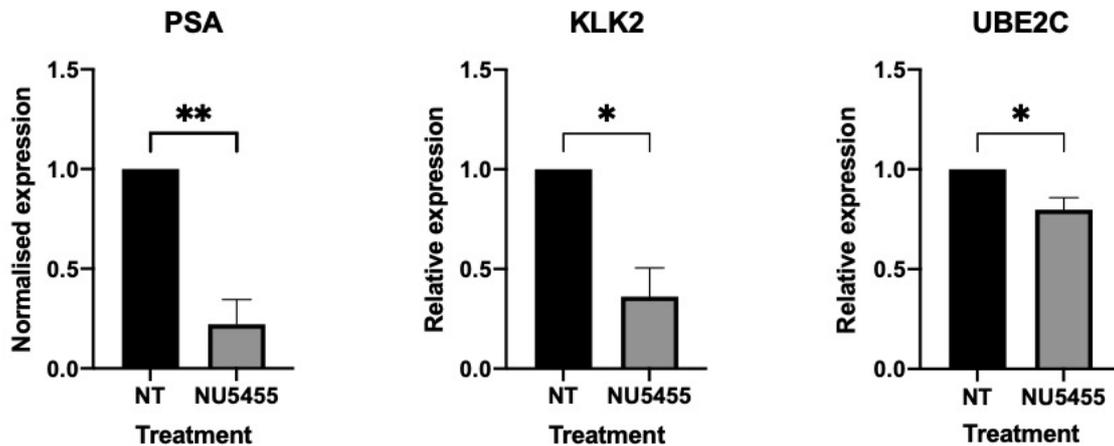


**Figure 4.26 DNA-PKcs inhibition using NU5455 causes a significant reduction in cell count as a consequence of G1 arrest.**

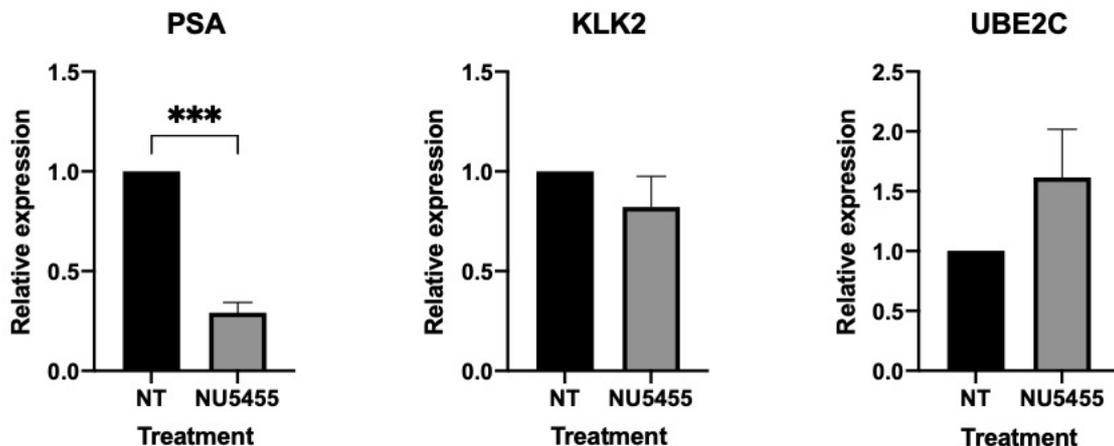
A. CWR22Rv1-AR-EK cells were cultured for 5 days in full media containing either DMSO, 1  $\mu\text{M}$  NU7441 or 1  $\mu\text{M}$  NU5455. Cell counts were then performed in quadruplicate. Data is representative of three independent repeats  $\pm$  SEM. One-way ANOVA using Bonferroni Post Hoc analysis was used to determine the statistical significance ( $*=p<0.05$ ,  $**=p<0.01$ ,  $***=p<0.001$ ). B. CWR22Rv1-AR-EK cells were cultured in full media containing either DMSO, 1  $\mu\text{M}$  or 5  $\mu\text{M}$  NU5455 for 48 hours. Cell cycle

analysis was performed using PI flow cytometry. Data represents mean number of cells in each cell cycle phase  $\pm$  SEM. Two-way ANOVA with Tukeys correction was used to determine statistical significance (\*\* =  $p < 0.01$ ).

**A. CWR22Rv1-AR-EK**



**B. VCaP**

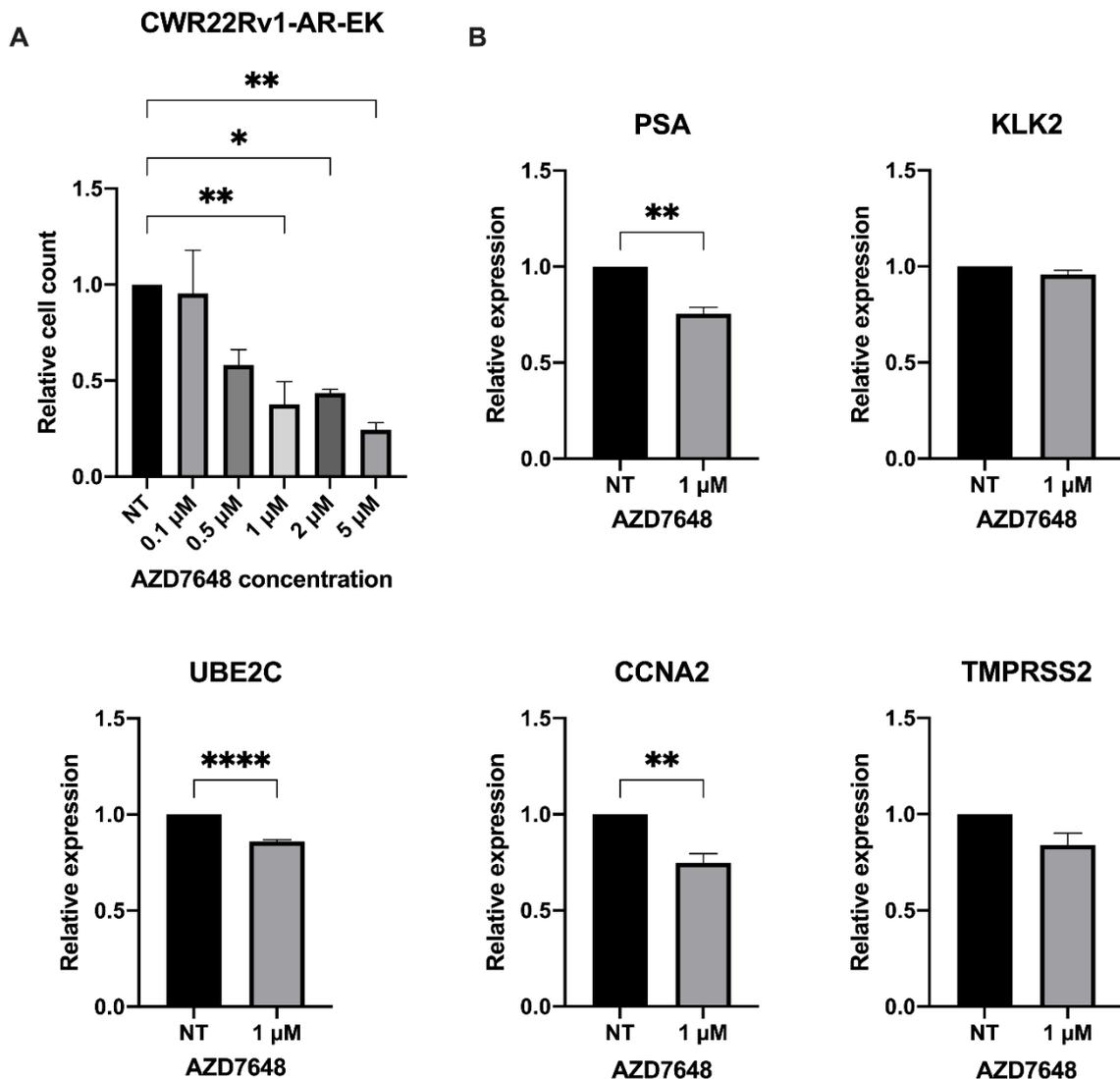


**Figure 4.27 DNA-PKcs inhibition with NU5455 caused a significant reduction in AR-V target genes in CWR22RV1-AR-EK cells and VCaP cells.**

A. CWR22Rv1-AR-EK cells were cultured in full media for 48 hours and then treated with 1  $\mu$ M NU5455 for 24 hours. B. VCaP cells were cultured in steroid-depleted media for 48 hours and then treated with 1  $\mu$ M NU5455 for 24 hours. Cells were then subject to AR target gene expression analysis using qRT-PCR. Data was normalised to the DMSO (NT) treatment arm for each target gene. Data is representative of three independent repeats  $\pm$  SEM. One-way ANOVA using Bonferroni Post Hoc analysis was used to determine the statistical significance (\* =  $p < 0.05$ , \*\* =  $p < 0.01$ ).

#### 4.2.10 DNA-PKcs inhibition with AZD7648 impacts proliferation and AR-V target gene expression in CWR22RV1-AR-EK cells

During the course of the project, an additional DNA-PKcs inhibitor AZD7648, developed by AstraZeneca, became commercially available (Fok *et al.*, 2019) and was incorporated into the study to further validate the role of DNA-PKcs in AR-V transcriptional regulation and help to overcome any batch variation that could have impacted outputs with the NU7441 and NU5455 compounds. Critically, AZD7648 had similar anti-proliferative effects to NU7441 and NU5455 and showed modest, but significant impact on AR-V target gene expression in CWR22Rv1-AR-EK cells.

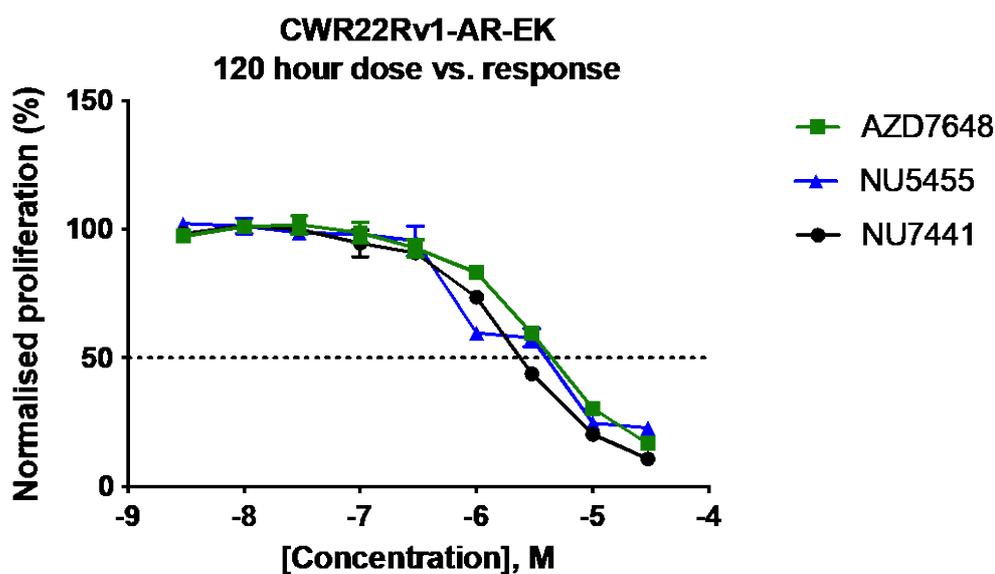


**Figure 4.28** DNA-PKcs inhibition with AZD7648 significantly reduces CWR22Rv1-AR-EK proliferation and several AR-V target genes.

A. CWR22Rv1-AR-EK cells were grown in full media with increasing concentrations of AZD7648 for 5 days before a manual cell count. Cell counts were performed in quadruplicate. B. CWR22Rv1-AR-EK cells were cultured in full media for 48 hours and then treated with 1 μM AZD7648 for 24 hours. Cells were then subject to AR target gene expression analysis using qRT-PCR. Data was normalised to the DMSO treatment arm for each target gene. Data is representative of three independent repeats ± SEM. One-way ANOVA using Bonferroni Post Hoc analysis was used to determine the statistical significance (\* =  $p < 0.05$ , \*\* =  $p < 0.01$ , \*\*\*\* =  $p < 0.0001$ ).

A direct comparison of the three DNA-PKcs targeting compounds was performed using an SRB assay. CWR22Rv1-AR-EK cells were seeded in 96-well plates and the following day, treated with a wide range of concentrations of NU7441, NU5455 and AZD7648. This confirmed DNA-

PKcs inhibition with NU7441 and NU5455 showed similar anti-proliferative effects, and AZD7648 showed a slightly reduced potency when comparing GI<sub>50</sub> values (Figure 4.29). AZD7648 had a GI<sub>50</sub> of 2.8 μM, NU7441 had a GI<sub>50</sub> of 1.8 μM, and NU5455 was the most potent, with a GI<sub>50</sub> of 1.4 μM.



Compound	AZD7648	NU7441	NU5455
GI <sub>50</sub> (μM)	2.8	1.8	1.4

Figure 4.29 CWR22Rv1-AR-EK cells are sensitive to DNA-PKcs inhibition with three DNA-PKcs inhibitors.

CWR22Rv1-AR-EK cells were treated with increasing concentrations of AZD7648, NU7441 and NU5455 for 120 hours before harvesting for an SRB assay to determine growth response to DNA-PKcs inhibition. Graph represents mean values ± SEM across three independent repeats as well as three technical replicates for each experimental arm. GI<sub>50</sub> values were determined using GraphPad PRISM software.

4.2.11 DNA-PKcs inhibition did not significantly downregulate PSA or KLK2 expression in steroid depleted conditions

The downregulation of AR-V target gene expression in CWR22Rv1-AR-EK cells grown in steroid proficient conditions was not mirrored in CWR22Rv1-AR-EK cells that were grown in steroid-

depleted conditions. The AR-V target genes *PSA* and *KLK2* were not downregulated by DNA-PKcs inhibition with NU7441 when grown in steroid-depleted conditions, in contrast to *UBE2C* and *CCNA2* which demonstrated significant reduction (Figure 4.30A). DNA-PKcs inhibition still significantly reduced cell number after 96-hour treatment, aligning with what was seen in steroid-proficient conditions (Figure 4.30B). It was hypothesised that the GR would be less active in these castrate-like conditions and DNA-PKcs inhibition could also be impacting on GR activity. Furthermore, similarly to AR-Vs, GR expression is upregulated following AR blockade and the GR and AR have been shown to regulate a subset of common target genes (Arora *et al.*, 2013). Therefore, in steroid-depleted conditions, diminished DNA-PKcs activity may not cause as significant a response as had been observed in steroid-proficient conditions. This led to some experiments in which the GR was knocked down using siRNA for 48 hours with a subsequent 24-hour DNA-PKcs inhibition using NU7441. However, GR knockdown did not lead to a more significant reduction in AR target genes, such as *PSA* (Figure 4.30C). This work requires some further investigation but could be a reason as to why DNA-PKcs inhibition in CWR22Rv1-AR-EK cells did not lead to a significant reduction in AR/AR-V target genes but did to the AR-V specific genes *CCNA2* and *UBE2C*.

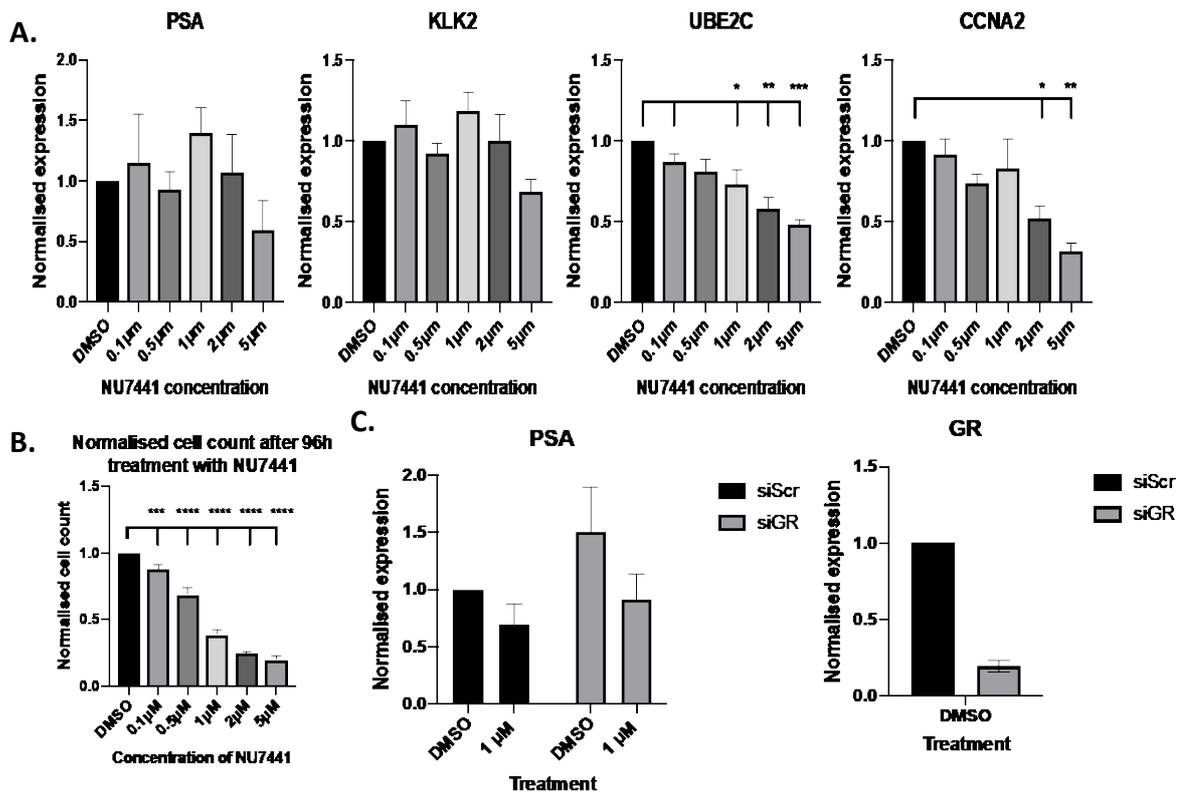


Figure 4.30 DNA-PKcs inhibition in CWR22Rv1-AR-EK cells grown in steroid-depleted media significantly reduced the expression of AR-V only target genes and proliferation.

CWR22Rv1-AR-EK cells were grown in SDM for 48 hours prior to increasing concentrations of NU7441 for 24 hours (A) or grown in SDM for 24 hours prior to 96h treatment with NU7441 (B). CWR22Rv1-AR-EK cells were treated with GR targeting or non-targeting siRNA for 48 hours prior to 24 hours 1  $\mu$ M NU7441 (C). Graphs show an average of three repeats  $\pm$  SEM (\* =  $p < 0.05$ , \*\* =  $p < 0.01$ , \*\*\* =  $p < 0.001$ , \*\*\*\* =  $p < 0.0001$ ).

### 4.3 Discussion

Although the gold standard treatment for advanced PCa patients using AR-targeting agents is initially successful, resistance is inevitable with the emergence of CRPC for which there remains no long-term, durable treatment options for patients. The AR continues to drive PCa progression into late-stage disease as a consequence of numerous molecular aberrations, including *AR* gene amplification or mutations that drive heightened AR activity despite therapeutic targeting with ADT and anti-androgen drugs. Furthermore, the generation of alternatively spliced AR-Vs help retain AR signalling and are a major resistance mechanism that is observed in patients undergoing hormone therapy (Sharp *et al.*, 2019). Critically, there remains no effective treatment options to block this signalling pathway. Due to the difficulty in generating compounds that directly target AR-Vs, indirect targeting of their co-regulators is a viable treatment option in CRPC. Therefore, there is a huge drive to better understand AR biology at this late stage of disease to facilitate the development of new therapies for patients.

DNA-PKcs is one of three proteins that compose the DNA-PK trimeric complex, along with Ku70 and Ku80. This complex is involved in the NHEJ pathway to repair DNA DSBs, that result from endogenous or exogenous insult, to maintain genomic stability.

Independent of its DDR related roles, DNA-PKcs has been shown to have roles in transcriptional regulation and has been identified as a FL-AR coregulatory protein in a number of studies (Goodwin *et al.*, 2015, Dylgjeri *et al.*, 2019). RNA sequencing in C4-2 PCa cells depleted of DNA-PKcs showed a downregulation in the androgen response, as well as DNA repair, cell cycle and pro-tumorigenic pathways. A recent report has also demonstrated a novel role for DNA-PKcs in regulating metabolism and how this could link to cancer progression (Dylgjeri *et al.*, 2022). Here, the first DNA-PKcs interactome was determined in the PCa cell line C4-2 using RIME. This identified several glycolytic enzymes that interacted with DNA-PKcs. Subsequent metabolomic analysis in response to NU7441-mediated DNA-PKcs blockade revealed numerous alterations to metabolite levels suggesting DNA-PKcs modulates cancer metabolism, further expanding its pleiotropic role in supporting CRPC progression.

In a study by Kothari *et al.*, the effect of DNA-PKcs inhibition *in vivo* using an LNCaP-AR xenograft was assessed. Here, NU7441 treatment caused a 44% reduction in tumour volume after 6 weeks, with minimal changes in body weight (Kothari *et al.*, 2019).

Given the role of DNA-PKcs in the DDR, there is a robust rationale and strong pre-clinical evidence to support the application of DNA-PKcs inhibitors in the clinical setting to sensitise cancer cells to radio- and chemotherapy. Consistent with this, AZD7648 was tested *in vivo* to investigate if DNA-PKcs blockade radio-sensitised syngeneic mouse models (Nakamura *et al.*, 2021). Strikingly, 75-100% of animals treated with radiotherapy plus AZD7648 in a MC38 colon cancer syngeneic mouse model achieved complete tumour regression and when rechallenged with tumour implantation, there was a significant delay in tumour formation. This suggests this combination is not only able to inhibit tumour growth but provides long term anti-tumour immunologic memory, because immune-compromised mice did not exert this same response. This was shown to be dependent on CD8<sup>+</sup> T-cells and the type 1 interferon response (Nakamura *et al.*, 2021). DNA-PKcs inhibitors have progressed to early phase clinical trials. Both AZD7648 and Pepsertib (M3814) are in trials to evaluate safety and tolerability which are currently ongoing (M3814 - NCT02516813, AZD7648 - NCT03907969).

Despite the relatively extensive research around DNA-PKcs in PCa pre-clinical models, an AR-V driven PCa model has not yet been studied to determine if DNA-PKcs is also a co-regulator of AR-Vs and is therefore a therapeutic target in AR-V positive CRPC patients. In this project, the aim was to investigate the co-regulatory role of DNA-PKcs on AR-Vs in PCa. Several PCa cell line models were used that express AR-Vs to model the castrate-resistant stage of the disease.

Analysis of publicly available patient data was performed to give an indication of how common *PRKDC* gene alterations are in PCa and CRPC. This showed that 18% of PCa patients have alterations in the *PRKDC* gene, and the proportion of patients with *PRKDC* alterations increased to 27% when analysing CRPC patient data. Furthermore, DNA-PKcs expression was also statistically significantly correlated with Gleason score (Figure 4.1). These findings indicate DNA-PKcs is potentially involved in the progression of PCa. Analysis of TCGA data, by fellow PhD student Nicholas Brittain, has shown *PRKDC* expression is significantly upregulated in matched tumour versus normal patient samples (Figure 4.2). Analysis of published

microarray data also revealed PRKDC expression is significantly upregulated in CRPC patients compared to localised PCa patients. Interrogation of patient data that includes AR-V7 status, indicated PRKDC expression was not significantly altered between AR-V7 positive versus negative patients. One possible reason for this could be that the AR-V7 information is limited to whether AR-V7 mRNA is present or not, and does not include any quantification of AR-V7 levels. Therefore, for future analyses it would be beneficial to retrieve this data to stratify patients into high and low AR-V7 expressing patients and then determine PRKDC expression changes between those patients.

The proliferative effects of DNA-PKcs inhibition using NU7441 was determined in several PCa cell lines, either grown in steroid-depleted conditions to encourage the emergence of AR-Vs (CWR22Rv1 and VCaP cells), or in the CWR22Rv1-AR-EK cell line that expresses only AR-Vs. In line with previous findings, DNA-PKcs inhibition caused a significant reduction in cell growth after 5 days that was dose-dependent. This was then validated with two newer, more selective DNA-PKcs inhibitors, NU5455 and AZD7648. The results with NU5455 and AZD7648 were comparable to the results observed with NU7441, suggesting the impact of the first-generation agent was not due to off-target effects at the dose range used.

Transcriptional activity of AR-Vs was assessed after DNA-PKcs inhibition using qRT-PCR to determine relative levels of AR/AR-V target genes. To study AR-V driven transcription, the cell lines CWR22Rv1 and VCaP were grown in steroid-depleted conditions to abrogate FL-AR activity. CWR22Rv1-AR-EK cells were also used as they do not express FL-AR protein and their canonical AR transcriptome is exclusively driven by AR-Vs so are a valuable model to study AR-Vs in isolation. Twenty-four hour DNA-PKcs inhibition caused a significant reduction in AR/AR-V target gene expression in multiple cell lines, which mimicked previous findings focusing on FL-AR-DNA-PKcs interplay; that showed AR target genes such as *PSA* and *TMPRSS2* were downregulated upon inhibition of DNA-PKcs catalytic activity. Critically, the AR-V driven genes *UBE2C* and *CCNA2* were downregulated in response to DNA-PKcs inhibition, indicating DNA-PKcs regulates AR-Vs as well as FL-AR. These findings suggest DNA-PKcs is a co-regulator of AR-V transcriptional activity. However, this was not replicated in CWR22Rv1-AR-EK cells that were grown in steroid-depleted conditions; with AR-V target genes *PSA* and *KLK2* not downregulated by NU7441, whereas *UBE2C* and *CCNA2* were significantly downregulated (Figure 4.30A). Proliferation of CWR22Rv1-AR-EK cells was reduced after 96-hour treatment,

aligning with what was seen in steroid-proficient conditions (Figure 4.30B). The GR may be less active in steroid-depleted conditions and DNA-PKcs inhibition could affect GR activity, therefore, in steroid-depleted conditions, diminished DNA-PKcs activity may not cause as significant of a response as had been observed in steroid-proficient conditions. Preliminary experiments to investigate this incorporated GR knockdown for 48 hours with subsequent 24-hour DNA-PKcs inhibition. However, GR knockdown did not lead to a more significant reduction in AR target genes such as *PSA* in response to DNA-PKcs inhibition (Figure 4.30C). Another preliminary study was carried out to investigate if DNA-PKcs inhibition also caused a reduction in GR transcriptional activity by profiling some GR target gene expression. DNA-PKcs inhibition did not appear to consistently cause a reduction in GR target gene expression. This hypothesis could be the reason why DNA-PKcs inhibition in CWR22Rv1-AR-EK cells did not lead to a significant reduction to AR/AR-V target genes *PSA* and *KLK2* in steroid-depleted conditions, but did downregulate the AR-V specific genes *CCNA2* and *UBE2C*. However, further experiments are required to confirm this.

AR-V protein levels were reduced upon DNA-PKcs inhibition in VCaP cells, indicating DNA-PKcs could regulate transcription of AR-Vs. The AR can regulate its own expression via recruitment to repressive elements in intron 2 of the *AR* gene (Cai *et al.*, 2011). To interrogate if DNA-PKcs was causing a build-up of AR at these regions and therefore repressing AR-V expression, an AR ChIP in VCaP cells grown in SDM was performed. This revealed DNA-PKcs depletion does not significantly impact AR-V levels at three separate regions around the downstream response elements. However, because this was not using an AR-V specific antibody, and although the cells were in steroid depleted conditions, it is not possible to determine if this was FL-AR or primarily AR-V enrichment at these loci. An additional study incorporating enzalutamide to further block FL-AR chromatin binding would have been useful but was not possible due to time constraints. In an effort to further understand the role of DNA-PKcs in controlling AR-V abundance in PCa cells, I investigated if DNA-PKcs inhibition was causing enhanced turnover of AR-V protein by the proteasome. To do this, VCaP cells were co-treated with NU7441 plus the proteasome inhibitor carfilzomib to determine if this would rescue the AR-V protein levels. In response to single agent carfilzomib treatment, AR-V protein levels were increased, possibly because of a build-up of mono-ubiquitylated forms of the protein that are not degraded by the proteasome. In response to NU7441 there was no change the

levels of these AR isoforms. Importantly, because transcription of AR-Vs is decreased in response to carfilzomib, as evidenced by decreased AR-V7 transcript levels (Figure 4.10), the western data suggests a degree of stabilisation of the higher molecular weight form of AR-V proteins at the expense of the unmodified forms. Work conducted by another laboratory member has characterised that the decrease in transcription of AR-Vs in response to proteasome inhibitor treatment is due to a build-up of AR-Vs at the AR downstream repressive element, previously discussed. The mechanism by which AR-V expression is impacted upon DNA-PKcs manipulation is interrogated further in Chapter 5.

The chromatin biology surrounding DNA-PKcs-AR-V interplay showed DNA-PKcs is recruited to *cis*-regulatory elements of AR and AR-V target genes in several PCa cell lines. CHIP experiments were used to investigate (i) AR-V recruitment to target genes and if this is dependent on DNA-PKcs and (ii) DNA-PKcs recruitment to AR/AR-V target genes to determine if DNA-PKcs presence is required for AR activity and if it is dependent upon AR for chromatin binding at the *cis*-regulatory regions of target genes. In the Goodwin *et al* study, DNA-PKcs recruitment to AR target genes in C4-2 cells was shown to be dependent on AR activation with DHT, as DNA-PKcs recruitment at *PSA* and *TMPRSS2* enhancers significantly increased 6-24 hours post DHT treatment (Goodwin *et al.*, 2015). To interrogate this mechanism in relation to AR-Vs, CWR22Rv1-AR-EK cells were used for AR and DNA-PKcs CHIP. AR-V recruitment to AR target genes was not significantly impacted upon NU7441 treatment at the *PSA*, *KLK2*, *UBE2C* and *CCNA2* enhancers, indicating DNA-PKcs kinase activity is not critical for AR-V recruitment to chromatin to activate transcription. This was reflected in VCaP cells with DNA-PKcs knockdown. However, because AR chromatin binding is not changed upon DNA-PKcs inhibition, that does not necessarily mean AR-Vs are able to actively transcribe genes in the absence of DNA-PKcs kinase activity. This could be investigated by interrogating phospho-RNA polymerase II CHIP to determine if active transcription is impacted by DNA-PKcs manipulation at the regions of interest. DNA-PKcs was robustly recruited to the enhancers of several AR-V target genes, that was modestly dependent on DNA-PKcs kinase activity as inhibition with NU7441 caused a statistically significant, but modest reduction in recruitment to the *PSA* enhancer but only a slight reduction in DNA-PKcs recruitment to *KLK2*, *UBE2C* and *CCNA2*. To determine if this was dependent on AR-V recruitment, the DNA-PKcs CHIP was repeated with AR-V knockdown using an AR exon 1 targeting siRNA. This slightly reduced DNA-PKcs although

this was not statistically significant. This indicates DNA-PKcs recruitment to AR-V target genes in AR-V driven cell lines is not as dependent as DNA-PKcs is on FL-AR in hormone-proficient conditions. To provide a genome wide investigation of DNA-PKcs recruitment to regulatory regions of genes, ChIP-sequencing could be performed. Unfortunately, this is under-reported in the literature, with only one report of DNA-PKcs ChIP-sequencing evident; using both phospho-DNA-PKcs and one DNA-PKcs ChIP-sequencing experiment in MCF7 cells in response to oestrogen (Liu *et al.*, 2014b, Bunch *et al.*, 2015). This paucity of information may indicate difficulties in achieving reproducible ChIP experiments using the currently available anti-DNA-PKcs antibodies. If successful though, DNA-PKcs ChIP-sequencing in prostate cells could be compared to AR/AR-V ChIP-sequencing to determine the level of overlap between DNA-PKcs and AR regulated genes and further support the notion of DNA-PKcs as an AR-V co-regulator.

To validate the effect of DNA-PKcs inhibition in CRPC cell line models, DNA-PKcs depletion using siRNA was performed. Firstly, a DNA-PKcs smartpool consisting of 4 individual siRNA sequences was used to determine the effect on AR-V target gene expression and proliferation of cell lines. Surprisingly, this did not reflect what was seen in response to DNA-PKcs inhibition, with expression of several AR target genes increasing after 48 and 72 hours. In VCaP cells, however, AR target gene expression and AR-V7 and FL-AR transcript levels were all significantly downregulated after 48- and 72-hour DNA-PKcs depletion. Furthermore, DNA-PKcs depletion caused a slight decrease in proliferation in CWR22Rv1 and CWR22Rv1-AR-EK cells, and a significant decrease in the VCaP cell line. The reason for the disparity between the cell lines and increased sensitivity the VCaP cells display in response to DNA-PKcs blockade remains largely unknown. However, I speculate that because of the AR gene amplification present in VCaP cells, making the cell line more heavily reliant upon AR/AR-V signalling, this could make them more sensitive to DNA-PKcs depletion. Alternatively, if there are off-target effects which could compromise the effect of DNA-PKcs depletion, CWR22Rv1 and CWR22Rv1-AR-EK cells may be more effected by this and that is why they show reduced impact of DNA-PKcs depletion.

To investigate why DNA-PKcs inhibition with the smartpool was causing unexpected results, a deconvoluted siRNA screen of the four oligonucleotides was performed to determine the individual responses. In all three cell lines tested, each siRNA showed different effects on AR-V target gene expression and proliferation. For example, siDNA-PK1 showed minimal

proliferative impact in CWR22Rv1-AR-EK, CWR22Rv1 and VCaP cells, or on AR-V target gene expression in CWR22Rv1-AR-EK and CWR22Rv1 cells. In contrast, individual siDNA-PKs 2-4 all caused a reduction in proliferation across the panel of cell lines as well as reductions in AR/AR-V target gene expression which aligned well to the effects observed using DNA-PKcs inhibitors. Because using a pool of siRNA is beneficial due to the reduced chances of off-target effects due to the lower concentration used for each of the oligonucleotides, a new siDNA-PKcs pool was generated, consisting of siDNA-PKs 2-4 and tested in subsequent experiments to interrogate DNA-PKcs depletion in CRPC. This was to decipher kinase-dependent and kinase-independent roles of DNA-PKcs in the panel of cell lines. The siDNA-PK2-4 pool of siRNAs showed significant down regulation of the AR-V target genes *UBE2C* and *CCNA2* and demonstrated robust anti-proliferative effects in the CRPC cell lines.

When the newer more selective DNA-PKcs inhibitors became available, these were compared to NU7441. SRB assays using NU5455, AZD7648 and NU7441 were performed to do a direct comparison of GI<sub>50</sub> values. This showed NU5455 was the most potent compound, followed by NU7441 then AZD7648. Also, the proliferative and cell cycle effects upon NU5455 treatment was compared to NU7441. This showed at 5 µM, NU7441 significantly increased the percentage of cells that are in the sub-G1 phase, suggesting these cells have an apoptotic phenotype. Whereas 5 µM NU5455 caused the percentage of cells in the G1 phase to increase and the cells in S phase decrease, without a notable induction of sub-G1, which was consistent with what was seen upon DNA-PKcs knockdown. This could suggest this is the specific effect of DNA-PKcs manipulation, whereas the 5 µM dose of NU7441 causing a large proportion of cells to be in the sub-G1 phase is likely due to the off-target effects of the compound at high doses.

The two newer compounds also significantly impacted AR/AR-V target gene expression. This validated the effects observed with NU7441 was due to DNA-PKcs inhibition and not due to off-target effects of this first-generation agent, such as PI3K inhibition which has been reported at higher compound doses (Willoughby *et al.*, 2020).

Due to the promising effects of DNA-PKcs manipulation on AR-V transcriptional activity, a more global investigation of the transcriptional effects of DNA-PKcs inhibition and knockdown using RNA-sequencing was required. These experiments would comprehensively investigate

differential gene expression and identify additional pathways that are altered in PCa cells, in response to compromised DNA-PKcs function that could further expand our understanding of the pleiotropic roles of DNA-PKcs in advanced PCa.

## Chapter 5 Investigating global transcriptomic effects of DNA-PKcs inhibition and knockdown in CRPC

### 5.1 Introduction

As described in Chapter 4, DNA-PKcs inhibition and knockdown leads to a reduction in AR target gene expression, suggesting DNA-PKcs plays a role in the regulation of AR-V transcriptional activity. Previously published microarray and RNA-sequencing datasets have shown DNA-PKcs regulates several other oncogenic pathways in PCa cell lines providing compelling evidence of its importance as a target in PCa. However, a limitation of these studies is that they have been conducted exclusively in FL-AR-driven cell lines and not in a specific AR-V driven cell line. The role of DNA-PKcs in controlling AR-Vs is a key knowledge gap that by filling, would help better understand its role in AR-V biology and potentially highlight it as a viable therapeutic target in AR-V positive CRPC patients.

Previous studies exploring the role of DNA-PKcs on global gene expression in the PCa cell line model C4-2 have demonstrated several pathways, such as Wnt signalling, MYC and E2F, are significantly downregulated upon DNA-PKcs blockade (Kothari *et al.*, 2019, Dylgjeri *et al.*, 2019). The published AR-V transcriptome from depletion of AR-Vs using an AR exon 1 targeting siRNA in the CWR22Rv1-AR-EK cell line was analysed alongside published microarray and RNA-sequencing data in C4-2 cells that were treated with 1  $\mu$ M NU7441 for 24 hours (Kounatidou *et al.*, 2019, Dylgjeri *et al.*, 2019, Goodwin *et al.*, 2015). The genes that are significantly differentially expressed (FC > 1.5 and adjusted p value < 0.05) in response to AR-V knockdown in CWR22Rv1-AR-EK cells or NU7441 treatment in C4-2 cells from microarray and RNA-sequencing were compared. This revealed that out of 500 genes differentially expressed in response to DNA-PKcs inhibition, 19% are part of the AR-V transcriptome (Figure 5.1). Given that the DNA-PKcs experiments were carried out in C4-2 cells that express AR-Vs at a lower level than CWR22Rv1 and CWR22Rv1-AR-EK cells, and that the experiment was carried out in hormone-proficient conditions where AR-Vs will not be particularly active, this overlap was quite significant. This provided the rationale to carry out RNA-sequencing in the CWR22Rv1-AR-EK cell line to determine the overlap of AR-V- and DNA-PKcs-regulated genes in an AR-V driven cell line model.

**CWR22Rv1-AR-EK** (Kounatidou *et al.*, 2019)    **C4-2** (Goodwin *et al.*, 2015, Dylgjeri *et al.*, 2019)



**Figure 5.1 A notable number of genes in response to DNA-PKcs inhibition in C4-2 cells overlap with the AR-V transcriptome in CWR22Rv1-AR-EK cells.**

Significantly differentially-expressed genes in response to AR-V knockdown in CWR22v1-AR-EK cells and NU7441 treatment in C4-2 cells (Fold change > 1.5 and p value < 0.05). Raw counts were downloaded from GEO (Dylgjeri, GSE116765, Goodwin, GSE63480) and analysed for differential gene expression using DESeq2 (AREx1 and NU7441 (2019)) for RNA-sequencing and Limma for the microarray (NU7441(2015)).

The generation of AR-Vs, such as AR-V7, arise due to alterations in splicing of *AR* pre-mRNA. This leads to the incorporation of CEs that translate to truncated versions of the FL-AR protein that lack the LBD while retaining the transcriptionally active NTD. It is not well-understood how these CEs are incorporated into mature transcripts. The splicing machinery, or spliceosome consists of small nuclear ribonucleoproteins and core proteins, and the inclusion or exclusion of exons and exclusion of introns is influenced by various splicing factors. As previously discussed, there has been several splicing factors identified that play a role in the generation of alternatively spliced AR-Vs, such as Sam68, SFPQ, U2AF65 and ASF/SF2 (Stockley *et al.*, 2015, Takayama *et al.*, 2017, Liu *et al.*, 2014a). Interestingly, there have been indications of DNA-PKcs regulating splicing-related genes from the two aforementioned microarray and

RNA-sequencing datasets in Figure 5.1. The authors analysed differentially expressed genes (DEGs) for pathways that were enriched or de-enriched in response to DNA-PKcs manipulation. Interestingly, this revealed that the spliceosome gene list was significantly downregulated in both data sets. This would be interesting to interrogate further as this could provide information on whether DNA-PKcs controls genes that are critical in regulating alternative splicing to enable AR-V generation in PCa.

#### 5.1.1 Aims

Given what has been shown in the previous chapter, that DNA-PKcs inhibition and knockdown leads to down regulation of several AR/AR-V target genes and decreased proliferation of several PCa cell lines, it was important to further interrogate the mechanisms behind this response. It was also important to determine the global transcriptomic effects of DNA-PKcs inhibition as well as knockdown to distinguish between kinase and kinase-independent roles of DNA-PKcs in transcriptional regulation.

To this end, two key studies will be undertaken:

- i. RNA-sequencing analysis will be conducted in CWR22Rv1-AR-EK cells compromised for DNA-PKcs function by either siRNA-mediated depletion or inhibition using NU7441, NU5455 or AZD7648 to provide a vital DNA-PKcs transcriptome in AR-V-driven PCa
- ii. Provide key mechanistic insight into DNA-PKcs-mediated regulation of AR-V abundance in PCa cells by interrogating publicly available datasets and splicing of AR-Vs

## 5.2 Specific materials and methods

### 5.2.1 RNA-sequencing experimental set-up

Both DNA-PKcs inhibition and siRNA-mediated knockdown experimental arms were included in the RNA-sequencing experiment. CWR22Rv1-AR-EK cells were seeded at a density of  $5 \times 10^5$  in 6 mm dishes and reverse transfected with either siScr or siDNA-PKcs siRNAs. In the DMSO, NU7441, NU5455 and AZD7648 samples, non-targeting scrambled siRNA oligonucleotides were transfected into the cells and in the DNA-PKcs knockdown samples, the siDNA-PK2-4 pool was transfected. 48 hours post siRNA transfection, DMSO or one of the three DNA-PKcs compounds were added to the media at a  $1 \mu\text{M}$  dose for 24 hours before subsequent RNA isolation using the protocol described in section 3.3.

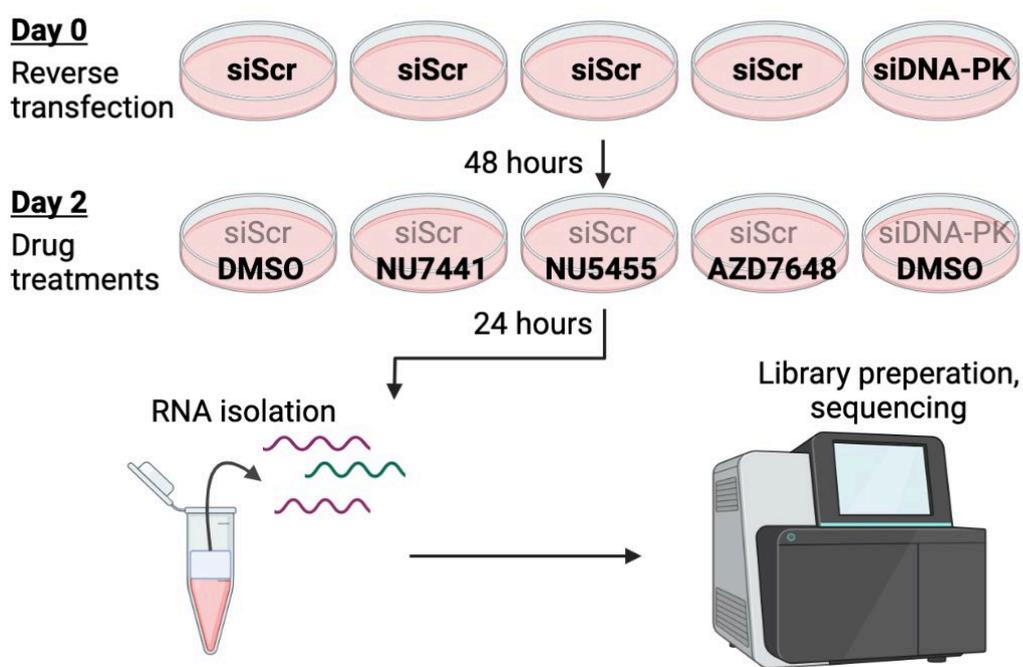


Figure 5.2 RNA-sequencing experimental set up to determine the global transcriptomic effect on CWR22Rv1-AR-EK cells upon DNA-PKcs inhibition with NU7441, NU5455 and AZD7648 and knockdown with an siRNA pool of siDNA-PK2-4.

### 5.2.2 RNA-sequencing analysis

RNA-sequencing was carried out at Genewiz using the Illumina NovaSeq, 2x150bp configuration, with an estimated data output of  $\sim 20\text{M}$  raw paired-end reads per sample. Raw sequencing reads were quality checked using FastQC and MultiQC to ensure the data passed the necessary QC requirements prior to analysis, such as adapter content (Andrews, 2010).

Reads were then aligned to the human genome using STAR (Spliced Transcripts Alignment to a Reference) and then gene level counts were generated using featureCounts (Dobin *et al.*, 2013). Principle component analysis (PCA) was used to show clustering of the biological replicates, to ensure all samples could be used in downstream analysis. DESeq2 was then used to determine the FC and p values of genes that were altered upon DNA-PKcs inhibition using any of the 3 compounds or upon DNA-PKcs depletion using siRNA (Love *et al.*, 2014). Stringent cut offs were then applied so final lists of DEGs included genes with an adjusted p value of < 0.05 and had a log2 fold change (log2FC) of more or less than 0.58 (1.5-fold change), which are hereby referred to as significant DEGs (SDEGs).

### 5.2.3 Gene set enrichment analysis

Gene set enrichment analysis (GSEA) was performed using the complete lists of DEGs without any filtering for p value or FC (Subramanian *et al.*, 2005). These lists were ranked based on FC and input into the Broad Institute gene set enrichment tool. The lists were run against Hallmark and Kyoto Encyclopaedia of Genes and Genome (KEGG) gene sets to determine if differentially expressed genes significantly enrich genes involved in certain pathways. The resulting pathways were ranked based on their Normalised Enrichment score (NES). Interpretation of an enrichment plot is presented in Figure 5.3.

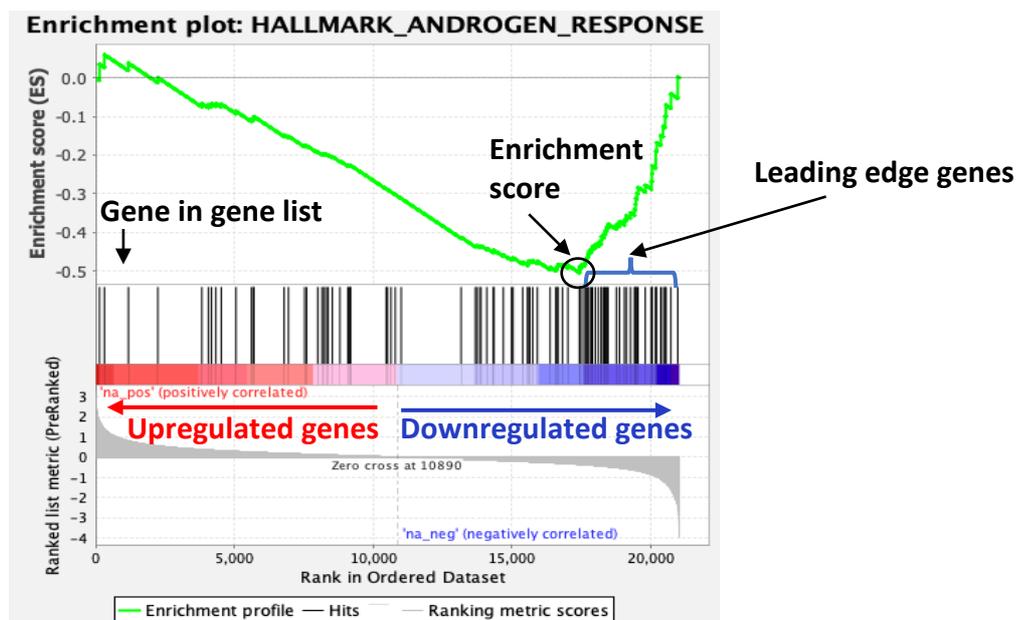


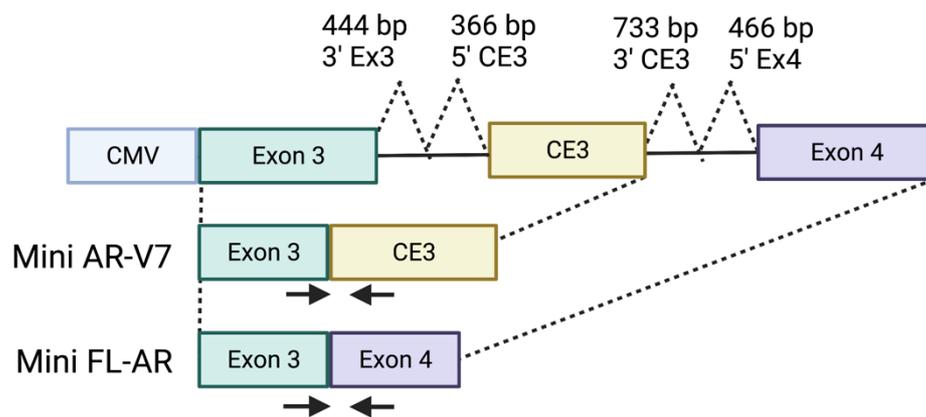
Figure 5.3 Interpretation of an enrichment plot from gene set enrichment analysis

#### 5.2.4 Immunoprecipitation

Immunoprecipitation (IP) was used to validate specificity of antibodies used for RNA immunoprecipitation. CWR22Rv1-AR-EK cells were reverse transfected on the day of seeding with 5 µg GFP-RBMX plasmid in 90 mm dishes, and then re-transfected 48 hours later. 24 hours after the second transfection, cells were trypsinised, pelleted and lysed in 1 mL of RIPA lysis buffer (Table 3.9) containing protease inhibitors (Roche) on ice for 30 mins. Samples were then sonicated using a Biorupter (Diagenode) on low for 3 x 30 seconds on/off cycles and then centrifuged at 15,000 x g for 15 minutes at 4°C. Supernatants were transferred to fresh Eppendorf tubes and 50 µL was taken from each sample as an input sample. 20 µL of protein G-Sepharose (PGS) beads were washed in RIPA lysis buffer (Table 3.9) and then added to remaining supernatants to pre-clear the lysates of non-specific proteins that bind to the beads and incubated at 4 °C with rotation for 1 hour. PGS beads were removed by centrifugation at 500 x g for 5 minutes and transfer of pre-cleared supernatant to fresh Eppendorf tubes, before adding 2 µg of anti-GFP (Abcam) or isotype IgG control antibodies. The following day, 20 µL PGS beads were added to the antibody-lysate samples to bind antibody-protein complexes and incubated for 1-2 hours at 4 °C with rotation. Samples were then centrifuged at 500 x g for 5 minutes to pellet the beads and the supernatant was removed. Bead-antibody-protein complexes were washed twice with wash buffer A (PBS, 0.2% Triton X-100, and 350 mM NaCl), centrifuged at 500 x g for 5 minutes and washed once in wash buffer B (PBS and 0.2% Triton X-100). PGS beads were centrifuged at 500 x g for 5 minutes, the supernatant was removed then 50 µL SDS sample buffer (Table 3.10) was added to the beads and boiled for 10 minutes at 100 °C.

#### 5.2.5 Minigene-based splicing analysis

PC3 cells were reverse transfected with 1 µg AR-V7 CMV-driven minigene reporter construct made up of exons 3, AR-V7-encoding CE3 and exon 4 separated by short stretches of adjoining intronic sequences. 8 hours post plasmid transfection, siRNA oligonucleotides were forward transfected into cells to a final concentration of 25 nM, as described in section 3.2, followed by a 72-hour incubation at 37 °C. RNA was then extracted as described in section 3.3, reverse transcribed (section 3.4) and qRT-PCR analysis was performed (section 3.5) using the primer pairs shown in Figure 5.4 and Table 3.6.



**Figure 5.4** Minigene construct diagram with the resulting mRNA for mini-AR-V7 and mini-FL-AR with the primer pair locations for each product.

#### 5.2.6 Differential splicing analysis

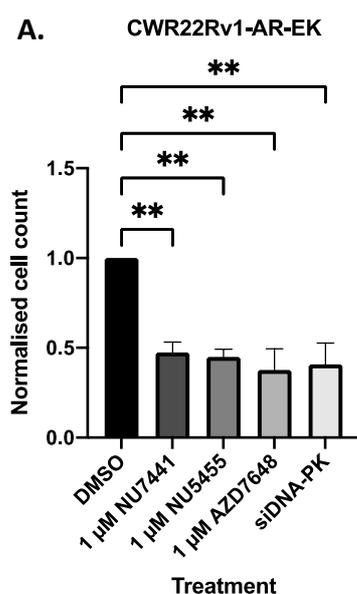
RNA-sequencing data was analysed using SUPPA2 to detect differential splicing events in response to DNA-PKcs knockdown (Trincado *et al.*, 2018). This was carried out by Graham Smith at the Newcastle University Bioinformatics support unit. The complete lists of differential splicing events plus a volcano plot of differential splicing events with the significant genes highlighted was returned, and I then quantified the number of splicing events in each category (e.g., alternative first exon or skipped exon) and applied a significance cut-off. These lists were input into PRISM to generate pie-charts.

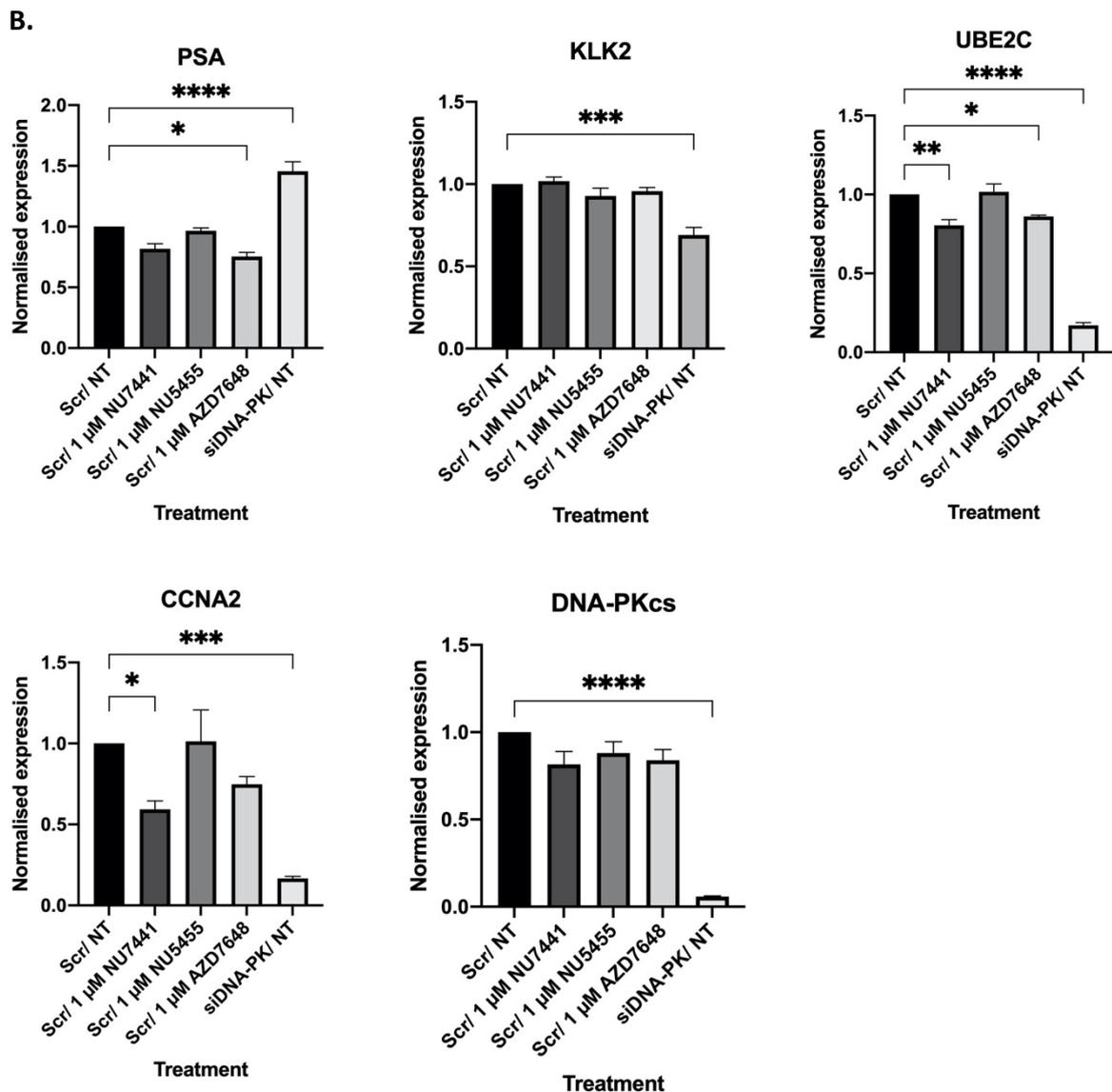
## 5.3 Results

### 5.3.1 Sample validation

Considering results with the inhibitors NU7441 and NU5455 varied at different times in the project, potentially by batch variation or by poor freeze-thaw stability of the compounds, an AstraZeneca commercialised compound, AZD7648 was purchased and included in the RNA-sequencing experiment. This inhibitor showed similar effects to the two NU compounds with a greater than 50% reduction of cell number when treated with 1  $\mu$ M for 5 days in CWR22Rv1-AR-EK cells (Figure 5.5A), and as previously shown, impacts AR/AR-V target gene expression after 24-hours treatment (Figure 4.28).

AR and AR-V target gene expression were also analysed in the samples that were being sequenced to confirm DNA-PKcs knockdown and to determine if the effects that had been seen previously on AR target genes were reproduced in these samples. Although the impact on AR target gene expression was not as pronounced as what had been observed earlier in the project (Figure 5.5), there was an impact on canonical AR and AR-V-driven gene expression. Furthermore, the anti-proliferative effects of DNA-PKcs blockade and depletion were consistent with previous observations in Chapter 4, so the samples were submitted for RNA-sequencing.





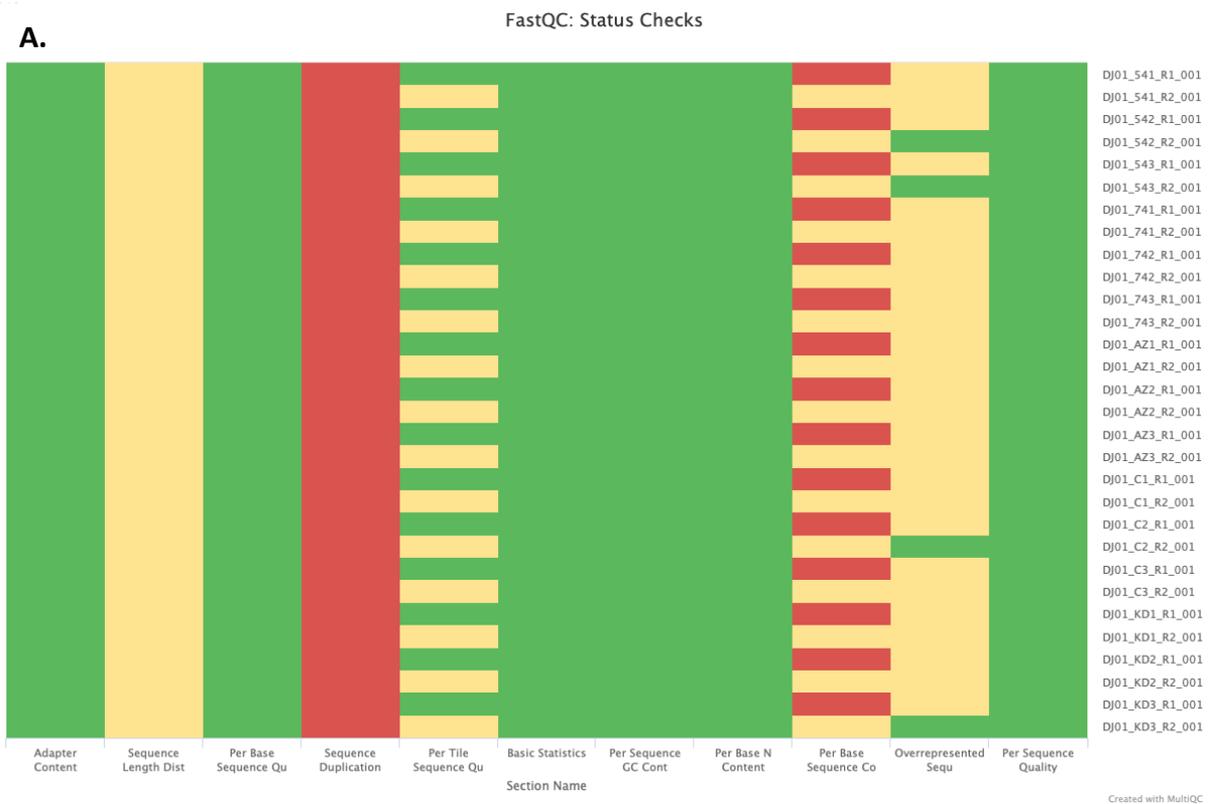
**Figure 5.5** NU7441, NU5455, AZD7648 and knockdown of DNA-PKcs all show similar effects on CWR22Rv1-AR-EK proliferation after 5 days and selectively impact AR/AR-V target gene expression.

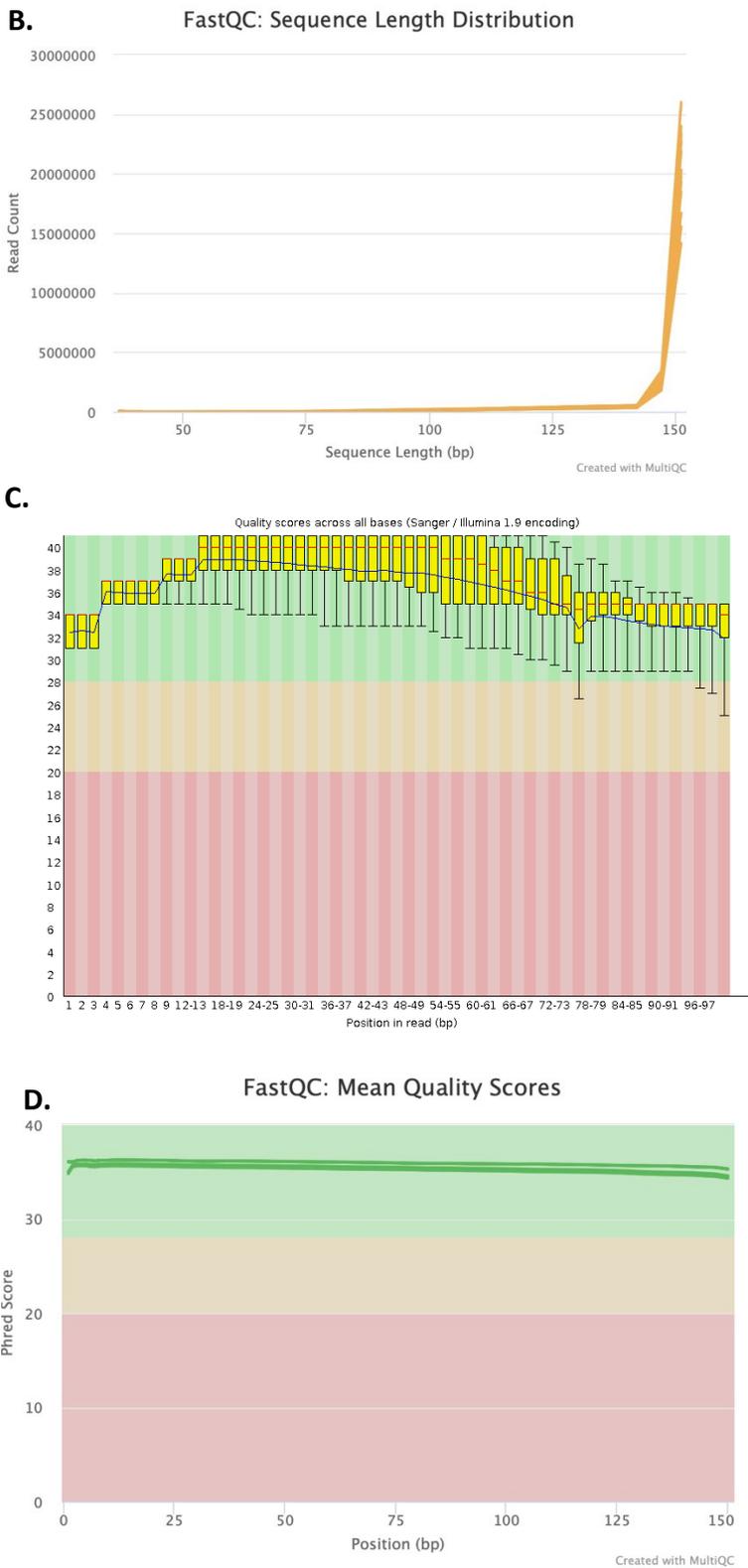
CWR22Rv1-AR-EK cells were reverse transfected with either siScr or siDNA-PK2-4 and incubated for 48 hours prior to 24 hour 1  $\mu$ M DNA-PK inhibitor treatment. Cell counts (A) and qRT-PCR analysis of AR target genes (B) was performed after 5 days and 24 hours, respectively. Graphs show an average of three repeats  $\pm$  SEM (\* =  $p < 0.05$ , \*\* =  $p < 0.01$ , \*\*\* =  $p < 0.001$ , \*\*\*\* =  $p < 0.0001$ ).

### 5.3.2 Quality control of RNA-sequencing samples

The quality of the RNA-sequencing data was analysed using the MultiQC and FastQC tools. As shown in the summary report in Figure 5.6, the samples passed the necessary requirements for downstream analysis. The parameter that appeared as 'failed' in the report was sequence duplication, which is expected to fail for RNA-sequencing due to the presence of highly expressed genes. For all samples, no adapter contamination over 0.1% was detected meaning

no trimming had to be done to the reads. Per sequence quality for all samples passed the quality checks, again meaning no trimming had to be carried out. Figure 5.6B shows the average lengths of the reads in the sample, confirming that 150 bp sequencing was performed. Figure 5.6C shows the per base quality score across the read for one of the control samples and Figure 5.6D shows the average per base quality scored across every sample and three replicates.





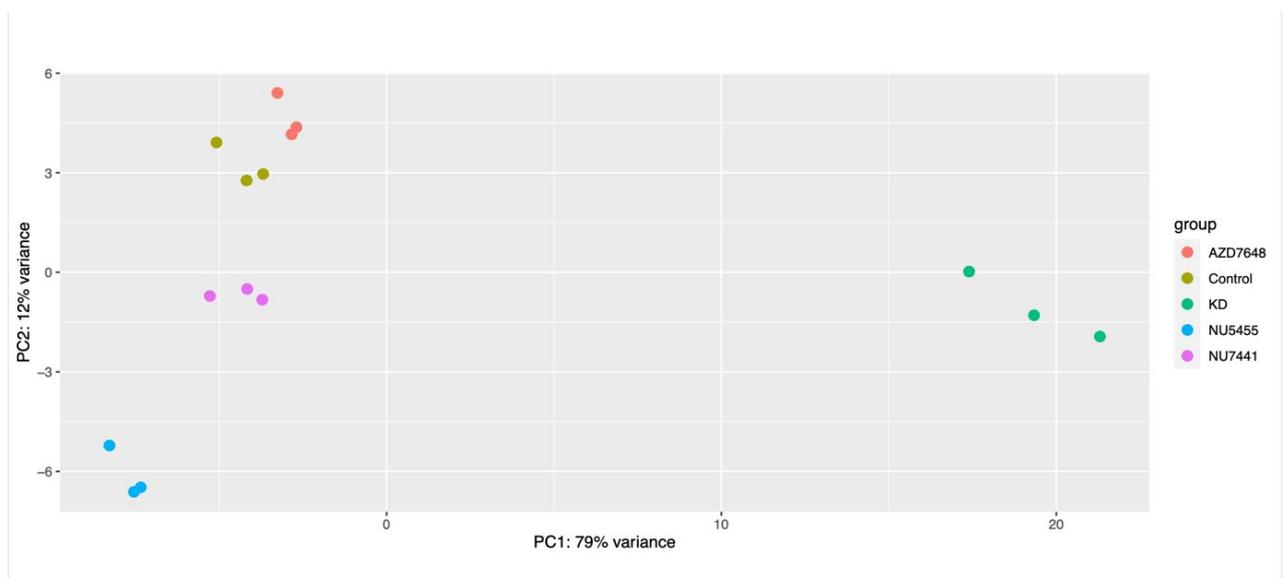
**Figure 5.6** FastQC report generated with MultiQC showed the RNA-sequencing data passed the necessary quality controls.

A. MultiQC summary report for the parameters checked by FastQC. Each sample and replicate is shown (for example DJ01\_541\_\_R1 is the forward strand for replicate 1 for the NU5455 samples) Individual

reports are shown for (B) sequence length distribution, (C) per base quality scored and (D) mean quality scored for all samples.

### 5.3.3 Principal component analysis

Using R Studio, PCA was performed to determine how much the inter-experimental repeats cluster together which will provide confidence in the data that will subsequently be derived from differential gene expression analysis. The PCA plot shows clustering of the individual triplicate samples of the 5 experimental arms, control (DMSO), NU7441, NU5455, AZD7648 and KD (siDNA-PKcs) indicating consistency between inter-experimental repeats (Figure 5.7). Furthermore, the data shows separation of the individual DNA-PKcs inhibitors and knockdown experimental arms suggesting distinct gene expression signatures were detected across the different DNA-PKcs manipulations.



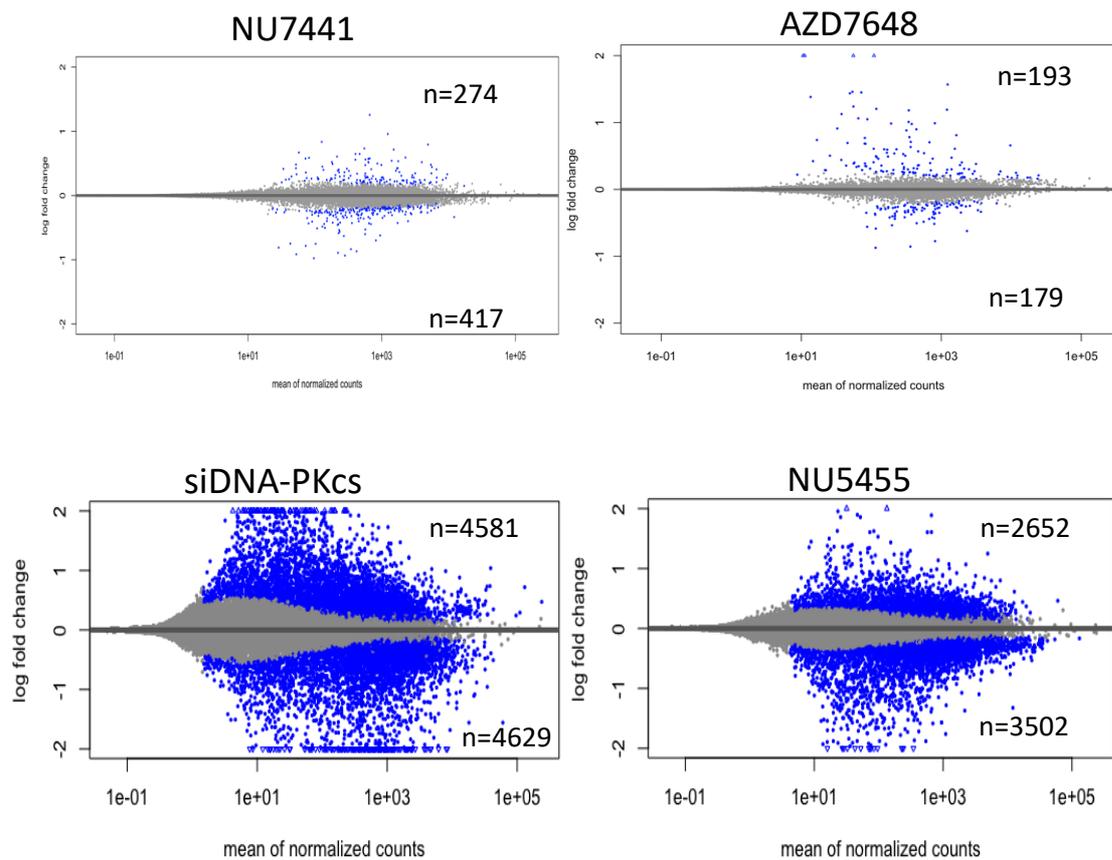
**Figure 5.7** Principal component analysis of RNA-sequencing data shows clustering of intra-experimental repeats of each experimental arm.

RNA-sequencing data was input into R studio after genome alignment and quantification. The gene count matrix was input to a principal component analysis to reveal clustering of each experimental arm. A key indicating the different experimental arms is provided to the right of the plot.

### 5.3.4 Differential gene expression using DESeq2

Differential gene expression analysis for all experimental arms was performed using the DESeq2 package in R Studio. Gene lists were filtered to capture only genes that were up- or down-regulated by more than 1.5-FC and were statistically significant (adjusted p value < 0.05). These genes are subsequently described as significant differentially expressed genes

(SDEG). Table 5.1 shows the numbers of differentially- and significantly DEGs. Importantly, DNA-PKcs knockdown was confirmed in the RNA-sequencing data with an average fold change of -9.78 ( $p$  adjusted =  $1.7 \times 10^{-177}$ ). DNA-PKcs inhibition with NU7441, NU5455 and AZD7648 caused 27, 1195 and 44 SDEGs, respectively, and DNA-PKcs siRNA knockdown caused 3827 SDEG. Critically, this is the first time that the distinct transcriptomic effects of different DNA-PKcs inhibitors has been compared head-to-head and the differences in SDEGs between experimental arms was not expected. I speculate that this difference in SDEG numbers between the treatment arms may reflect different activities of the compounds in the cells, such as differences in intracellular concentration of the compounds over time that could impact the duration of DNA-PKcs inactivation. Alternatively, differences in DNA-PKcs structure when bound to the distinct compounds could cause differences in the DNA-PKcs interactome or DNA-binding capacity that could alter biological effects downstream. Because NU5455 and DNA-PKcs depletion caused the greatest level of change in the transcriptome, these arms were plotted as an MA plot, alongside NU7441 and AZD7648 treatments, to visualise the direction and extent of change in all the DEGs (Figure 5.8). This showed that DNA-PKcs depletion not only caused a greater number of genes to change, but the fold up- and down-regulation was greater as evidenced by the larger spread of points towards the outer edges of the plots.



**Figure 5.8 MA plots show the number of up and down regulated genes in response to DNA-PKcs knockdown and inhibition with NU7441, AZD7648 and NU5455.**

N=number of genes either up or down regulated LFC  $\pm 0$ . Blue dots represent any genes that have a p value < 0.05, with no fold change cut off applied.

**Table 5.1 Lists of differentially expressed and significant differentially expressed genes in response to DNA-PKcs inhibition and knockdown in CWR22Rv1-AR-EK cells**

Treatment	No. of DE genes	No. of sig DE genes (padj < 0.05, FC > 1.5 / < -1.5)
AZD7648	16632	44
NU7441	14050	27
NU5455	18569	1195
siDNA-PKcs	21792	3827

The SDEG lists from the three DNA-PKcs inhibitors were compared to determine if there were overlapping genes. The largest overlap was seen between NU7441 and NU5455; with 89% of SDEGs in response to NU441 also observed in the NU5455-treated SDEG list. Given the low number of DEGs in the AZD7648 and NU7441 treated experimental arms, there was just 2 genes that overlapped between all 3 compounds, *TWIST1* and *PDK4* (Figure 5.9). Interestingly, when determining the direction of change, *PDK4* was upregulated across the three data sets and *TWIST1* was downregulated across the three data sets, providing confidence that these are genuine DNA-PKcs-regulated genes, and that the response was consistent, independently of which inhibitor was used.

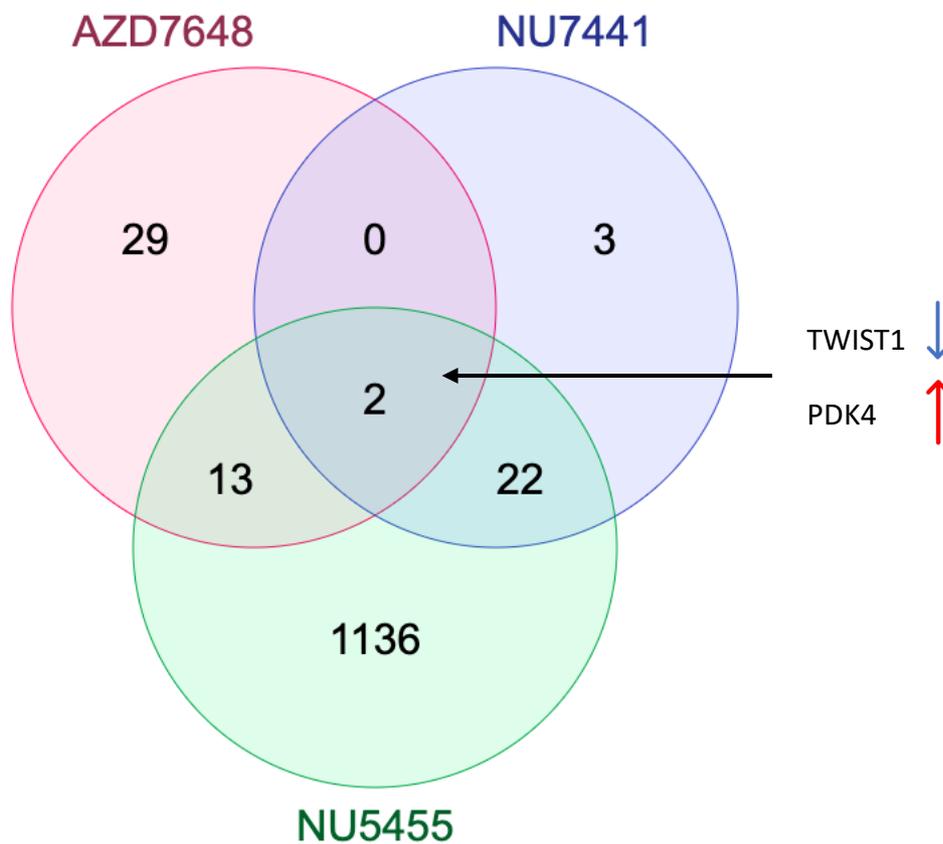
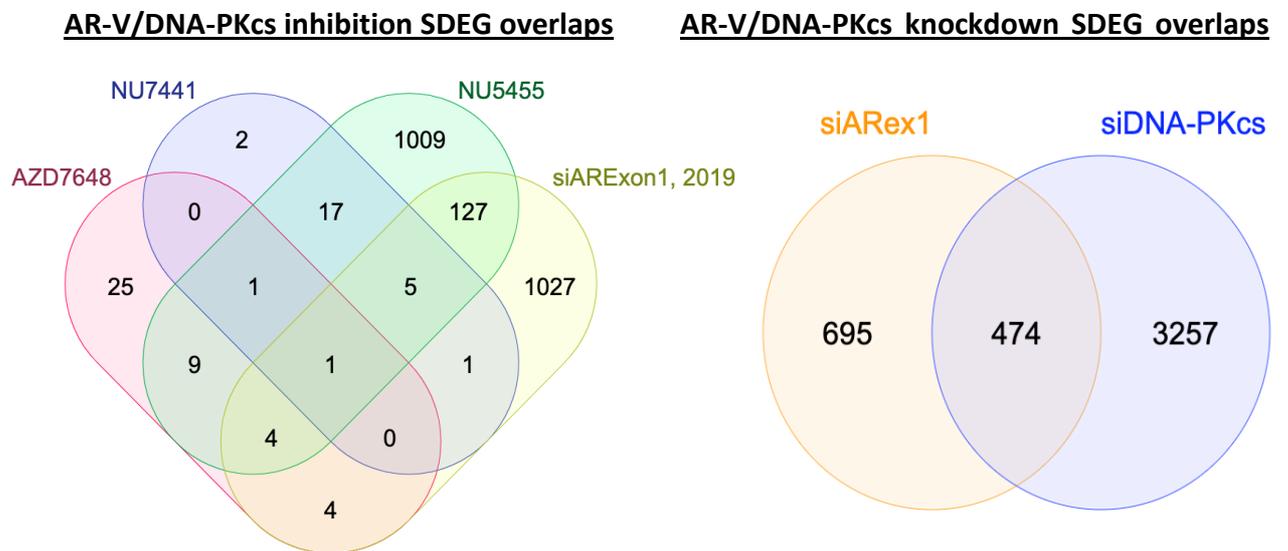


Figure 5.9 Comparison of overlapping significantly differentially expressed genes in response to DNA-PKcs inhibition shows *TWIST1* and *PDK4* are commonly down- and upregulated, respectively.

Given our interest in examining global regulation of AR-V transcriptional activity by DNA-PKcs, gene lists were subsequently compared with a previously published in-house AR-V gene-list, generated by the host laboratory, in which CWR22Rv1-AR-EK cells were subject to 48-hour siRNA-mediated knockdown of AR-Vs prior to RNA-sequencing (Kounatidou *et al.*, 2019). Importantly, this RNA sequencing data was re-analysed using the same DESeq2 bioinformatic pipeline and filtering that had been applied to analyse the DNA-PKcs datasets.

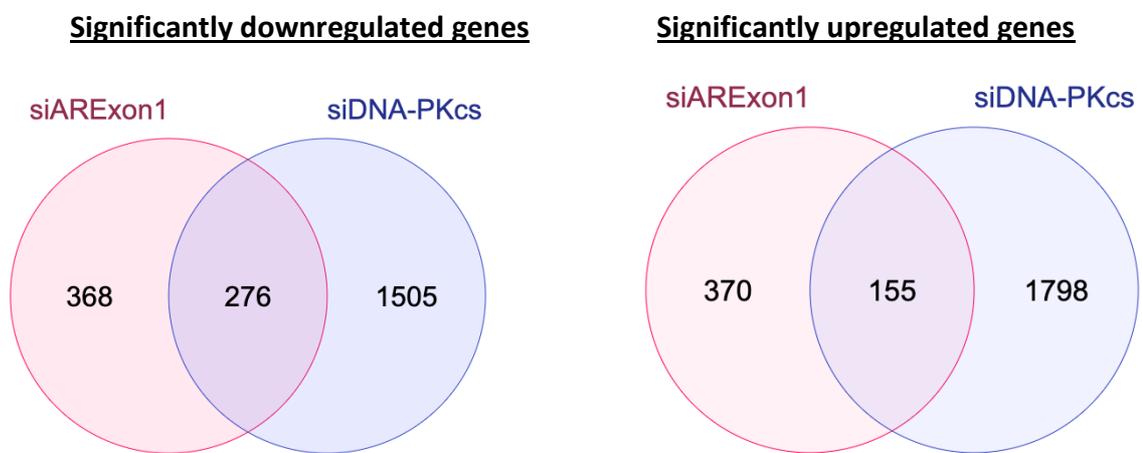
Firstly, the effect of DNA-PKcs inhibition using NU7441, NU5455 and AZD7648 was compared to the in-house AR-V regulated gene-set (referred to as siAR-V gene-set). Of the 1169 genes that are significantly differentially expressed by AR-V depletion, 12% (142/1169) were significantly differentially expressed upon DNA-PKcs inhibition and 41% (474/1169) were significantly differentially expressed upon DNA-PKcs depletion (Figure 5.10). This level of overlap suggests AR-Vs require DNA-PKcs presence more than its catalytic activity for transcriptional regulation. Due to the higher percentage overlap of the AR-V transcriptome with DNA-PKcs depletion, this suggests DNA-PKcs may play an important scaffolding role in its regulation of AR-V activity.



**Figure 5.10** Significantly differentially expressed genes comparisons between the AR-V transcriptome and genes differentially expressed in response to DNA-PKcs inhibition and depletion reveal a significant overlap between AR-V and DNA-PKcs knockdown transcriptomes.

Differentially expressed genes were determined using DESeq2 and the significantly differentially expressed genes ( $p < 0.05$ ,  $\log_2$  fold change  $> \pm 1.5$ ) were compared using Molbiotools.com.

To examine the direction of change in the gene lists, siAR-V downregulated genes and siDNA-PKcs downregulated genes were compared and conversely, siAR-V upregulated and siDNA-PKcs upregulated genes were compared to determine the percentage overlaps. This revealed that 43% (276/644) of genes significantly downregulated by AR-V depletion were also significantly downregulated by DNA-PKcs depletion. Furthermore, 30% (155/525) of genes significantly upregulated by AR-V depletion were also significantly upregulated by DNA-PKcs depletion (Figure 5.11). This data further supports the hypothesis that DNA-PKcs presence is required for AR-V transcriptional activity.

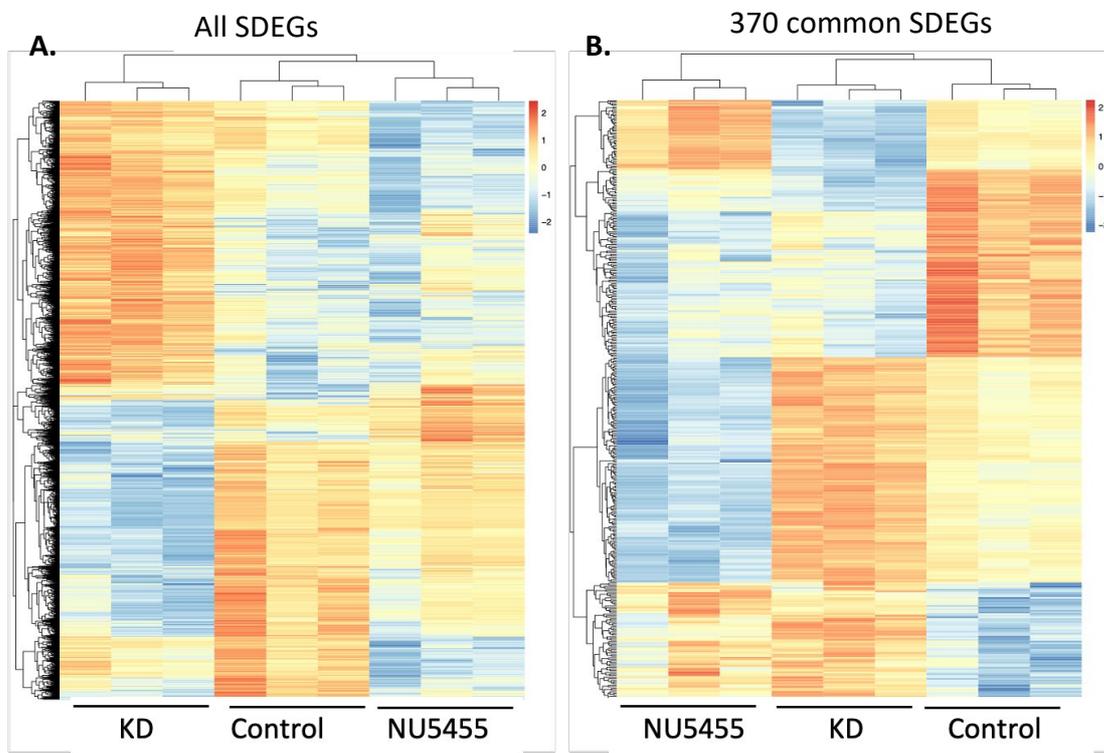


**Figure 5.11** Comparisons between common significantly downregulated and upregulated genes between siAR-V and siDNA-PKcs show a considerable number of SDEGs shared between the two transcriptomes.

Differentially expressed genes were determined using DESeq2 and the significantly differentially expressed genes ( $p < 0.05$ , fold change  $> \pm 1.5$ ) were separated into up- and downregulated lists and compared using Molbiotools.com.

For the three DNA-PKcs inhibition arms, the greatest number of SDEGs were as result of the NU5455 compound. Therefore, this gene list was taken forward to compare with the DNA-

PKcs knockdown arm. There were 370 commonly SDEGs between the two datasets. To further interrogate the difference in roles between DNA-PKcs kinase and non-kinase activity, the clustering of genes were determined compared to the control arm (Figure 5.12A). There are distinct clusters of genes that show opposite directions of gene expression changes in response to DNA-PKcs manipulation. For example, there are clusters of genes that are downregulated in response to DNA-PKcs knockdown but upregulated in response to DNA-PKcs inhibition. Conversely, there are clusters of genes that are upregulated in response to DNA-PKcs knockdown but downregulated in response to DNA-PKcs inhibition. Similarly, when looking at just the 370 common SDEG lists, the clusters of genes that have opposite directions of gene expression changes can again be seen in the heat map (Figure 5.12B).



**Figure 5.12 Heat map showing significant differentially expressed genes in response to DNA-PKcs inhibition with NU5455 and DNA-PKcs depletion**

Lists of all significant differentially expressed genes in response to DNA-PKcs inhibition (NU5455) and DNA-PKcs depletion (KD) were determined using DESeq2 were plotted as a heat map to determine the direction of change of gene expression compared to the control samples.

### 5.3.5 Gene set enrichment analysis reveals pathways that are impacted in response to DNA-PKcs manipulation

To identify what pathways were altered in response to DNA-PKcs manipulation, GSEA was performed using KEGG and Hallmarks gene lists. The complete gene lists without any pre-filtering for all four-treatment arms from DESeq2 were input into the GSEA tool by the Broad Institute (Subramanian *et al.*, 2005, Mootha *et al.*, 2003).

**Table 5.2 Number of significantly differentially expressed KEGG pathways in response to DNA-PKcs inhibition and depletion of DNA-PKcs**

Treatment	Number of downregulated pathways (nominal p value <0.05)	Number of upregulated pathways (nominal p value <0.05)
siDNA-PKcs	34	8
NU5455	48	1
NU7441	30	3
AZD7648	9	7

Focussing on the downregulated pathways upon DNA-PKcs manipulation, the top 10, or in the case of AZD7648, the only significant pathways (nominal p value < 0.05 and false discovery rate (FDR) < 25%) that were downregulated pathways have been tabulated (Table 5.3) with their normalised enrichment score (NES). Comparison between the DNA-PKcs inhibition pathways revealed three were common between all three compounds and are DNA replication, p53 signalling pathway and cell cycle.

**Table 5.3 Downregulated KEGG pathways with their normalised enrichment scores in response to three DNA-PKcs inhibitors and depletion of DNA-PKcs**

Downregulated gene set (NU5455)	NES	Nominal p value	FDR q value
Ribosome	-2.43	0.000	0.000
Neuroactive Ligand Receptor Interaction	-2.24	0.000	0.000
Drug Metabolism Cytochrome P450	-2.18	0.000	0.000
Metabolism of Xenobiotics by Cytochrome P450	-2.17	0.000	0.000

Drug Metabolism other Enzymes	-1.98	0.000	0.001
Cytokine Cytokine Receptor Interaction	-1.93	0.000	0.004
Starch and Sucrose Metabolism	-1.93	0.000	0.003
Cardiac Muscle Contraction	-1.91	0.000	0.004
Dilated Cardiomyopathy	-1.90	0.000	0.003
Parkinsons Disease	-1.89	0.000	0.004

Downregulated gene set (NU7441)	NES	Nominal p value	FDR q value
DNA Replication	-2.51	0.000	0.000
Parkinsons Disease	-2.44	0.000	0.000
Spliceosome	-2.30	0.000	0.000
Cell Cycle	-2.24	0.000	0.000
Oxidative Phosphorylation	-2.21	0.000	0.000
Base Excision Repair	-2.05	0.002	0.002
Huntingtons Disease	-2.05	0.000	0.002
Mismatch Repair	-2.00	0.000	0.003
Cardiac Muscle Contraction	-1.94	0.000	0.006
Ribosome	-1.92	0.000	0.007

Downregulated gene set (AZD7648)	NES	Nominal p value	FDR q value
Cell Cycle	-1.95	0.000	0.016
Mismatch Repair	-1.93	0.002	0.011
P53 Signaling Pathway	-1.86	0.000	0.024
Homologous Recombination	-1.85	0.007	0.020
DNA Replication	-1.81	0.000	0.027

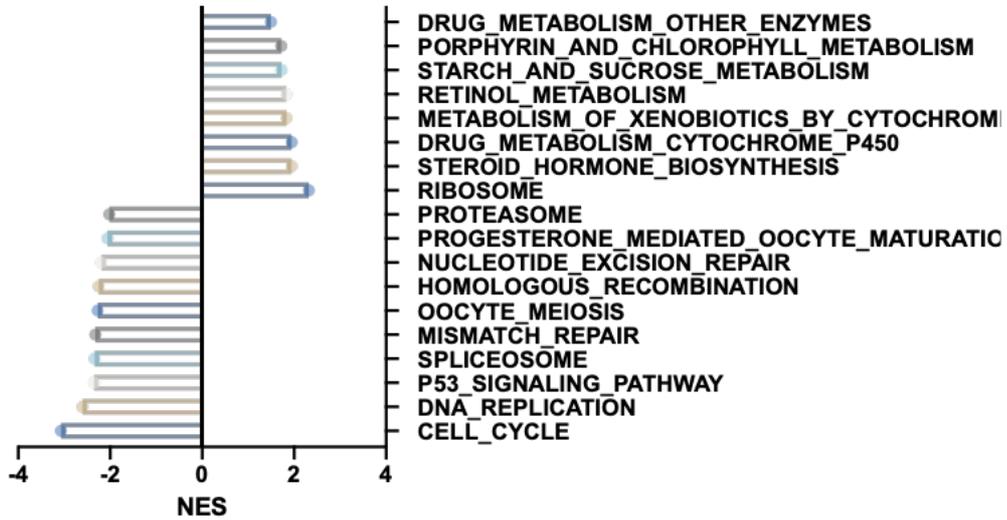
Oocyte Meiosis	-1.74	0.000	0.051
Nucleotide Excision Repair	-1.62	0.011	0.134
Progesterone Mediated Oocyte Maturation	-1.61	0.004	0.12

Downregulated gene set (siDNA-PKcs)	NES	Nominal p value	FDR q value
Cell Cycle	-3.07	0.000	0.000
DNA Replication	-2.60	0.000	0.000
P53 Signalling Pathway	-2.35	0.000	0.000
Spliceosome	-2.34	0.000	0.000
Mismatch Repair	-2.32	0.000	0.000
Oocyte Meiosis	-2.27	0.000	0.000
Homologous Recombination	-2.25	0.000	0.000
Nucleotide Excision Repair	-2.19	0.000	0.000
Progesterone Mediated Oocyte Maturation	-2.06	0.000	0.001
Proteasome	-2.02	0.000	0.001

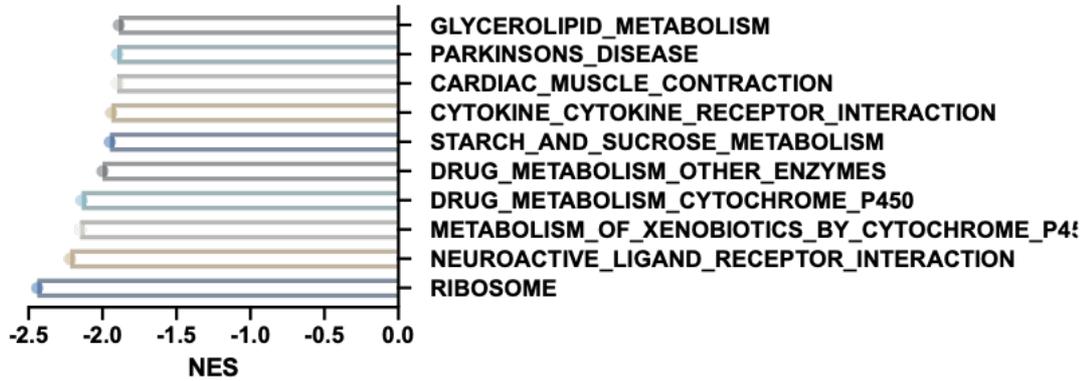
**Table 5.4 Commonly downregulated KEGG pathways between DNA-PKcs inhibition with NU5455 and depletion of DNA-PKcs**

DNA Replication	P53 Signalling Pathway	Spliceosome	Homologous Recombination	Proteasome
Pyrimidine Metabolism	Parkinson's Disease	Base Excision Repair	N Glycan Biosynthesis	Oxidative Phosphorylation
Huntington's Disease	Purine Metabolism	Cell Cycle		

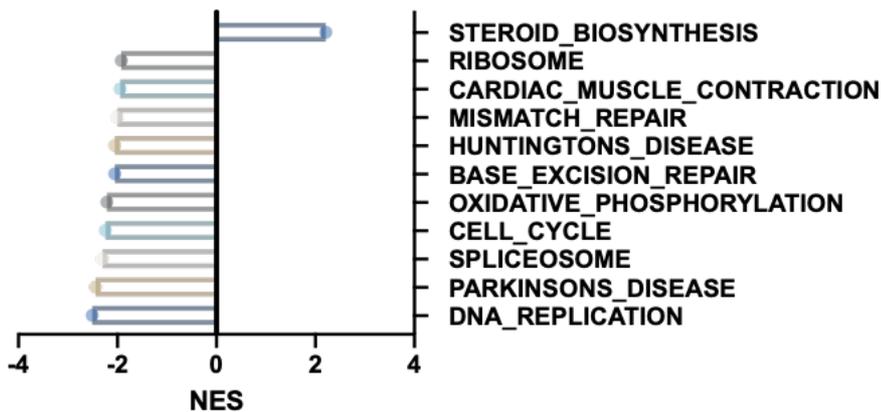
**siDNA-PKcs Enriched Pathways**



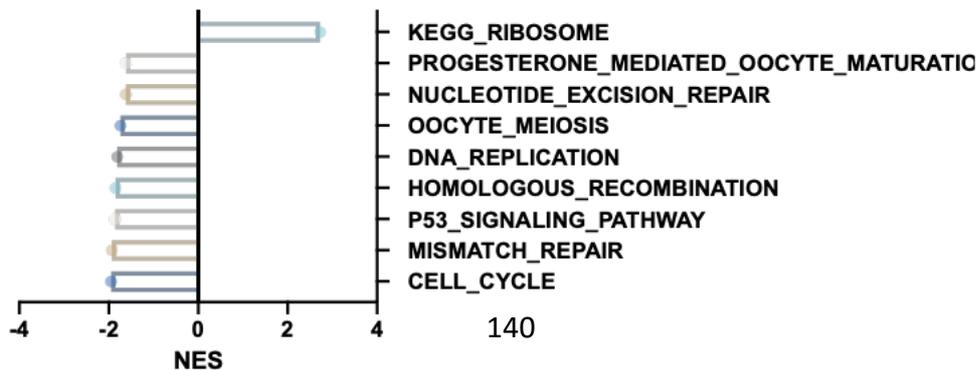
**NU5455 Enriched Pathways**



**NU7441 Enriched Pathways**



**AZD7641 Enriched Pathways**



**Figure 5.13 DNA-PKcs regulates several KEGG gene-sets.**

DEG lists from NU7441, NU5455, AZD7648 and siDNA-PKcs treatment in CWR22Rv1-AR-EK cells were compared to KEGG gene lists using GSEA. Graphs show the top 10 negatively enriched pathways (8 for AZD7648) with a p value < 0.05 and FDR < 25% and the only positively enriched pathways with a p value < 0.05 and FDR < 25%. NES = normalised enrichment score.

The hallmarks gene sets from GSEA were also analysed for enrichment post DNA-PKcs inhibition using NU5455, as this was the DNA-PKcs inhibitor dataset that caused the most genes to be differentially expressed, and DNA-PKcs knockdown. There were 26 hallmark pathways downregulated by DNA-PKcs depletion and 31 by DNA-PKcs inhibition. There were 17 commonly downregulated pathways in both the knockdown and inhibition arms, including MYC targets V1 and V2 (Table 5.7). DNA-PKcs has previously been shown to phosphorylate MYC and modulate the stability of MYC (An *et al.*, 2008, Iijima *et al.*, 1992). These pathways, as well as E2F targets and others, were also shown to be significantly down regulated in previous RNA sequencing data in the prostate cancer cell line C4-2 after DNA-PKcs inhibition with NU7441 (Dylgjeri *et al.*, 2019)(Figure 5.14).

**Table 5.5 Number of significantly differentially expressed Hallmark pathways in response to DNA-PKcs inhibition and depletion of DNA-PKcs**

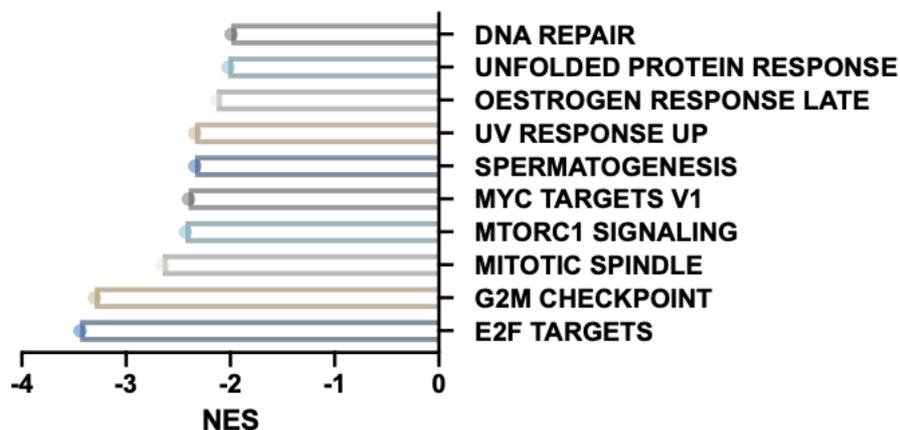
Treatment	Number of downregulated pathways (nominal p value <0.05)	Number of upregulated pathways (nominal p value <0.05)
siDNA-PKcs	26	0
NU5455	31	1

**Table 5.6 Common downregulated Hallmark pathways with their normalised enrichment scores in response DNA-PKcs inhibition with NU5455 and depletion of DNA-PKcs**

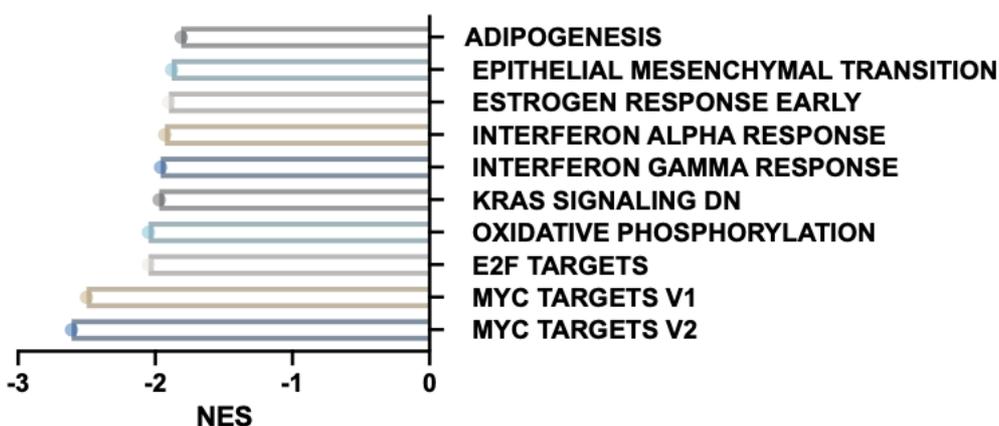
<u>Hallmarks</u>	NES	Nominal p value	FDR q value
<b>Downregulated gene set (siDNA-PKcs)</b>			
E2F Targets	-3.44	0.000	0.000
G2M Checkpoint	-3.30	0.000	0.000
Mitotic Spindle	-2.65	0.000	0.000
MTORC1 Signalling	-2.43	0.000	0.000
MYC Targets V1	-2.40	0.000	0.000
Spermatogenesis	-2.34	0.000	0.000
UV Response Up	-2.34	0.000	0.000
Oestrogen Response Late	-2.13	0.000	0.000
Unfolded Protein Response	-2.02	0.000	0.000
DNA Repair	-1.99	0.000	0.000

<u>Hallmarks</u>	NES	Nominal p value	FDR q value
<b>Downregulated gene set (NU5455)</b>			
MYC Targets V2	-2.61	0.000	0.000
MYC Targets V1	-2.50	0.000	0.000
E2F Targets	-2.05	0.000	0.000
Oxidative Phosphorylation	-2.05	0.000	0.000
KRAS Signalling Dn	-1.97	0.000	0.000
Interferon Gamma Response	-1.96	0.000	0.000
Interferon Alpha Response	-1.93	0.000	0.000
Oestrogen Response Early	-1.90	0.000	0.000
Epithelial Mesenchymal Transition	-1.88	0.000	0.001
Adipogenesis	-1.81	0.000	0.002

**siDNA-PKcs Enriched Hallmark Pathways**



**NU5455 Enriched Hallmark Pathways**



**Figure 5.14 DNA-PKcs regulates dysregulation to several hallmark gene sets.**

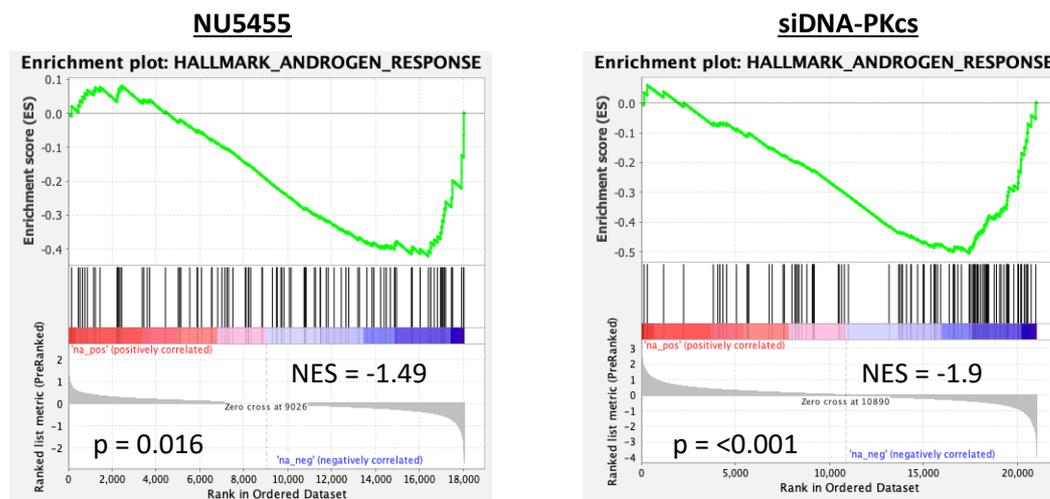
DEG lists from both NU5455 and siDNA-PKcs treatment in CWR22Rv1-AR-EK cells were compared to Hallmark gene lists using GSEA. Graphs show the top 10 negatively enriched pathways with a p value < 0.05 and FDR < 25%. NES = normalised enrichment score.

**Table 5.7 Commonly downregulated hallmark pathways between DNA-PKcs inhibition and depletion.**

G2M Checkpoint	MTORC1 Signalling	MYC Targets V1	UV Response UP	Oestrogen Response Late
Unfolded Protein Response	DNA Repair	Glycolysis	MYC Targets V2	Androgen Response
Hypoxia	Oxidative Phosphorylation	Oestrogen Response Early	P53 Pathway	E2F Targets

## 5.3.6 DNA-PKcs inhibition and knockdown downregulates the androgen response

DNA-PKcs knockdown and inhibition using NU5455 both caused a significant downregulation in the Hallmark androgen response gene-set. DNA-PKcs inhibition caused a NES of -1.49 (p value 0.016) and DNA-PKcs depletion caused a NES of -1.9 (p value <0.001) for the androgen response pathway, further supporting the hypothesis that DNA-PKcs is a transcriptional regulator of the AR/AR-Vs. Given the level of overlapping SDEGs between the AR transcriptome and DNA-PKcs depletion datasets, DNA-PKcs depletion caused a greater, more significant reduction on the androgen response. The Hallmark pathway analysis includes an androgen response gene list, whereas the KEGG pathway analysis do not, hence this pathway was not detected in the KEGG GSEA described previously.



**Figure 5.15** DNA-PKcs depletion and inhibition causes a significant downregulation of the androgen response in CWR22Rv1-AR-EK cells.

Unfiltered differentially expressed gene lists from both NU5455 and siDNA-PKcs treatment in CWR22Rv1-AR-EK cells were analysed for differentially enriched Hallmark gene lists using GSEA.

Interestingly, known regulators of the androgen response such as *AURKA*, encoding Aurora Kinase A, and *EZH2*, Enhancer of zeste homolog 2, were significantly down-regulated upon DNA-PKcs knockdown with a fold change of -4.5,  $p = 7 \times 10^{-70}$  (*AURKA*) and -2.2,  $p = 7 \times 10^{-41}$  (*EZH2*). Aurora Kinase A regulates the cell cycle and is associated with AR expression in PCa. Its expression was shown to be induced by DHT stimulation in LNCaP cells as a consequence of AR binding to an intronic enhancer region of the *AURKA* gene (Kivinummi *et al.*, 2017).

Furthermore, Aurora kinase A has been shown to regulate AR-Vs by modulating splicing; depletion of Aurora kinase A reduced AR-V protein levels and therefore AR-V target genes such as *UBE2C* (Jones *et al.*, 2017). EZH2 has also been identified as an AR co-regulator by directly binding to the AR, and is involved in PCa progression (Liu *et al.*, 2019, Varambally *et al.*, 2002). Interestingly, EZH2 depletion was shown to enhance a double stranded RNA-STING pathway leading to increased antigen presentation and interferon response and therefore increasing the anti-tumour response of checkpoint inhibition in PCa (Morel *et al.*, 2021). This could be important when determining other mechanisms by which DNA-PKcs manipulation exerts anti-cancer effects and may provide a rationale for drug combination approaches. However, it is not known what role EZH2 plays in AR-V regulation.

#### 5.3.7 DNA-PKcs inhibition and knockdown downregulates genes involved in the spliceosome

The emergence of AR-Vs during hormonal therapy (ADT and/or anti-androgens) is due, principally, to alternative splicing of full-length AR pre-mRNA transcripts. Interestingly, the spliceosome KEGG pathway was significantly downregulated for all but the AZD7648 treatment arms: DNA-PKcs depletion caused a NES of -2.3 (p value < 0.0001); inhibition with NU5455 caused a NES of -1.8 (p value < 0.0001); and NU7441 caused a NES of -2.3 (p value < 0.0001). All DEGs with their corresponding LFC values in the spliceosome gene list are presented in Appendix 8.1, and the DEGs that are 'core enriched' and contribute to the leading-edge subset of genes that are shared between the NU5455 and siDNA-PKcs treatment arms are tabulated (Table 5.8). Given the downregulatory effect of compromising DNA-PKcs activity on AR-V7 transcript levels which was particularly evident in VCaP cells (Figure 4.10), it was speculated that differentially-expressed genes within this spliceosome gene set may offer insight into splicing factors that regulate the synthesis of AR-Vs. In addition, it was necessary to determine what the impact of this change in spliceosome-associated genes would be on global splicing. To establish this, SUPPA2 was applied to the siDNA-PKcs RNA sequencing data to perform differential splicing analysis. SUPPA2 provides a value that represents the difference in percentage spliced-in (PSI) abundances between two conditions ( $\Delta$ PSI). The siDNA-PKcs data was used for the splicing analysis as this was the treatment that caused the lowest NES and has the largest number of splicing genes that are significantly differentially expressed (Figure 5.16).



**Figure 5.16 Differential splicing analysis reveals genes that are significantly alternatively spliced in response to DNA-PKcs knockdown**

CWR22Rv1-AR-EK cells were depleted of DNA-PKcs before RNA-sequencing. Data was analysed for differential splicing using SUPPA2. A. Events that passed a  $\Delta\text{PSI} \pm 0.2$  were plotted. B. Events that passed a p value cut off of  $< 0.05$  were plotted. C. Volcano plot of differential splicing events with events that passed the cut offs of false discovery rate  $< 0.05$  and  $\Delta\text{PSI}$  of 0.6 were annotated with their gene ID.

The differential splicing analysis revealed there were 358 significant ( $p < 0.05$ ) alternative splicing events in response to DNA-PKcs depletion, although the androgen receptor was not present in this list. However, in the previous chapter, VCaP cells showed the most noticeable change in AR-V transcript levels upon DNA-PKcs manipulation whereas in CWR22Rv1-AR-EK and CWR22Rv1 the changes were more subtle (Figure 4.18 and Figure 4.19), so this may not be unexpected.

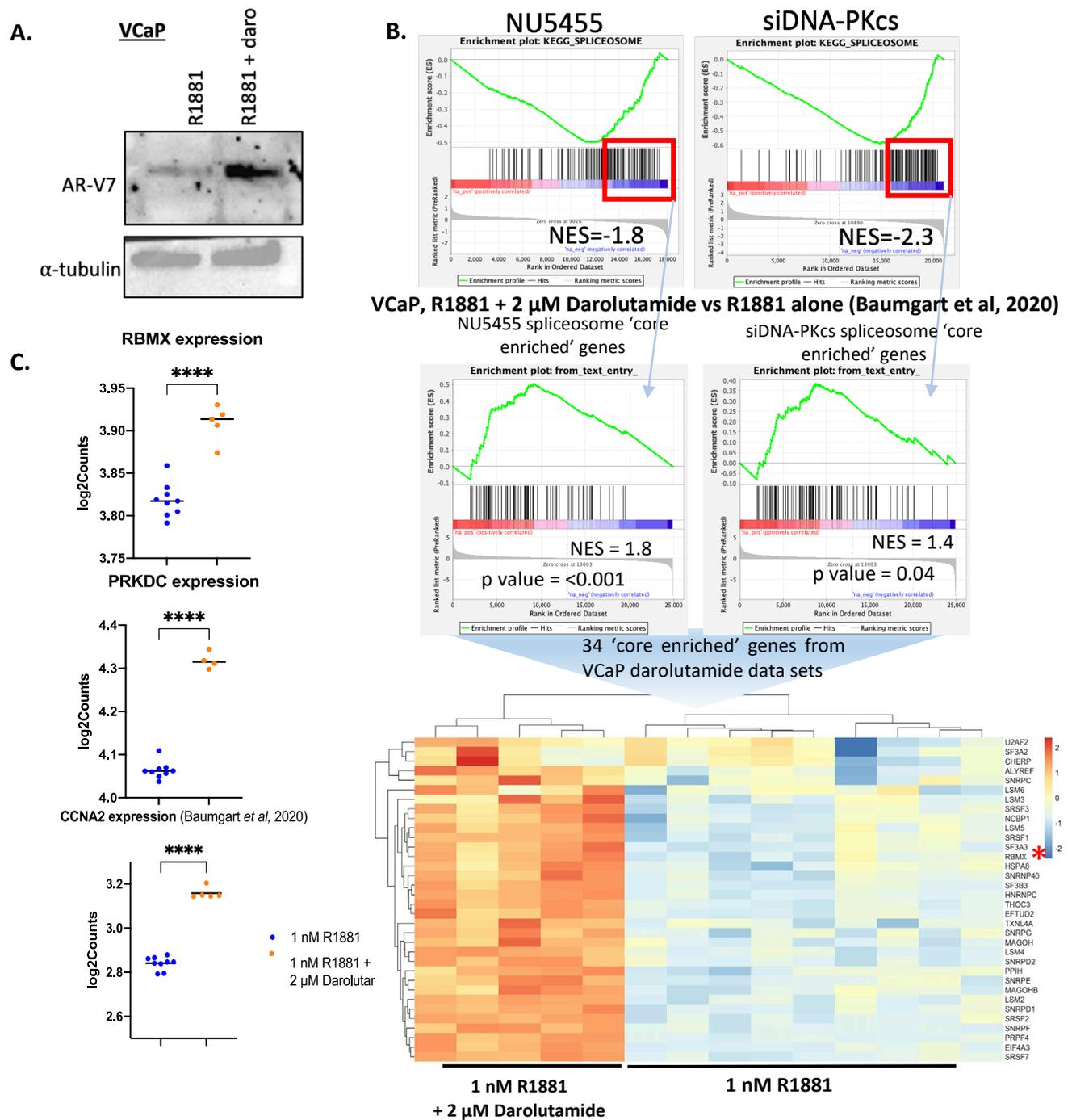
Focussing on NU5455 and siDNA-PKcs, as these treatment arms had the most DEGs that were included in the leading edge 'core enriched' spliceosome-associated genes, the gene lists were compared with a published RNA-sequencing data set to provide a list of spliceosome genes that may be involved in AR-V splicing. The published RNA sequencing dataset is from VCaP cells that had been treated with either synthetic androgens (R1881) or R1881 plus the anti-androgen darolutamide (Baumgart *et al.*, 2020). Treatment with anti-androgens in the VCaP cell line has been shown to upregulate the production of AR-Vs in response to FL-AR inhibition. This was confirmed in house, where 24-hour darolutamide treatment caused an increase in AR-V7 protein levels (Figure 5.17A). Therefore, genes from the spliceosome gene list that are upregulated in response to darolutamide, and show concurrent downregulation in response to DNA-PKcs inhibition/knockdown, may be responsible for or contribute to alternative splicing required to produce AR-Vs. This comparative exercise could support validation studies of splicing factors deemed important for AR-V synthesis.

GSEA of the most downregulated spliceosome-associated genes by DNA-PKcs inhibition and depletion showed a positive and statistically significant enrichment in the VCaP-Darolutamide data set. The downregulated spliceosome genes in response to DNA-PKcs inhibition showed a NES of 1.8,  $p < 0.001$  and DNA-PKcs depletion showed a NES of 1.4,  $p = 0.04$  when aligned to the splicing-associated genes up-regulated in response to darolutamide treatment in VCaP

cells. The overlapping upregulated genes from both lists were then plotted in a heatmap (Figure 5.17B) with the splicing-associated genes shown to the right. To reiterate, the genes up-regulated in the heat-map (shown in red) are those that show: (i) elevated expression in response to darolutamide (AR-V synthesis is activated) and (ii) down-regulated expression when DNA-PKcs activity is compromised (AR-V levels are decreased).

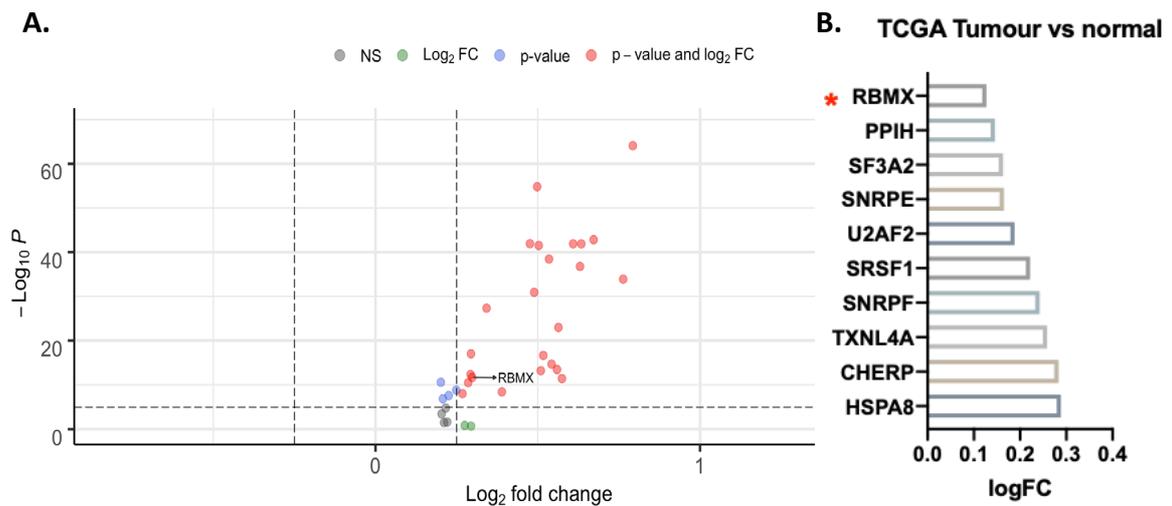
To visualise the level of upregulation in response to darolutamide in VCaP cells, the genes were plotted in a volcano plot (Figure 5.18A). Then, to narrow down the list of splicing factors that could be investigated further, using the tumour vs normal dataset from TCGA that was previously analysed (Chapter 4), it was determined if each of the 34 splicing genes were upregulated in tumour samples. The genes with a 10% increase in expression in tumour samples ( $\log_2$  fold change over 0.13) are plotted in Figure 5.18B. From this list, RBMX was taken forward for validation as a regulator of AR-V splicing.

Reassuringly, when DNA-PKcs counts were extracted from the RNA-sequencing data, this showed DNA-PKcs is also significantly upregulated in response to darolutamide, correlating with expression of the splicing factor, RBMX, as well as the AR-V regulated gene *CCNA2*. This further backed-up the hypothesis that DNA-PKcs regulation of RBMX expression leads to AR-V generation (Figure 5.17C).



**Figure 5.17 DNA-PKcs regulated genes involved in the spliceosome, which are upregulated in response to darolutamide in VCaP cells**

Lists of differentially expressed genes in response to DNA-PKcs manipulation were input into GSEA. The most downregulated genes were compared to a published RNA-sequencing dataset. Overlapping genes were plotted in a heat map. Log<sub>2</sub> counts and p values from DESeq2 of *RBMX*, *PRKDC* and *CCNA2* were extracted and plotted using Prism

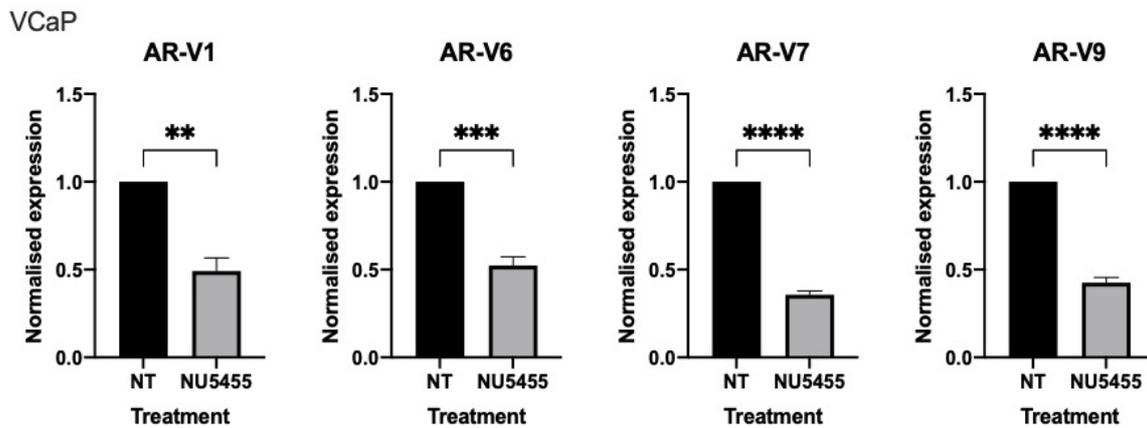


**Figure 5.18 RBMX was significantly upregulated in response to darolutamide in VCaP cells and in tumour vs normal samples.**

A. The 34 spliceosome genes that are downregulated in response to DNA-PKcs inhibition and depletion and upregulated in response to darolutamide are presented in a volcano plot and highlighting RBMX. B. The genes that are upregulated by 10% in matched tumour vs normal samples from the list of 34 genes are presented, highlighting RBMX with an Asterix.

### 5.3.8 Splicing regulation involving DNA-PKcs and RBMX controls androgen receptor variant synthesis

To validate that DNA-PKcs impacted AR-V mRNA levels, the VCaP cell line, that selectively expresses AR-Vs under certain conditions, was used. This allows interrogation of AR-V generation as they are largely upregulated in response to being cultured in steroid-depleted conditions, rather than CWR22Rv1-AR-EK cells that constitutively express AR-Vs irrespective of growth conditions. VCaP cells were incubated in steroid-depleted conditions for 48 hours prior to treatment with 1  $\mu$ M of the DNA-PKcs inhibitor NU5455 for 24 hours. DNA-PKcs inhibition using NU5455 significantly downregulated the expression of several AR-Vs, AR-V1, AR-V6, AR-V7 and AR-V9 (Figure 5.19). Downregulation of AR-V mRNA following DNA-PKcs knockdown was shown in Chapter 4 and is validated further in Section 5.3.9.

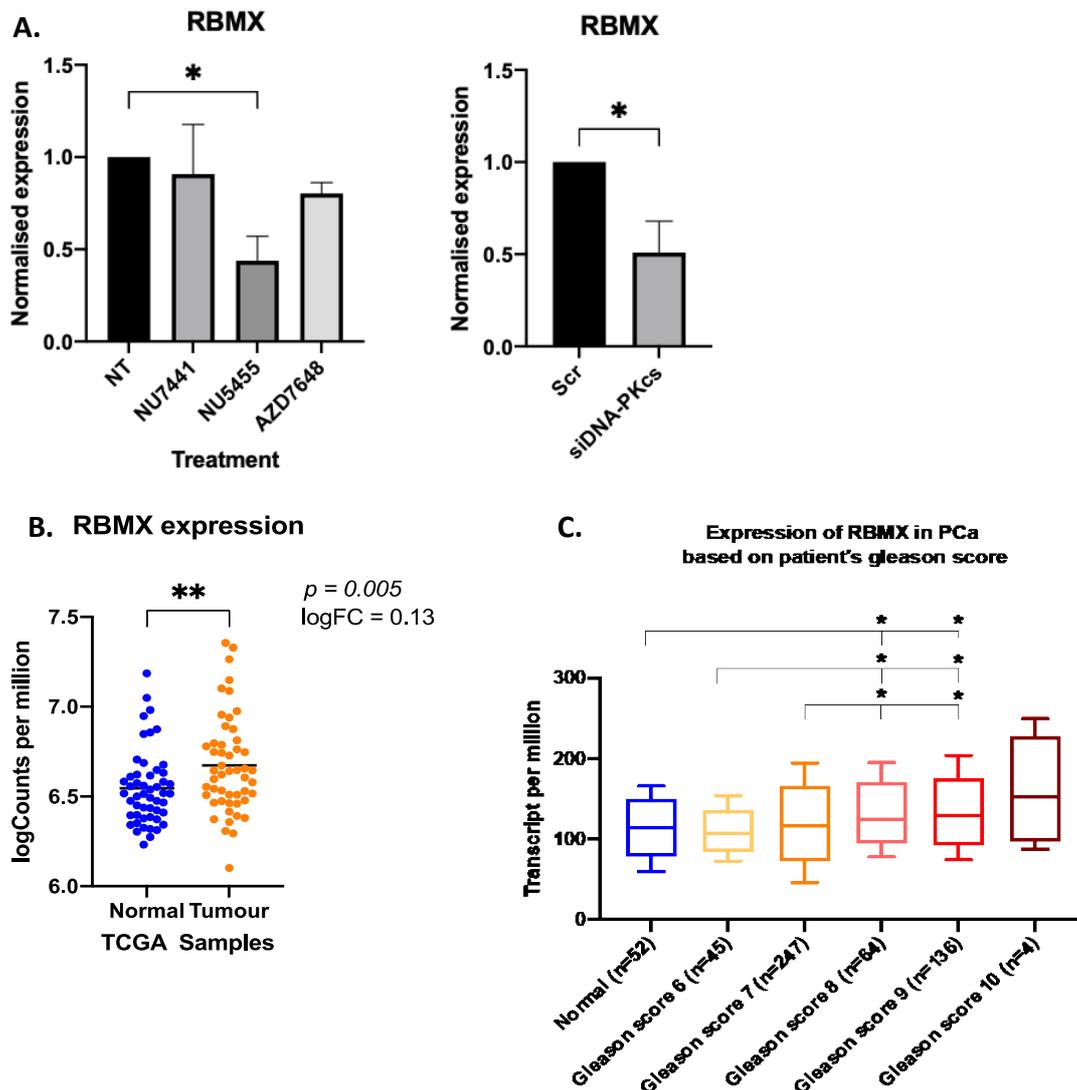


**Figure 5.19 DNA-PKcs inhibition and depletion decreases expression of several AR-V transcripts.**

VCaP cells seeded and grown in steroid-depleted conditions for 48 hours then 24 hours treatment with 1  $\mu$ M NU5455 before being subject to RT-qPCR analysis of AR-V1, AR-V6, AR-V7 and AR-V9 levels. Data represents the mean of three repeats  $\pm$  SEM. An unpaired t test was used to determine the statistical significance (\*\* =  $p < 0.01$ , \*\*\* =  $p < 0.001$ , \*\*\*\* =  $p < 0.0001$ ).

To confirm expression of RBMX is impacted by DNA-PKcs inhibition and depletion, and correlated with AR-V synthesis, CWR22Rv1-AR-EK cells were subject to DNA-PKcs inhibitor treatment and DNA-PKcs depletion for 24 hours and 72 hours, respectively. Consistent with the *in silico* predictions made from the RNA-sequencing data, RBMX transcript levels were significantly reduced in response to selective DNA-PKcs inhibition with NU5455 and DNA-PKcs knockdown (Figure 5.20A). NU7441 and AZD7648 did not significantly impact RBMX levels, which was also consistent with *in silico* predictions made from the RNA-sequencing data.

Interestingly, RBMX expression changed in the TCGA tumour vs normal dataset; RBMX expression is significantly upregulated ( $p = 0.005$ ) in tumour samples with a 10% increase in expression compared to normal prostate tissue (Figure 5.20B). RBMX expression was also analysed across PCa Gleason grades; a scoring system applied to indicate how advanced the tumour is with highest score of 10 representing the most advanced stage of disease. This showed RBMX expression significantly correlates with increased Gleason score (Figure 5.20C), suggesting that RBMX could be associated with tumour progression in PCa.

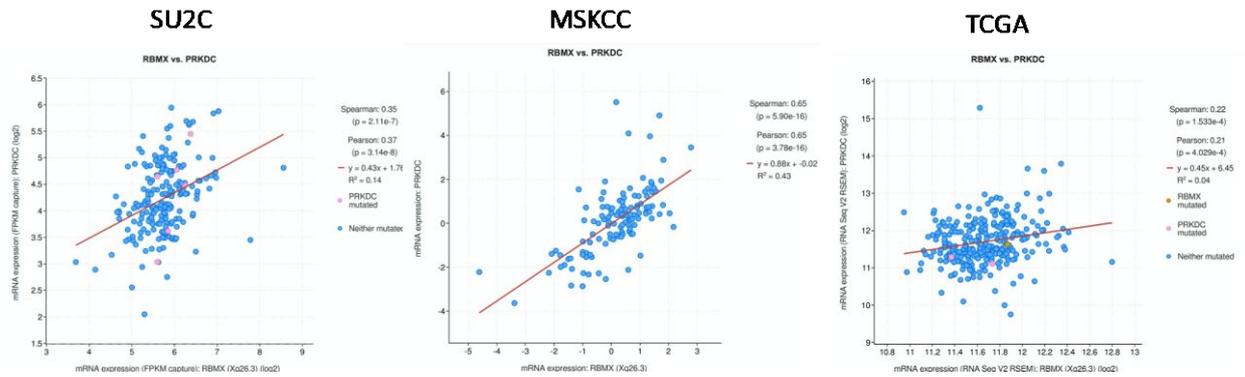


**Figure 5.20** Validation that RBMX identified from RNA-sequencing experiment are downregulated by DNA-PKcs inhibition and knockdown in separate independent experiments in CWR22Rv1-AR-EK cells.

CWR22Rv1-AR-EK cells were either reverse transfected with siScr or siDNA-PK2-4 for 72 hours or seeded, incubated for 48 hours, then treated for 24 hours with 1  $\mu$ M NU7441, NU5455 or AZD7648. A. RBMX expression was analysed using qRT-PCR. B. RBMX counts extracted from TCGA was plotted against matched normal vs tumour samples or C. Gleason score, using transcript per million average values from ULACAN (\* =  $p < 0.05$ , \*\* =  $p < 0.01$ ).

The investigate the relationship between DNA-PKcs and RBMX, the correlation between RBMX and PRKDC mRNA expression was assessed *in silico* using the cBioPortal database by interrogating the Stand-up to cancer/ PCF (SU2C), Memorial Sloan Kettering Cancer Centre (MSKCC) (Taylor *et al.*, 2010) and TCGA datasets. RBMX mRNA expression was positively

correlated with PRKDC mRNA expression with a spearman correlation coefficient of 0.35 in the SU2C dataset, 0.65 in the MSKCC dataset and 0.22 in the TCGA dataset. Taken together with the previous data that shows RBMX expression is downregulated upon DNA-PKcs inhibition and depletion, this could suggest the *RBMX* gene is regulated by DNA-PKcs.



**Figure 5.21 DNA-PKcs mRNA expression is correlated with RBMX mRNA expression.**

cBioPortal was utilised to interrogate if RBMX and the DNA-PKcs encoding gene *PRKDC* mRNA were correlated in three independent datasets. These included SU2C, MSKCC and TCGA (Abida *et al.*, 2017, Cancer Genome Atlas Research, 2015, Taylor *et al.*, 2010) .

### 5.3.9 RBMX is a key regulator of AR-V synthesis

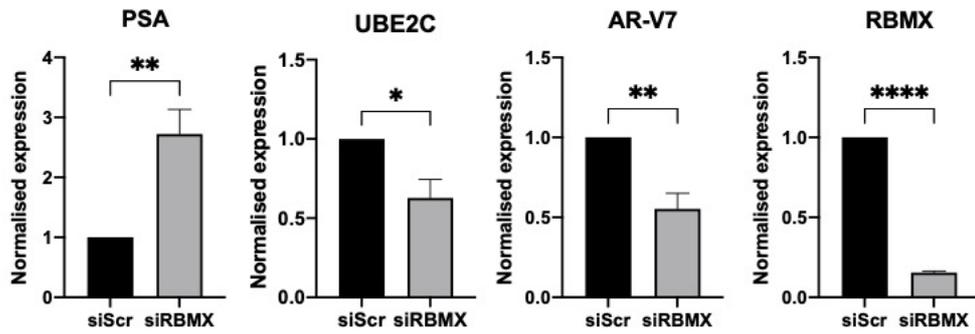
Given that RBMX: (i) was down regulated by both DNA-PKcs inhibition and depletion in the RNA sequencing data set; (ii) was confirmed as significantly down-regulated in validation qRT-PCR experiments with DNA-PKcs inhibition and knockdown; (iii) is significantly increased in TCGA patient data in primary tumours compared to normal; and (iv) it was positively correlated with DNA-PKcs expression and Gleason score, it was taken forward for further interrogation.

RBMX (RNA Binding Motif protein, X-linked) is an RNA binding protein belonging to the group of heterogenous nuclear ribonuclear proteins (hnRNPs) (Geuens *et al.*, 2016). RBMX is thought to play a role in splicing as it has been shown to interact with several splicing regulators, such as Tra2-beta in a yeast 2-hybrid screen and was identified as part of the supraspliceosome (Heinrich *et al.*, 2009).

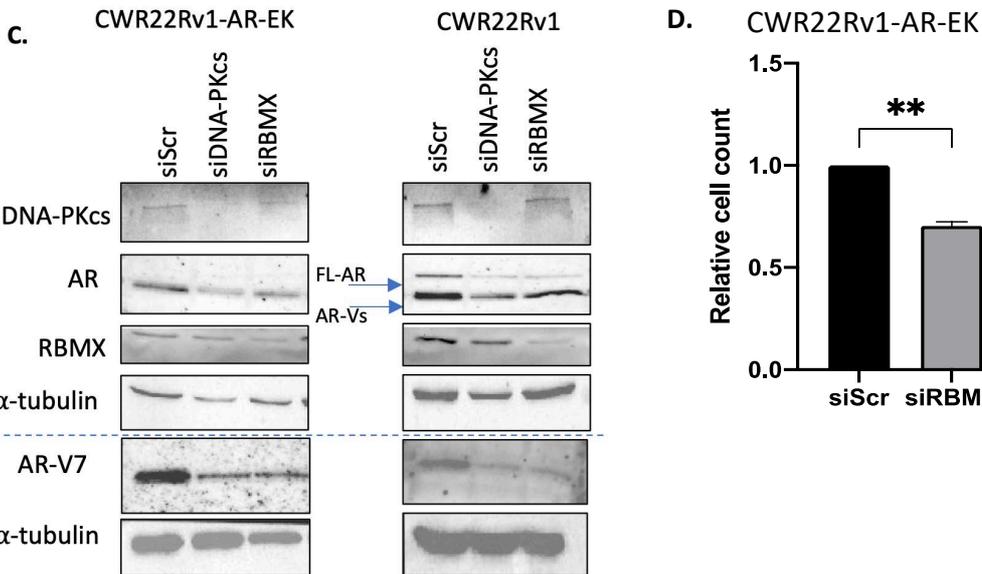
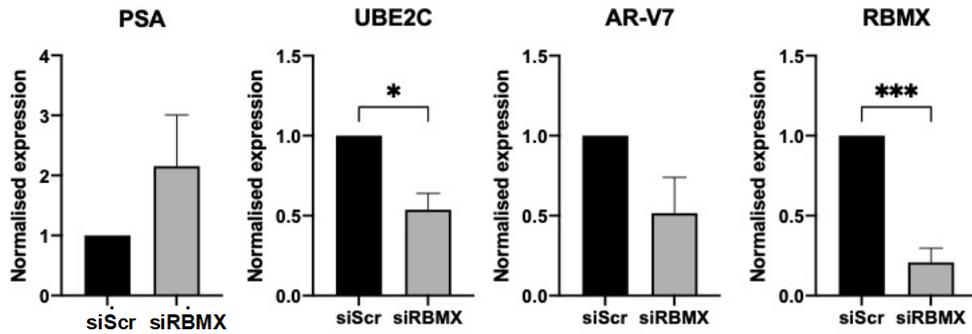
To establish a potential role for RBMX in regulation of AR-V levels, and by proxy AR-V target gene expression, RBMX was depleted for 72 hours in CWR22Rv1-AR-EK and CWR22Rv1 cells, respectively, prior to qRT-PCR (Figure 5.22). RBMX knockdown in both cell lines was confirmed at the mRNA and protein level. In CWR22Rv1-AR-EK cells grown in full media, RBMX depletion significantly decreased AR-V7 mRNA levels and the AR-V target gene *UBE2C*. AR-V7 transcript levels were able to be interrogated as the reverse primer is located in the AR-V7 specific cryptic exon, CE3, meaning other AR/AR-V mRNA species will not be detected. Intriguingly, *PSA* was significantly increased, which is at odds with reduced AR-V protein levels observed by western analysis. The reason for this up-regulation of *PSA* mRNA is currently unknown, but we speculate that possible alterations to splicing and/or post-splicing mRNA metabolism of the *PSA* transcript, as a consequence of RBMX knockdown, may give rise to the observed effect.

Consistent with downregulation at the transcript level, AR-V protein was also decreased in response to RBMX depletion in CWR22Rv1-AR-EK cells (Figure 5.22C). This reduction to AR-V levels translated to a significant 30% reduction in CWR22Rv1-AR-EK cell growth after 5 days post-RBMX depletion (Figure 5.22D). CWR22Rv1 cells grown in steroid-depleted conditions, to induce the production of AR-Vs, and depleted of RBMX showed similar results to CWR22Rv1-AR-EK cells in which there was an increase in *PSA* expression, significant decrease in the AR-V target *UBE2C*, and a reduction in AR-V7 mRNA levels, although this was not statistically significant. The western blot in Figure 5.22C (right) showed reductions in both AR-V and FL-AR protein levels.

**A. CWR22Rv1-AR-EK**



**B. CWR22Rv1**

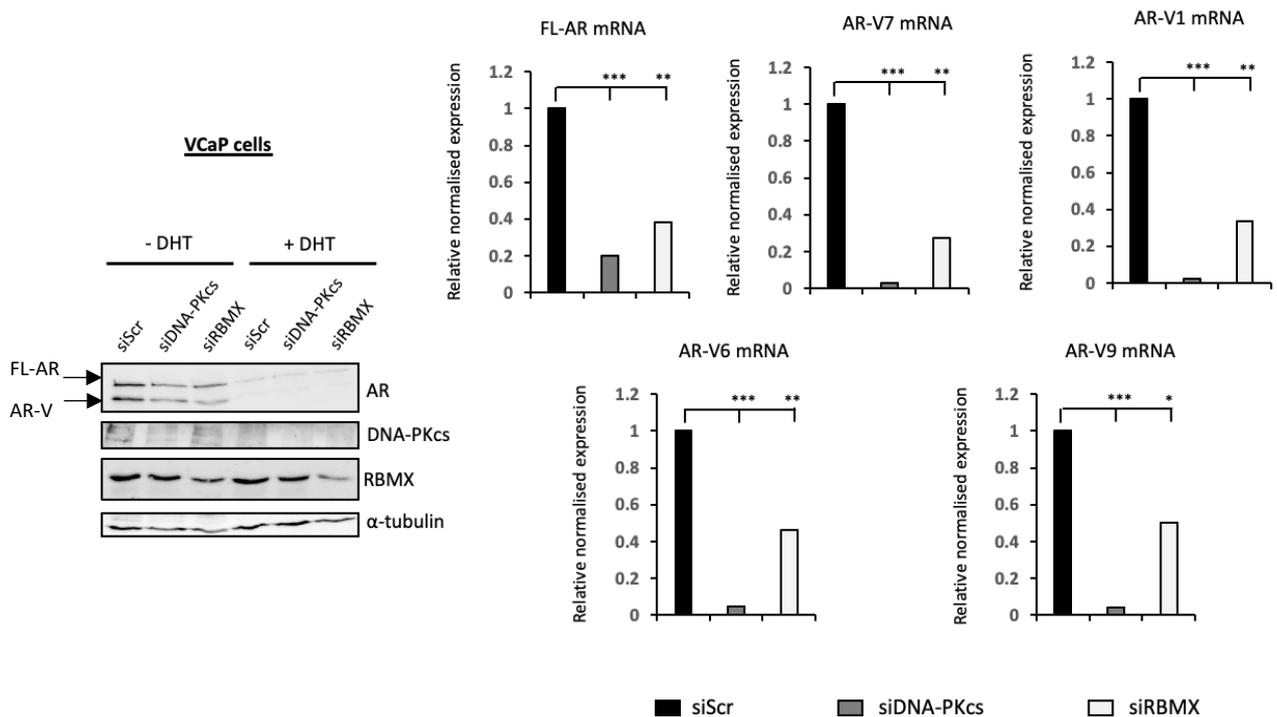


**Figure 5.22 RBMX depletion significantly reduces AR-V and AR-V target gene expression and impacts PCa cells proliferation.**

CWR22Rv1-AR-EK (A) and CWR22Rv1 (B) were grown in full and steroid-depleted media, respectively, after siScr, siRBMX or siDNAPK2-4 reverse transfection, for 72 hours. AR target gene expression was analysed using qRT-PCR. Data represents the mean of three repeats  $\pm$  SEM. An unpaired t-test was used to determine the statistical significance (\* =  $p < 0.05$ , \*\* =  $p < 0.01$ , \*\*\* =  $p < 0.001$ ). C. CWR22Rv1-AR-EK and CWR22Rv1 treated as above and AR, AR-V7, DNA-PKcs and RBMX protein levels were analysed using western blot. D. CWR22Rv1-AR-EK cells were cultured for 5 days in full media after transfection of siScr or siRBMX and cell counts were performed. Data is representative of two

independent repeats  $\pm$  SEM. An unpaired t-test was used to determine the statistical significance (\*\* =  $p < 0.01$ ).

VCaP cells were then depleted of DNA-PKcs or RBMX to ensure the previous results were not cell line specific. VCaP cells were seeded in steroid-depleted conditions and transfected with either siScr, siDNA-PK2-4 or siRBMX. AR and AR-V protein levels and AR isoform mRNA levels were then analysed. This validated that RBMX depletion significantly reduces FL-AR and several AR-V transcript levels which is also reflected at the protein level (Figure 5.23).

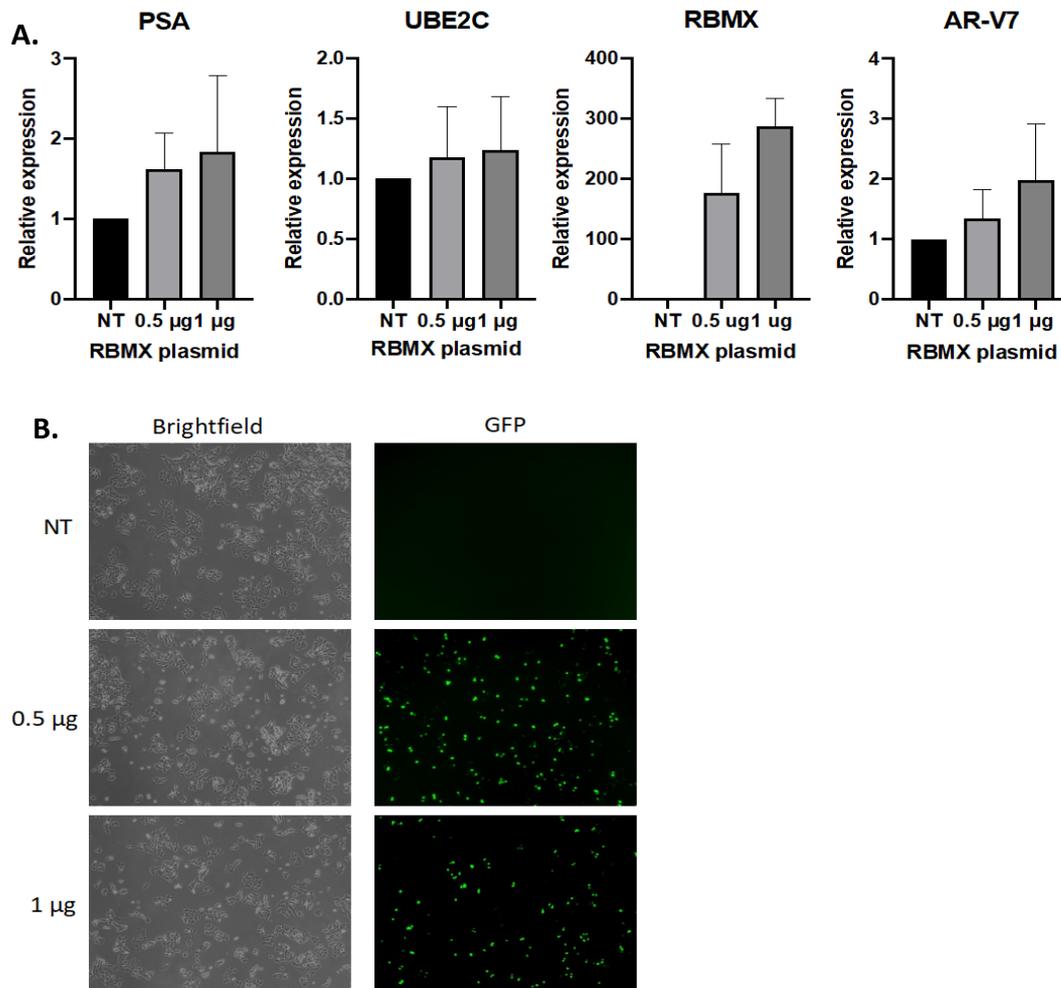


**Figure 5.23 RBMX depletion significantly reduces AR and AR-V protein and transcript levels.**

VCaP cells were grown in steroid-depleted media and transfected with siScr, siDNAPK2-4 or siRBMX for 72 hours. (A) AR, DNA-PKcs and RBMX protein levels were analysed using western blot and (B) AR target gene expression was analysed using qRT-PCR. Data represents the mean of three repeats  $\pm$  SEM. An unpaired t-test was used to determine the statistical significance (\* =  $p < 0.05$ , \*\* =  $p < 0.01$ , \*\*\* =  $p < 0.001$ ).

RBMX overexpression was also investigated to determine if it would lead to an increase in AR-V7 mRNA and protein levels. After 72 hours transfection with a GFP-RBMX-expressing plasmid, both AR/AR-V target genes *PSA* and *UBE2C*, and AR-V7 transcript, were increased but not statistically significantly. RBMX over-expression was confirmed at the mRNA and protein level and GFP expression was confirmed using microscopy (Figure 5.24B). Due to low transfection efficiency of CWR22Rv1-AR-EK cells, the data shown does not reflect a full population of RBMX

over-expressing cells. Although time constraints prevented this work continuing, the experiment could be improved by GFP-based cell sorting to select only those which were successfully transfected and repeating the qRT-PCR experiments in that sub population. This will give a better reflection of the effects of RBMX overexpression in this cell line.



**Figure 5.24. RBMX overexpression marginally increases AR-V mRNA and AR-V target gene expression.**

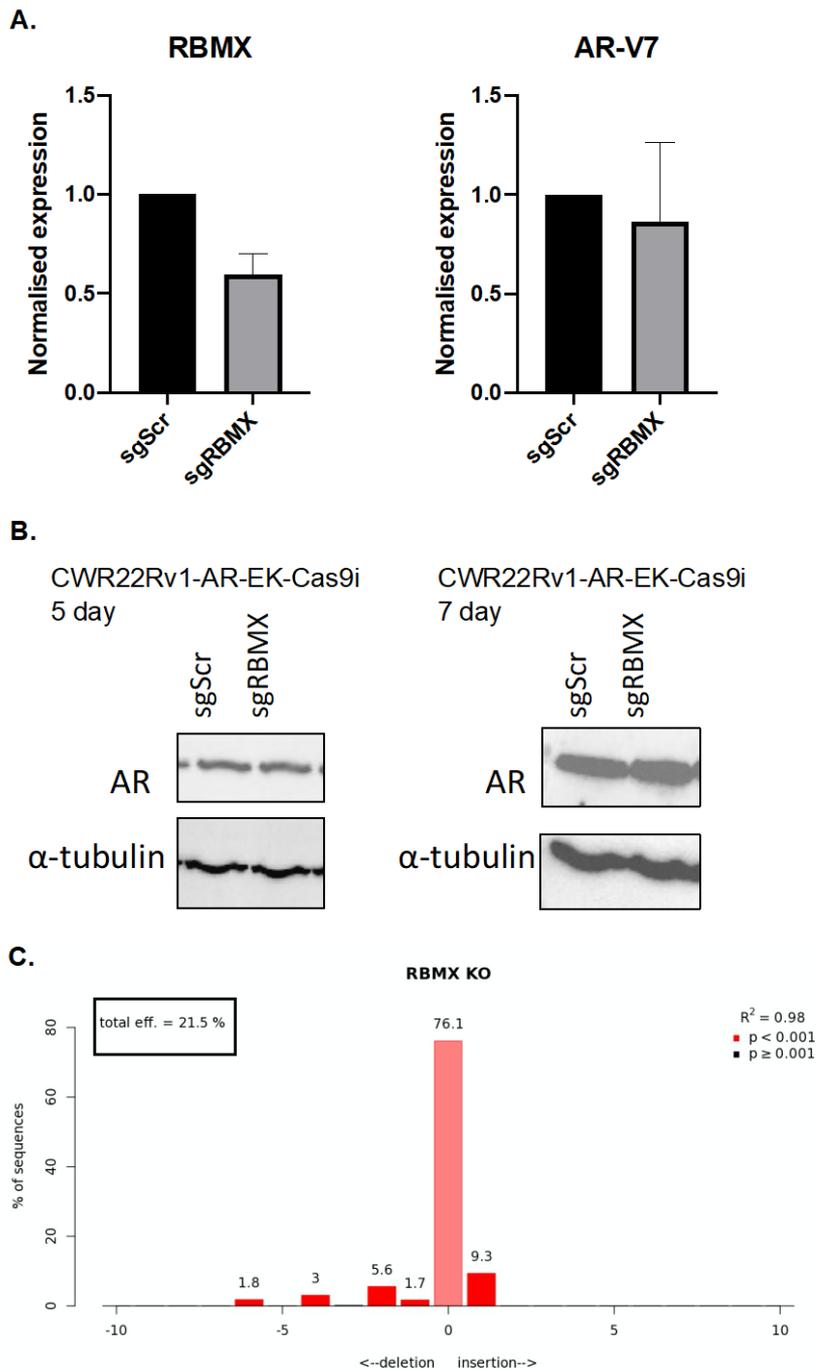
A. CWR22Rv1-AR-EK cells were transfected with a GFP-RBMX expression plasmid for 72 hours. Cells were then subject to qRT-PCR analysis of AR target genes *PSA* and *UBE2C*, as well as AR-V7 and RBMX expression. B. CWR22Rv1-AR-EK cells were transfected as in (A) and imaged using a NikonTE2000 to both demonstrate successful GFP-RBMX expression (using GFP as a surrogate for RBMX) and to capture brightfield and fluorescent images (10x magnification).

To further assess the role of RBMX in controlling AR-V generation, CRISPR-based RBMX knockouts were performed using a CWR22Rv1-AR-EK cell derivative that expresses Cas9 in response to doxycycline (made in-house by PhD student Laura Walker; called CWR22Rv1-AR-EK-iCas9). Utilising this cell-line, which has been used to successfully knock out targets by

combining induction of Cas9 expression with transfection of a suitable synthetic guide RNA (sgRNA), the impact of RBMX knockout on AR-V levels was also interrogated, using Sigma pre-designed MISSION™ sgRNA. RBMX and AR-V7 levels were analysed using qRT-PCR 72 hours after transfection with either an RBMX-targeting (sgRBMX) or scrambled control (sgScr) sgRNA. As shown in Figure 5.25A, a 40% reduction in RBMX mRNA was observed in the RBMX sgRNA experimental arm compared to sgScr control, but this did not impact on AR-V7 mRNA levels. AR-V protein levels were also assessed 5- and 7-days post-transfection but, consistent with the transcript analysis data, AR-V abundance remained at the same level as sgScr control (Figure 5.25B). In light of these findings, the efficiency of the RBMX-targeting sgRNA was determined using TIDE analysis of genomic DNA at the RBMX-targeting sgRNA locus, that was amplified using PCR following genomic DNA extraction 72 hours post sgRNA transfection using the primers in Table 5.9. Critically, CRISPR-induced indel efficiency, using the RBMX-targeting sgRNA, was only 21.5% indicating the chosen RBMX sgRNA is not efficient at enabling CRISPR targeting of the *RBMX* gene (Figure 5.25C). This experiment, therefore, was unable to effectively determine the biological effects of RBMX knockout on AR-V generation. Although time constraints meant these experiments could not be repeated, it would be important to improve these studies by using multiple optimised sgRNAs to generate more successful knockouts prior to examining the effects on AR-V generation and splicing.

**Table 5.9 Primer sequences for amplification of genomic DNA for TIDE analysis**

Gene	Forward (5'-3')	Reverse (5'-3')
<i>RBMX</i>	GCTCCTGTATCACGTGGAAG	ACCTCCACTTGGATGATCTGAAT

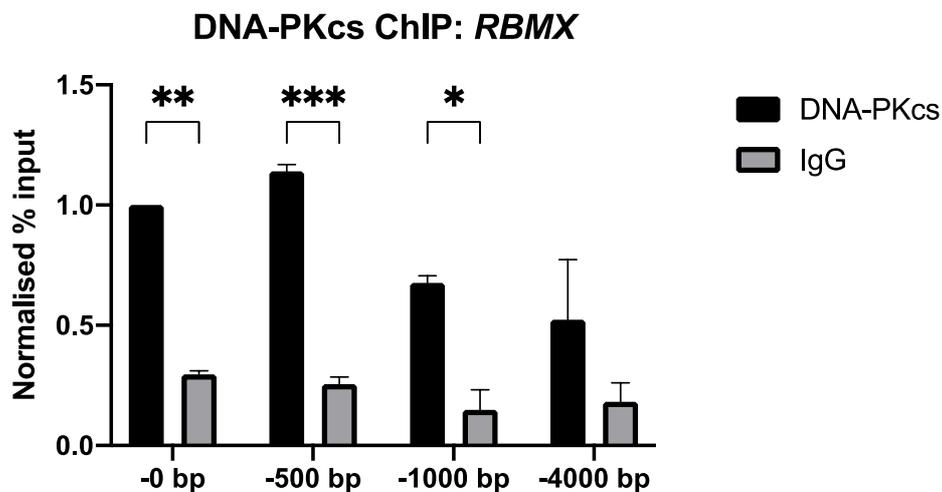


**Figure 5.25 RBMX CRISPR knockout was not achieved in CWR22Rv1-AR-EK-Cas9i cells.**

A. CWR22Rv1-AR-EK cells were treated with 1  $\mu$ M doxycycline, to induce Cas9 expression, for 24 hours before seeding and transfection with sgRNAs sgScr and sgRBMX for 72 hours. Cells were then subject to qRT-PCR analysis of RBMX and AR-V7 mRNA expression. B. CWR22Rv1-AR-EK cells were doxycycline-induced for 24 hours before seeding and transfection with sgScr and sgRBMX for 5 and 7 days. Cells were then lysed and subject to western blot analysis of AR-V protein levels. C. CWR22Rv1-AR-EK cells

treated as in (A) prior to genomic DNA extraction and PCR to amplify the sgRBMX-targeting locus. Amplicons were purified, sequenced and subject to TIDE analysis to determine indel efficiency.

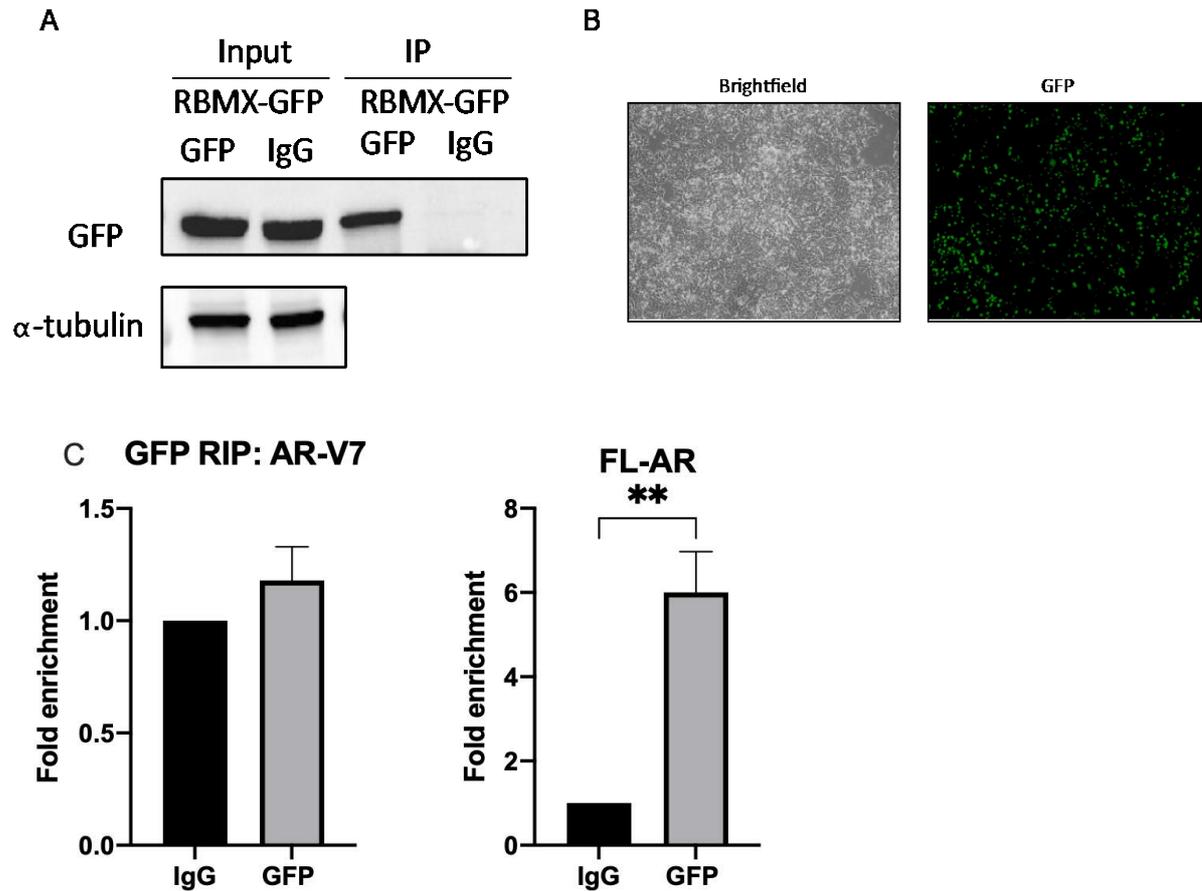
As previously shown, DNA-PKcs inhibition and depletion led to the downregulation of RBMX levels in both RNA-sequencing data and subsequent independent validation experiments suggesting DNA-PKcs is a transcriptional regulator of the *RBMX* gene. To examine this further, we next assessed if DNA-PKcs was directly involved in the regulation of *RBMX* transcription by determining if it is recruited to *cis*-regulatory elements upstream of the *RBMX* gene. To this end, ChIP experiments were performed with a DNA-PKcs antibody and regions -0 bp, -500 bp, -1000 bp and -4000 bp upstream from the *RBMX* transcriptional start site were analysed for DNA-PKcs enrichment. This showed that DNA-PKcs enrichment was around 3-4-fold over the IgG control at proximal promoter sites of the *RBMX* gene which decreased as the primer pairs ventured further upstream (Figure 5.26).



**Figure 5.26 DNA-PKcs is recruited to regions upstream of the *RBMX* transcriptional start site.**

CWR22Rv1-AR-EK cells were seeded in full media and allowed to grow to 70-80% confluency and used for a DNA-PKcs ChIP. ChIP-qPCR readouts show recruitment of DNA-PKcs to the *RBMX* transcriptional start site, -500, -1000 and -4000 bp upstream compared to the IgG control. Data shown represents the normalised percentage input to the DNA-PKcs ChIP at the -0 bp site and represents 2 independent repeats  $\pm$  SEM. Two-way ANOVA using Šídák's multiple comparisons test was used to determine the statistical significance (\* =  $p < 0.05$ , \*\* =  $p < 0.01$ , \*\*\* =  $p < 0.001$ ). Table shows the average fold change of the % inputs over the IgG control and represents 2 independent repeats.

Given that RBMX has pleiotropic cellular roles, including regulation of splicing and transcription, it was important to determine if the effect of manipulating RBMX levels on AR-V transcript levels was a consequence of altered splicing of the AR transcript or transcription of the *AR* gene. Therefore, RNA immunoprecipitation (RIP) experiments were performed to determine if there was an interaction between RBMX and AR-V transcripts. RIP was initially attempted on endogenous levels of RBMX incorporating an RBMX antibody in CWR22Rv1-AR-EK cells. However, the RBMX antibody was not suitably concentrated to perform the immunoprecipitation. To address this, ectopic expression of GFP-tagged RBMX was utilised for RIP using a GFP antibody. CWR22Rv1-AR-EK cells were seeded in full media, transfected with GFP-RBMX expressing constructs on the day of transfection, then 48 hours later cross-linked using formaldehyde and harvested for RIP. The GFP antibody was tested using conventional co-immunoprecipitation to confirm GFP was being pulled out specifically in the over-expressing cells (Figure 5.27A), and the transfection efficiency after two rounds of transfection was determined to confirm a large population of the cells were taking in and expressing the fusion protein (Figure 5.27B). The GFP RIP showed a significant enrichment of the GFP-RBMX protein at the FL-AR mRNA transcripts, but no enrichment at the AR-V7 transcript. (Figure 5.27C). This was not expected due to the downregulation in AR-V7 mRNA and protein seen in response to RBMX depletion. The reason for this could be that RBMX is acting upon pre-mRNA and is involved in early splicing events. This can be interrogated using some different primer sets that will only bind to pre-mRNA transcripts.

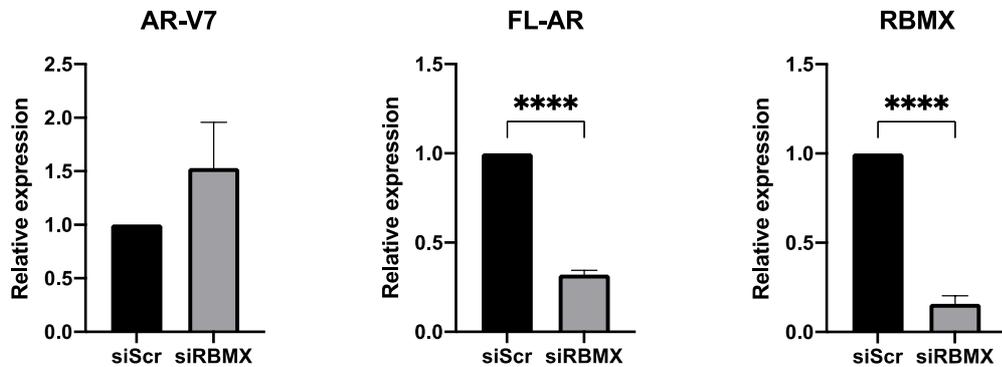


**Figure 5.27 RNA Immunoprecipitation shows RBMX enrichment at FL-AR mRNA transcripts.**

A. CWR22Rv1-AR-EK cells were reverse transfected with 5  $\mu$ g GFP-RBMX plasmid on the day of seeding and 48 hours later before being harvested for GFP IP (n=2). B. Brightfield and GFP images were taken using a Nikon2000 microscope 24 hours after the second transfection (10x magnification). C. CWR22Rv1-AR-EK cells were reverse transfected with 10  $\mu$ g GFP-RBMX plasmid on the day of seeding and 48 hours later before being harvested for RNA immunoprecipitation.

To further interrogate if RBMX is involved in splicing, a minigene reporter system was used. This is a construct that includes exons 3, CE3 and 4 of the AR gene with intronic sequences between them meaning splicing should occur to generate the products exon 3-CE3, representing mature AR-V7 transcript, or exon3-exon4, representing mature FL-AR mRNA. The construct was transfected into the AR-negative cell line PC3, because this cell line should not have detectable AR mRNA transcripts, meaning only the minigene transcripts can be analysed. 8 hours post plasmid transfection, siScrambled or siRBMX siRNAs were transfected into cells and incubated for 72 hours. qRT-PCR confirmed RBMX knockdown, and AR-V7 (exon3-CE3) and FL-AR (exon3-exon4) readouts showed RBMX depletion significantly impacted the splicing of FL-AR transcripts but not AR-V7. Consistent with the RIP experiments, this was not expected due to the decrease in AR-V7 mRNA in response to RBMX knockdown.

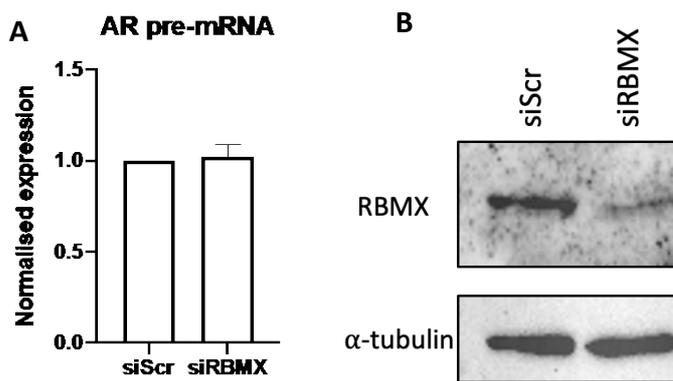
This again could be due to RBMX being involved in earlier splicing events and the intronic sequences of the minigene not containing the RNA consensus sequences to which RBMX binds.



**Figure 5.28 RBMX depletion significantly impacts splicing of FL-AR minigene.**

PC3 cells were transfected with 1  $\mu$ g AR-V7 minigene plasmid at the time of seeding prior to siScr or siRBMX transfection 8 hours later. Cells were incubated for 72 hours and harvested for qRT-PCR analysis of AR-V7, FL-AR and RBMX expression. Data represents the mean of four repeats  $\pm$  SEM. An unpaired t-test was used to determine the statistical significance (\*\*\*\* =  $p < 0.0001$ ).

To determine if RBMX depletion was causing a transcriptional effect on AR, qRT-PCR was performed using pre-mRNA primers located in intron 3 and CE3. This will provide a read out of the impact on the whole AR transcript, which if RBMX is acting co- or post-transcriptionally, the relative mRNA expression level should not change. CWR22Rv1 cells were seeded in steroid-depleted conditions to stimulate the production of AR-Vs and transfected with siScrambled or siRBMX for 72 hours. RNA was extracted using DNA-free columns to ensure DNA would not contaminate the qRT-PCR readouts so pre-mRNA could be analysed. This showed AR pre-mRNA levels do not change in response to RBMX depletion so it can be assumed RBMX does not impact transcription.



**Figure 5.29 RBMX depletion does not impact transcription of AR pre-mRNA.**

CWR22Rv1 cells were reverse transfected with siScr or siRBMX and seeded in steroid-depleted conditions for 72 hours. Cells were then subject to (A) qRT-PCR analysis of AR pre-mRNA levels or (B) western blot analysis of RBMX protein levels. Data is representative of three replicates  $\pm$  SEM and statistical significance was determined using an unpaired t-test.

## 5.4 Discussion

Alternative splicing of AR transcripts leads to the generation of AR-Vs. AR-Vs are constitutively active and drive resistance to the current repertoire of AR-targeting agents that target the LBD of FL-AR, such as enzalutamide. DNA-PKcs has been established as a FL-AR co-regulator and a driver of progression and metastasis of PCa (Goodwin *et al.*, 2015). What has not been reported in the literature is if DNA-PKcs acts as a co-regulator of AR-Vs, and if this can be attenuated through pharmacological intervention to inhibit AR-V activity. In Chapter 4, we show that DNA-PKcs manipulation has significant anti-proliferative effects and downregulates a panel of key AR/AR-V target genes in several AR-V expressing CRPC cell lines. We also show DNA-PKcs is recruited to *cis*-regulatory regions of AR/AR-V target genes so could act as an AR-V co-regulator during transcription. Here, using global transcriptomic analysis of RNA-sequencing data, we show DNA-PKcs inhibition and depletion significantly downregulates the androgen response hallmark, confirming DNA-PKcs regulation of AR-V transcriptional activity. Furthermore, the spliceosome gene list was also significantly downregulated in response to DNA-PKcs inhibition and knockdown. This was interesting due to previously identifying DNA-PKcs manipulation depletes AR-V protein levels in VCaP cells in Chapter 4. It was hypothesised that the genes from the spliceosome gene list could be involved in splicing AR pre-mRNA leading to the incorporation of CEs in mRNA that is translated into AR-Vs.

Splicing is an essential process and is required for the expression of most genes. There are several splicing patterns for the same gene which enhances the diversity of the proteome meaning there are many more proteins than there are genes in the genome. Pre-mRNA is transcribed from genes and contains coding regions (exons) and non-coding regions (introns). Coupled with transcription, the highly selective exclusion of introns and splicing of exons is necessary to enable generation of mature mRNA transcripts and subsequent translation of the protein (Bentley, 1999). The splicing process is co-ordinated by a large ribonucleoprotein complex called the spliceosome, which assembles on pre-RNA during transcription. Four spliceosome complexes, called the supraspliceosome, coordinate simultaneous exclusion of four introns as well as mRNA processing such as 3' polyadenylation, 5' capping and m6A methylation. Each spliceosomal complex is made up of five small nuclear ribonucleoproteins (snRNPs) U1, 2, 4, 5 and 6, RNA-dependent ATPases/helicases and several splicing factors

(Butcher and Brow, 2005). U1 snRNP recognition of 5' splice sites (SS) at exon-intron boundaries is required to initiate splicing (Du and Rosbash, 2002). 3' SS are recognised by U2 snRNP, splicing factor 1 and U2AFs (Wu and Manley, 1989). There are also *cis*-acting elements termed intronic splicing enhancers (ISEs), intronic splicing silencers (ISSs), exonic splicing enhancers (ESEs) and exonic splicing silencers (ESSs) that further dictate the splice site strength and usage (Sheth *et al.*, 2006). *Trans*-acting splicing regulators include heterogeneous nuclear ribonuclear proteins (hnRNPs) and serine-rich and/or arginine-rich (SR) proteins that interact with the core spliceosome complex which promote or repress splicing, depending on the *cis*-acting element they interact with (Tacke and Manley, 1999, Paschalis *et al.*, 2018). A complex multi-step process then cuts intronic regions from pre-mRNA and ligates exons together.

Deregulation of splicing can occur in disease and is commonly associated with cancer and therapy resistance. Alternatively spliced mRNAs lead to the generation of protein variants. This can be caused by alternative 3' or 5' splice site selection, mutually exclusive exons, alternative first and last exons, retained introns or exon skipping. As well as alternative splicing of the AR, there are other known splicing alterations that are implicated in cancer and therapy evasion. In melanoma, a splice variant of the BRAF variant V600E that lacks exons 4 through 8 is associated with resistance to vemurafenib due to loss of RAS signalling (Poulikakos *et al.*, 2011). In PCa, the inclusion of CEs into AR mRNA transcripts leads to premature stop codons being incorporated into transcripts and translation of truncated AR-Vs. AR-Vs lack the LBD, meaning directly targeting them is challenging. Therefore, targeting the splicing machinery that is involved in their generation is an attractive in-direct therapeutic strategy. Research into SS utilising an AR minigene reporter construct identified that intron three encompasses a 5' ISE sequence and CE3 encompasses an ESE sequence at the 3' SS that are required for splicing of exon 3 to CE3 in VCaP and LNCaP cell lines (Liu *et al.*, 2014a). The core splicing protein U2AF65 bound to the ISE and SRSF1 bound to the ESE and their depletion reduced expression of AR-V7 protein.

There are published studies that have identified genes that are altered in response to DNA-PKcs inhibition and depletion using microarray and RNA-sequencing in PCa cell lines. As previously indicated, the differentially-expressed genes in these datasets were compared to the AR-V transcriptome in CWR22RV1-AR-EK cells. This revealed that 19% of genes altered in

response to NU7441 were part of the AR-V transcriptome. There are no published whole transcriptome RNA-sequencing datasets analysing the effects of DNA-PKcs inhibition and knockdown in an AR-V driven PCa cell line. Therefore, we performed RNA-sequencing analysis of CWR22Rv1-AR-EK cells using the routinely used DNA-PKcs inhibitor NU7441, the newer, more selective compounds NU5455 and AZD7648 and DNA-PKcs depletion using siRNA. DNA-PKcs inhibition caused significant differential expression ( $FC > 1.5$  or  $< -1.5$ ,  $p < 0.05$ ) of 27 genes with NU7441, 1195 genes with NU5455, 44 genes with AZD7648, and DNA-PKcs siRNA knockdown caused differential expression of 3827 genes. RNA-sequencing data in C4-2 cells in response to NU7441, when analysed using the same bioinformatic pipeline applied to my data, caused significant differential expression of just 66 genes (Dylgjeri *et al.*, 2019). This supports our RNA-sequencing data, as NU7441 did not cause a vast number of gene expression changes, so this is most likely a genuine effect of NU7441 and not due to reduced potency of the compound. The reason for the higher number of DEG observed in response to NU5455 is not currently known. We suspect it is due to the greater selectivity of this compound causing a greater level of DNA-PKcs inhibition and therefore the enhanced transcriptional impact. It could also be due to NU5455 having a higher intracellular concentration if it is able to remain in the cells for longer and avoid metabolism. This could be investigated using MS analysis of cell lysates after different lengths of exposure to each of the DNA-PKcs inhibitors.

Comparison of genes that are differentially expressed in response to DNA-PKcs inhibition with NU7441, NU5455 and AZD7648 showed that 2 genes were common in all three experimental arms (Figure 5.9). *TWIST1* was commonly downregulated and *PKD4* was commonly upregulated in response to the three inhibitors. *TWIST1* (Twist Family BHLH (basic helix-loop-helix) Transcription Factor 1) is a transcription factor essential for embryonic development and has been implicated in cancer. It has been shown to influence many stages of cancer including initiation, primary tumour growth and metastasis (Morel *et al.*, 2012, Kang and Massague, 2004). Numerous studies have reported over-expression of *TWIST1* in many cancers and is related to a poorer prognosis and metastasis and may be used as a prognostic biomarker (Martin *et al.*, 2005, Yu *et al.*, 2010, Shibata *et al.*, 2008). *TWIST1* has also been investigated in PCa. One study reported that 90% of PCa tissues had high expression of *TWIST1* in contrast to just 6.7% of benign prostate hyperplasia. Expression was also positively

correlated with Gleason scores and metastasis through promotion of epithelial-to-mesenchymal transition (Kwok *et al.*, 2005). Epithelial-to-mesenchymal transition has been shown to be promoted in TWIST1-overexpressing cells by reducing the expression of adherens junction proteins, such as E-cadherin and N-cadherin, and inducing the expression of fibroblast markers (Yang *et al.*, 2004). A more recent study has shown TWIST1 is an androgen-regulated gene as mRNA and protein levels increase in response to the synthetic androgen R1881, and siRNA-mediated depletion of AR caused downregulation of TWIST1 mRNA in LNCaP cells (Eide *et al.*, 2013). Furthermore, the same authors showed that NKX3-1 mediated this androgen-regulation of TWIST1 as NKX3-1 binds to the *TWIST1* promoter to repress expression. DNA-PKcs expression has previously been linked to TWIST1 in cervical cancer. They showed Twist1 enhanced DNA damage repair as Twist1 knockdown cells showed persistent DNA damage after radiation that was proposed to be due to reduced nuclear levels of DNA-PKcs (Xiong *et al.*, 2017).

The *PDK4* (Pyruvate dehydrogenase kinase 4) gene encodes a mitochondrial protein that regulates glucose and fatty acid metabolism. As described by Warburg *et al.*, tumours can switch from oxidative phosphorylation to glycolysis pathways, termed “the Warburg effect” (Warburg, 1956), and PDKs, through inhibition of pyruvate dehydrogenase activity, promote this switch to cytoplasmic glycolysis (Bonnet *et al.*, 2007). *PDK4* expression is increased in colorectal patients, along with decreased methylation of CpG dinucleotides in its 5’ region in the normal colon of colorectal cancer patients compared to controls. Moreover, inhibition of *PDK4* reduces features such as migration, invasion and apoptosis *in vitro* (Leclerc *et al.*, 2017). Upregulation of *PDK4* has also been observed in high-grade bladder cancer when compared to low-grade disease, and again, inhibition and knockdown of *PDK4* inhibited bladder cancer cell proliferation *in vitro*. Additionally, combination treatment of *PDK4* inhibition and cisplatin resulted in a significant reduction in viable tumour burden *in vivo* (Woolbright *et al.*, 2018). Conversely, in a study looking at *STAT3* (signal transducer and activator of transcription 3) expression and markers associated with earlier biochemical recurrence, it was identified that low *PDK4* expression was significantly associated with a higher risk of biochemical recurrence and could be a prognostic marker in prostate cancer (Oberhuber *et al.*, 2020). PDKs are associated with therapy resistance as therapy-resistant cell lines have been either re-sensitised following knockdown or expression has been shown to be elevated in resistant cell

lines (Atas *et al.*, 2020). In contrast to *TWIST1*, downregulation of *PDK4* drives EMT in EGFR mutant lung cancer cells (Sun *et al.*, 2014).

Differentially-expressed gene lists were then input into a gene set enrichment analysis to identify alterations in pathways using KEGG and Hallmark gene sets. KEGG pathway analysis revealed pathways such as DNA replication, p53 signalling pathway, homologous recombination, oxidative phosphorylation and cell cycle were commonly de-enriched in response to both NU5455 and siDNA-PKcs (Figure 5.13). Hallmark pathway analysis showed G2M checkpoint, mTORC1 signalling, MYC and E2F targets, oestrogen response early and late, DNA repair and hypoxia. Importantly, the androgen response hallmark was significantly downregulated in response to NU5455 and DNA-PKcs depletion (Figure 5.14). As several AR and AR-V target genes are shared, this strengthened the hypothesis that DNA-PKcs is a regulator of AR-V transcriptional activity. To determine how DNA-PKcs co-regulates AR-V transcriptional activity specifically, the AR-V differentially expressed genes in response to siAR Exon1 in CWR22Rv1-AR-EK cells were compared to differentially expressed genes in response to DNA-PKcs inhibition and depletion. This showed DNA-PKcs kinase activity is involved in the expression of 11% of AR-V regulated genes and DNA-PKcs kinase-independent activity determined by siRNA-mediated depletion is involved in 41% of the AR-V transcriptome (Figure 5.10). DNA-PKcs depletion causes a greater level of AR-V transcriptional repression over inhibition as suggested by the overlapping SDEG (41% vs 11%) and the greater downregulation of the androgen response hallmark (NES -1.9 vs -1.49) (Figure 5.15). This suggests that DNA-PKcs kinase-independent activity or simply DNA-PKcs presence at *cis*-regulatory elements of AR-V target genes may have a scaffolding role that enables maximal AR-V activity. Further investigation into what enables/supports this scaffolding role could reveal targets that could enhance DNA-PKcs targeting. A comparison of AR-V and DNA-PKcs ChIP sequencing using both total DNA-PKcs and phospho-DNA-PKcs antibodies could reveal genome-wide binding sites and the overlap between cistromes would be valuable for future studies.

The spliceosome KEGG gene list was significantly downregulated in response to NU7441, NU5455 and siDNA-PKcs (Figure 5.17). This was also significantly down regulated in the DNA-PKcs RNA-sequencing data in C4-2 cells (Dylgjeri *et al.*, 2019), increasing confidence that this was a genuine DNA-PKcs-regulated gene set. This was further investigated as I hypothesised that splicing regulators that are downregulated by DNA-PKcs could be involved in the splicing

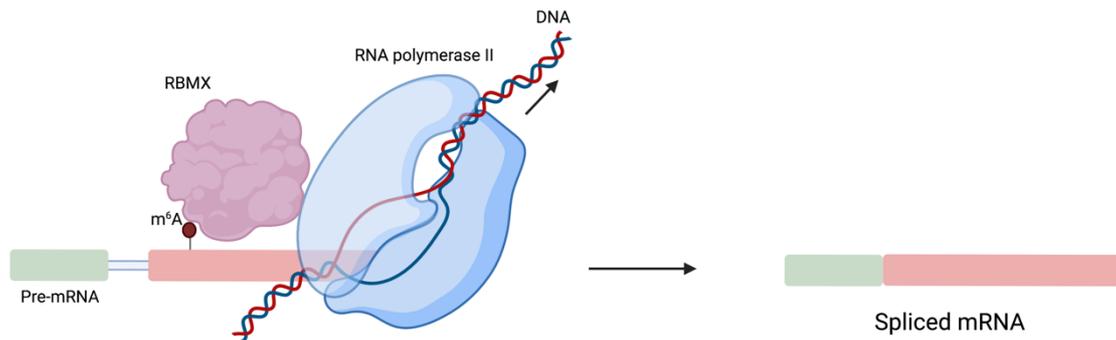
of AR-Vs. The global differential splicing effects in response to DNA-PKcs depletion was interrogated using SUPPA2. This experimental arm was taken forward to be analysed using SUPPA2 because DNA-PKcs depletion caused the greatest downregulation of the spliceosome hallmark. SUPPA2 is a bioinformatic programme that can detect changes in splicing events between two conditions. This showed that there were 358 statistically significant splicing events that occur in response to DNA-PKcs depletion. The greatest proportion of differential splicing events were alternative first exon and exon skipping (Figure 5.16).

VCaP cells do not express a high level of AR-Vs in hormone-proficient conditions, however in steroid-depleted conditions, or in response to anti-androgens, AR-Vs are upregulated. A recent publication carried out RNA-sequencing in response to the anti-androgen, darolutamide (Baumgart *et al.*, 2020). Darolutamide is a new second-generation anti-androgen that has recently been approved for use in CRPC (Fizazi *et al.*, 2019). In response to darolutamide treatment of VCaP cells, AR-V production is upregulated (Figure 5.17). The rationale for using this VCaP-Darolutamide dataset to compare with the DNA-PKcs regulated genes was based on the simple presumption that if AR-Vs are upregulated in response to darolutamide, the splicing-associated genes that also demonstrate elevated expression in response to darolutamide, could be involved in the generation and splicing of AR-Vs. Therefore, the genes that are the most downregulated in response to both DNA-PKcs inhibition and knockdown, were input as a custom gene list into the gene set enrichment tool to determine if these genes were also significantly positively-enriched differentially-expressed genes in response to darolutamide.

The 'core enriched' spliceosome genes in response to NU5455 and siDNA-PKcs treatments were enriched by 1.4 and 1.8 respectively in response to darolutamide in VCaP cells; with 34 genes commonly upregulated which were investigated further (Figure 5.17). To further refine this list of 34 genes, I determined if the genes were upregulated in tumour versus normal samples from the TCGA dataset. The top 10 genes that were upregulated by >10% were plotted in Figure 5.18. RBMX was one of the genes in this refined list and was taken forward for further validation.

RBMX, also known as hnRNPG (heterogeneous nuclear ribonucleoprotein G), is an RNA binding protein that has roles in modulating alternative splicing, genome stability and

chromatin cohesion. RBMX is 43 kDa protein located on the X chromosome and contains an N-terminal RNA recognition motif and C-terminal RNA binding domain. N6-methyladenosine (m6A) is the most abundant mRNA modification and is deposited on mRNA by a methyltransferases ‘writer’ complex that includes methyltransferase-like 3 (METTL3) and METTL14. RBMX has been reported as an m6A ‘reader’ protein. It binds purine-rich sequences that m6A modifications expose when the structure of RNA is altered, leading to modulation of alternative splicing (Liu *et al.*, 2017). Another group showed RBMX bound m6A sites near splice sites of exons associated with increased RNA polymerase II occupancy and exon inclusion (Zhou *et al.*, 2019). Additionally, RBMX interacts with the phosphorylated C-terminal domain of RNA polymerase to co-transcriptionally regulate splicing (Zhou *et al.*, 2019)(Figure 5.30). Taken together, these findings provide a mechanism by which RBMX reads and interacts with m6A sites in mRNA to modulate RNA polymerase II occupancy and alternative splicing.



**Figure 5.30 RBMX is involved in co-transcriptional splicing of pre-mRNA**

RBMX interacts with the c-terminal of RNA polymerase II and recognises m6A modifications on pre-mRNA to regulate splicing (Zhou *et al.*, 2019). Image adapted from Zhou *et al.*, 2019.

Here, we validate that RBMX is involved in the generation of AR-Vs and FL-AR. Using siRNA-mediated depletion of RBMX, we have shown this reduces FL-AR and AR-V mRNA and protein levels in CWR22Rv1-AR-EK, CWR22Rv1 and VCaP cells (Figure 5.22 and Figure 5.23). We have also shown RBMX depletion causes an anti-proliferative effect in CWR22Rv1-AR-EK cells.

To investigate if overexpression of RBMX led to the opposite effect on AR/AR-V protein and mRNA level, an RBMX-GFP construct was transfected into CWR22Rv1-AR-EK cells. However, because CWR22Rv1 cells are a difficult to transfect cell line, results were variable and did not cause a significant impact on AR-V transcripts. This experiment could be improved by selecting cells for GFP positivity to compare cells that have been successfully transfected with RBMX-

GFP to cells that express endogenous RBMX levels to determine if this causes an upregulation in AR/AR-V transcript and protein levels.

RBMX knockout using the clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated protein 9 (Cas9) system was also attempted to help validate observations made with siRBMX. However, the *RBMX* gene knockout efficiency using a sgRNA was around 20% meaning a sufficient knockout of RBMX was not achieved so could not provide a reliable read out of AR-V generation. This could be optimised by testing different sgRNAs to achieve a greater level of knockout.

To confirm that the *RBMX* gene is transcriptionally regulated by DNA-PKcs and is not a secondary or tertiary transcriptional effect when DNA-PKcs is depleted or inhibited, a DNA-PKcs ChIP was performed in CWR22Rv1-AR-EK cells. Primer sets were designed adjacent and upstream from the *RBMX* transcriptional start site to interrogate enrichment of DNA-PKcs at *cis*-regulatory elements of the *RBMX* gene. This showed DNA-PKcs is enriched near the proximal promoter regions of *RBMX*, and the level of enrichment diminished as the primer sets moved further upstream, demonstrated by the variability of percentage enrichment values at the -4kb region. The greatest level of enrichment was seen at the -500 bp and -1000 bp site, with a normalised fold enrichment over the IgG isotype control of 5.

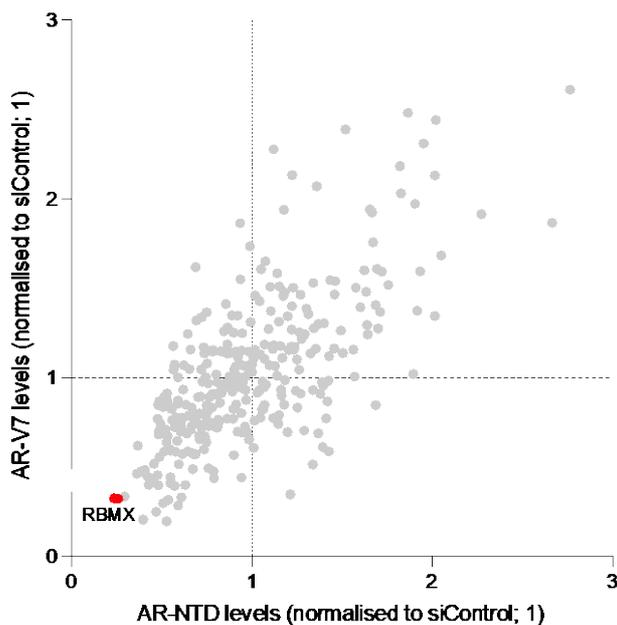
To determine if RBMX is directly involved in the splicing of AR transcripts, RNA immunoprecipitation of ectopically expressed GFP-RBMX was performed. Unexpectedly, this showed enrichment of GFP-RBMX at FL-AR mRNA transcripts but not at AR-V7 mRNA. This should be repeated to determine if RBMX is enriched at AR-V7 pre-mRNA and if splicing of AR-V7 by RBMX is an earlier splicing event to explain why there was no enrichment over the IgG isotype control at mature AR-V7 transcripts. This would reflect what has been reported in the literature where it has been shown RBMX regulates co-transcriptional splicing (Zhou *et al.*, 2019). An RBMX ChIP would also confirm if the splicing events are coupled with transcription in our cell line models of CRPC.

It was key to determine if the downregulation of AR/AR-V mRNA transcripts was due to splicing alterations and not a transcriptional effect on AR pre-mRNA transcripts. To do this, pre-mRNA primers were designed in intron 3 and CE3 so pre-spliced but post-transcription

transcripts could be measured. No change in pre-mRNA transcripts was observed, so it can be assumed RBMX is acting either co- or post-transcriptionally to splice AR/AR-V transcripts. A FL-AR/AR-V7 minigene reporter system was then utilised to validate that AR isoform splicing is controlled by RBMX. This system is a CMV driven 'mini' AR gene that contains exon 3, CE3 and exon 4 with ~800 bp intronic sequences incorporated between exons meaning splicing occurs to ligate exon 3 to CE3 (representative of AR-V7) and exon 3 to exon 4 (representative of FL-AR). The results from this experiment was contradictory to what was anticipated as AR-V7 minigene transcripts were not altered in response to RBMX depletion. A potential reason for this is that this is an artificial reporter system in PC3 cells that do not typically express AR/AR-Vs. Furthermore, the intronic sequences do not contain the entire endogenous introns of the AR gene. The consensus sequences to which RBMX typically binds could therefore be missing from the introns so alterations of splicing may not occur for this reason.

The next steps to validate RBMX as a splicing regulator would be to perform a rescue experiment to determine if in cells that have been depleted of endogenous RBMX, ectopic re-expression rescues the AR/AR-V protein levels. This could be achieved using GFP-RBMX as successfully transfected cells that have been depleted of endogenous RBMX could be compared with non-transfected cells by using the GFP tag and the levels of AR-V7 or all AR-Vs could be compared. If the AR/AR-V protein levels are higher in the GFP positive cells it could be assumed that RBMX has rescued their expression. Both immunofluorescence and flow cytometry approaches could be utilised for this experiment.

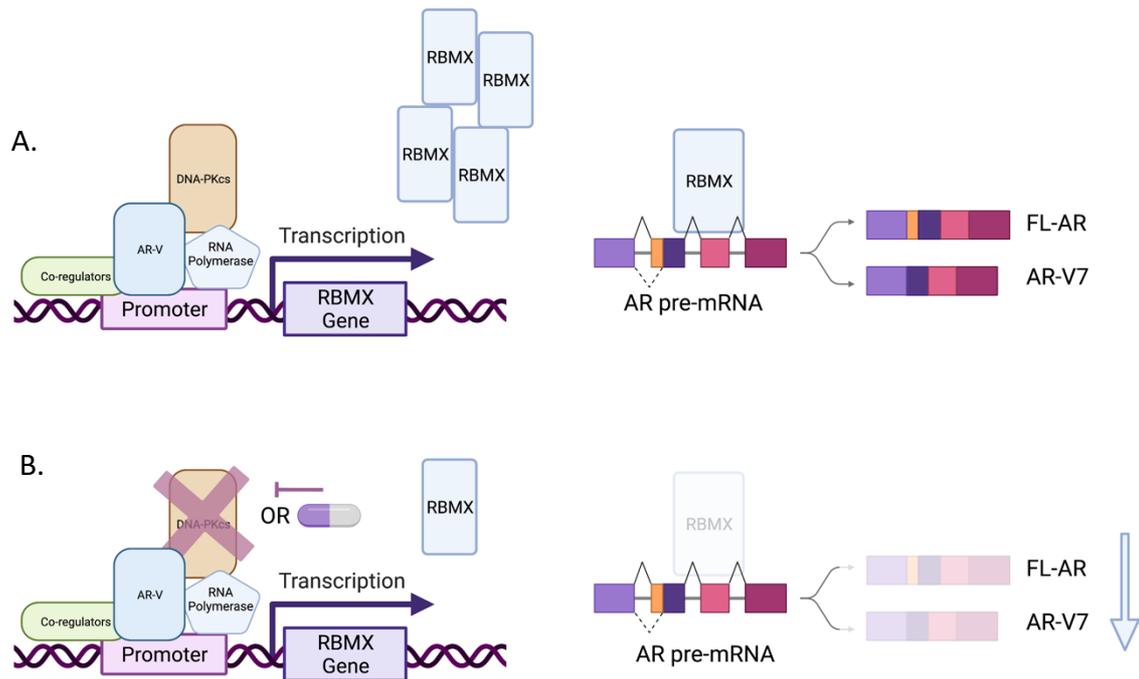
Encouragingly, a collaborator, Adam Sharp, from the Institute for Cancer Research, has performed an siRNA library screen of splicing regulators, and RBMX was one of the hits that impacted AR-V and FL-AR protein levels in CWR22Rv1 cells (Figure 5.31). This increases confidence that our results are robust as the results are consistent with what we have shown.



**Figure 5.31 RBMX depletion decreases FL-AR/AR-V and AR-V7 protein levels in CWR22Rv1 cells**

CWR22Rv1 cells were depleted of RBMX or one of several other splicing genes as part of an siRNA screen. RBMX was one of the targets that decreased expression of all AR isoforms by analysed using an N-terminal antibody and AR-V7 analysed using an AR-V7 specific antibody. This work was carried out by Adam Sharp, Institute of cancer research.

Taken together, this project has provided key mechanistic information on how DNA-PKcs inhibition and depletion causes a downregulation of AR/AR-V transcript and protein levels. AR/AR-Vs expression is altered in response to expression changes in key splicing regulators caused by DNA-PKcs manipulation. Specifically, siRNA mediated depletion of DNA-PKcs using a pool of three siRNA oligonucleotides or inhibition with NU5455 caused a downregulation of DNA-PKcs-regulated transcription of RBMX. RBMX is involved in AR/AR-V splicing, so this causes a decrease in AR/AR-V mRNA and protein expression. AR and AR-Vs are key drivers in the CRPC cell lines used so this is detrimental to the growth and survival of the cells. This proposed mechanism is presented in Figure 5.32.



**Figure 5.32 Proposed mechanism of DNA-PKcs modulated AR/AR-V splicing.**

A. In normal conditions, DNA-PKcs is involved in the transcription of the RNA binding protein RBMX, which is directly involved in splicing FL-AR and AR-V mRNA transcripts. B. In DNA-PKcs deficient conditions or when DNA-PKcs is pharmacologically inhibited, RBMX levels are downregulated which hinders splicing of FL-AR and AR-V mRNA causing their mRNA and protein levels to decrease.

## Chapter 6 Determining AR-V involvement in DNA damage repair

### 6.1 Introduction

Radiotherapy is a commonly used treatment option for PCa patients who primarily present with locally-confined disease, but more recently is also being applied to advanced metastatic lesions (Warde *et al.*, 2011). The Systemic Therapy for Advanced or Metastatic Prostate cancer: Evaluation of Drug Efficacy (STAMPEDE) clinical trial, found that radiotherapy improves overall survival from 73% to 81% at three years in selected patients that have a low metastatic burden (Parker *et al.*, 2018). However, patients can be *de novo* radio-resistant as a consequence of PCa cells acquiring a higher tolerability of DNA damage and/or of DNA being successfully repaired; preventing the catastrophic effects of pro-longed DNA damage on cell cycle checkpoint and apoptotic signalling typically observed in sensitive patients. Consistent with the theme of this thesis, it has been shown that the AR signalling pathway regulates the expression of several DDR genes, including *PARP1*, *ATR* and *CHEK1*. Polkinghorn *et al.*, identified 32 DDR-associated genes that are directly regulated by the AR in the LNCaP PCa cell line; and were down-regulated upon ADT or AR antagonist treatment (Polkinghorn *et al.*, 2013a). Pre-clinical studies demonstrated greater efficacy of ADT plus radiotherapy over single agent treatments, therefore, they suggested that patients could be stratified for ADT plus radiotherapy, rather than radiotherapy alone to improve response. Critically, this mechanistic insight into interplay between AR signalling and the DDR using pre-clinical models of PCa was driven by observations made in a number of earlier clinical trials assessing the synergistic effects of ADT and radiotherapy. Such trials have demonstrated that combining ADT plus radiotherapy improves a number of oncological outcomes in PCa patients, such as overall survival, metastasis-free survival, biochemical progression-free survival, and local failure, as reported in a meta-analysis of seven trials (Bria *et al.*, 2009). A specific example includes the phase III GETUG-AFU 16 clinical trial (NCT00423475; enrolment from 2006 – 2010) where patients were randomised to receive either short-term ADT (goserelin on the first day of irradiation and then again after 3 months) plus radiotherapy or radiotherapy alone. The 120-month progression-free survival was 64% for ADT plus radiotherapy patients and 49% for radiotherapy only patients (Carrie *et al.*, 2019). Therefore, it is evident from pre-clinical studies that the AR is important for facilitating DDR and as such, provides the mechanistic rationale

for why inhibition of the AR using ADT enhances the effect of IR. Crucially, it also indicates AR may be central to acquired resistance to IR and other DNA damaging therapies in PCa patients. Consistent with the FL-AR, identification of the AR-V transcriptome using CWR22Rv1-AR-EK cells has shown AR-Vs also control expression of numerous DDR genes involved in HR, base-excision, and mismatch repair pathways (Kounatidou *et al.*, 2019). In this study, they also compared the DDR-associated genes to an independent and in-house AR-V transcriptome by Jones *et al.* and He *et al.* (Jones *et al.*, 2015, He *et al.*, 2018). The overlaps between DDR associated genes were 59% and 95% respectively, validating that these genes are driven by AR-Vs. The FL-AR driven DDR gene signatures in VCaP cells were also compared to the Jones *et al.*, AR-V transcriptome to determine which DDR genes are regulated by both FL-AR and AR-Vs. This showed 27 out of 47 (57%) DDR genes driven by AR-Vs are also regulated by FL-AR. This data indicates that AR-V positive patients could show less favourable responses to IR plus ADT due to ADT not targeting AR-Vs. Therefore, it is important to better understand how AR-Vs control DDR and if they have a direct or indirect role in facilitating the repair of DNA breaks. Furthermore, given that ADT induces expression of AR-Vs (Yu *et al.*, 2014), longer durations of adjuvant ADT treatment prior to radiotherapy may diminish efficacy of radiotherapy if AR-Vs are contributing to resolving DNA damage in response to radiotherapy.

AR-Vs have been shown to interact with key DDR-associated proteins, including DNA-PKcs and  $\gamma$ H2AX, post-treatment with DNA damaging agents, such as IR (Yin *et al.*, 2017). This aforementioned study used an IP-MS approach to define the interactome of AR<sup>V567es</sup> in the presence and absence of IR, using the cell line R1-D567 which expresses only the single AR-V, AR<sup>V567es</sup> and no FL-AR. This revealed DNA-PKcs was the most abundant interactor of AR<sup>V567es</sup> in both steady-state and in response to DNA damage, and was validated using co-IP which demonstrated an enhanced interaction between AR<sup>V567es</sup> and DNA-PKcs in response to IR. Intriguingly, it was shown using proximity ligation assays (PLA) that AR<sup>V567es</sup> directly interacted with  $\gamma$ H2AX upon IR treatment and this was diminished upon DNA-PKcs inhibition suggesting AR-Vs interact with sites of DNA damage and this is dependent upon the kinase activity of DNA-PKcs. Current literature supports the concept that AR-Vs can enhance DNA damage repair by (i) up-regulating expression of DDR-associated genes (Kounatidou *et al.*, 2019); and, although less defined, (ii) interacting with sites of damage to support repair of DNA lesions directly (Yin *et al.*, 2017).

As such, there remains a major knowledge gap in our understanding of the role of AR-Vs in DNA damage repair, particularly if they play a specific role at the damaged DNA locus. To this end, by conducting an unbiased AR-V interactome study using a novel proximity-biotinylation labelling approach in the presence and absence of IR, I aim to provide key insight into the role of AR-Vs in direct DNA repair outside of their characterised role as transcriptional regulators of DDR-associated genes.

Proximity labelling has several advantages over antibody-based IP-MS-based approaches for identification of candidate protein interactomes. These include reduced likelihood of weaker or transient interactions being missed and non-specific proteins being immunoprecipitated. In recent years, proximity-based labelling techniques have been developed to chemically tag proteins in proximity of a bait protein to determine its interactome in live cells. This novel method was first developed incorporating a BioID biotin protein ligase tagging protocol that directly labels proteins in close proximity (~10-15 nm) to the bait protein with biotin (Kim *et al.*, 2014). This novel approach enables the detection of weak and transient bait-interacting protein interactions that could be missed using conventional antibody-based IP methods (Roux *et al.*, 2012). The strong affinity of streptavidin for biotin is then utilised to pull-down biotinylated proteins that can then be identified using MS-based techniques. More recently, a technique using ascorbate peroxidase (APEX2) was developed for the proteomic mapping of mitochondria, and of mitochondrial DNA in living cells (Rhee *et al.*, 2013, Han *et al.*, 2017). APEX2 has advantages over BioID as it is considerably quicker to label interacting proteins, occurring in minutes instead of hours; and APEX2 is smaller in size (28 kDa versus 35 kDa of the BioID ligase) which reduces impact on cellular activity of the tagged/bait protein. APEX2, with the addition of the co-substrate hydrogen peroxide (H<sub>2</sub>O<sub>2</sub>) and biotin-tyramide, catalyses the oxidation of biotin-phenol to a short-lived biotin-phenoxy radical, resulting in biotinylation of neighbouring proteins within 20 nm by reaction with electron-rich amino acids, primarily tyrosine (Hung *et al.*, 2016).

So far, unbiased IP-based approaches have been used to define the AR and AR-V interactomes, using both IP-MS and RIME. A recent AR and AR<sup>V567es</sup> interactome defined by RIME was performed in R1-AD1 and R1-D567 PCa cells, and GRHL2 was identified as a novel AR and AR-V binding partner; which was further validated using co-IP (Paltoglou *et al.*, 2017). Furthermore, a biotinylation approach has been utilised to define the interactome of

overexpressed FL-AR in response to DHT treatment using BioID in HEK293 cells which identified 32 DHT-dependent AR-proximity interactions (Lempiainen *et al.*, 2017). This was also employed in the androgen dependent cell line LAPC4 that stably express the transgene BirA-AR to define the AR interactome in an androgen responsive cell line. This approach identified a novel AR binding partner, Krüppel-like factor 4, KLF4 (Velot *et al.*, 2021). Although these novel approaches have provided key protein-protein interactions of the FL-AR, there are no published AR-V interactomes that use the proximity biotin-labelling methodology. In the host laboratory, RIME approaches have been used to begin to define the AR-V7 interactome, however, like previously mentioned, there are technical limitations to this approach meaning functionally important proteins could be missed using this method. Furthermore, the AR-V7 interactome in both the presence or absence of IR is yet to be determined using a proximal biotinylation approach which could help better understand AR-V function during the DDR.

#### 6.1.1 CRISPR/Cas9 genome editing

The ability to precisely edit the genome has been one of the most important advancements in science and research over the past couple of decades. It allows the introduction of genomic alterations in the form of insertions and/or deletions (indels) or base substitutions. There are several techniques that have been developed, including zinc finger nucleases, transcriptional activator like effector nucleases (TALENs) and the CRISPR/Cas9 system. The CRISPR/Cas9 system is the most recently discovered technique to edit genomes and has many advantages over the older gene editing systems, including being less labour intensive, less expensive and less time consuming. The CRISPR/Cas9 system originates from the bacterial adaptive immune system that utilises a number of specific DNA endonucleases to degrade bacteriophage genomic DNA to prevent re-infection of the host cell. Critically, this system was harnessed for use in genomic editing by determining that the Cas9 nuclease can be guided to specific regions of the genome (Jinek *et al.*, 2012). CRISPR pioneers Emmanuelle Charpentier, Jennifer Doudna and Feng Zhang discovered that the system can be adapted for human genome editing (Cong *et al.*, 2013, Jinek *et al.*, 2013).

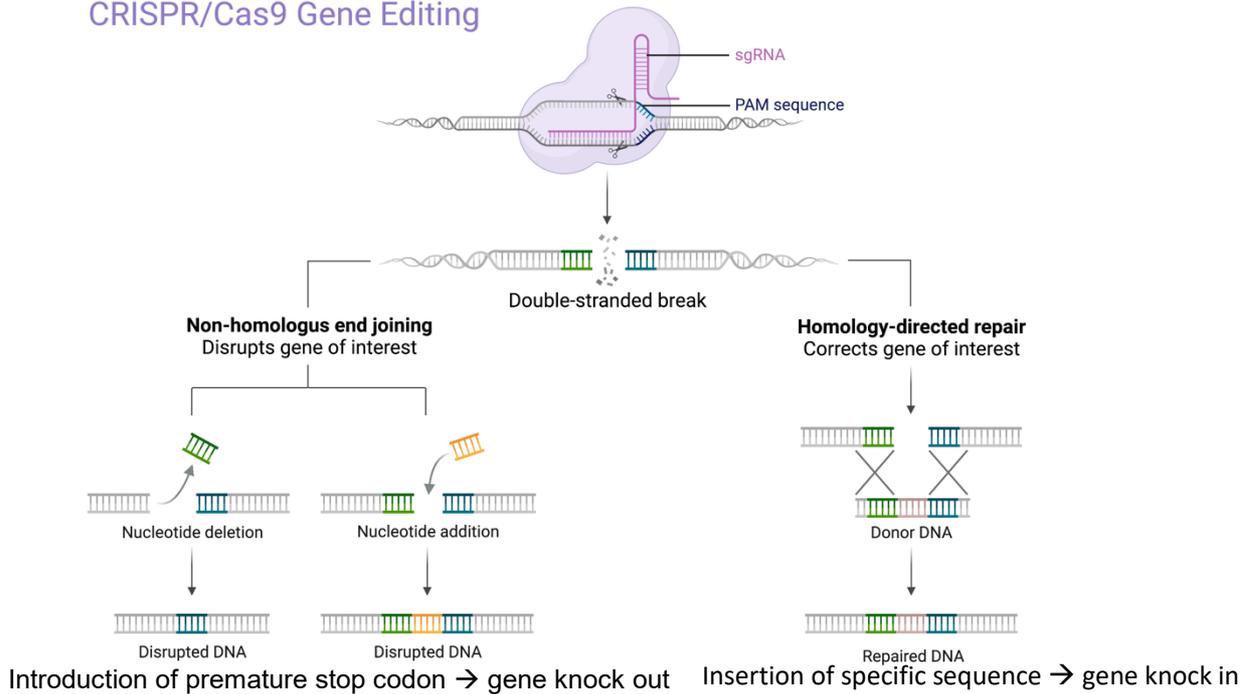
Cas9 is an RNA guided DNA endonuclease and has the DNA cleavage domains RuvC and His-Asn-His (HNH), which cuts the dsDNA site located 3 base pairs upstream of protospacer

adjacent motif (PAM) sequences (NGG) in the target DNA (Mojica *et al.*, 2009, Westra *et al.*, 2013). To cleave the DNA efficiently and specifically at the correct site, the Cas9 protein forms a ribonucleoprotein complex with CRISPR RNA (crRNA) complementary to the target locus and transactivating RNA (tracrRNA) that binds to the crRNA and Cas9. Once the DNA has been cleaved, it will be repaired by one of two endogenous DNA repair pathways: (i) the error prone NHEJ pathway; or (ii) the homology-directed repair (HDR) pathway. By exploiting the cells own DNA repair pathways, application of CRISPR/Cas9 has enabled targeted gene knockout and precise knock-in approaches to be successfully undertaken in mammalian cell lines. To knockout a specific gene of interest, Cas9 is targeted to a desired locus, using gRNA, to cleave the DNA, which will primarily be repaired by NHEJ given most cells will be in interphase. Because NHEJ is error prone, this leads to introduction of indels into the target locus, disrupting the reading frame of the gene and causing incorporation of premature stop codons and ultimately knockout of the protein of interest.

For targeted gene knock-in, Cas9 is directed to a specific locus and cleaves the DNA. Then, by providing cells with a repair template that encompasses regions that are homologous to the region of interest, the sequence can be inserted (knocked-in) into the genome by cells that undergo the HDR pathway (Figure 6.1). This approach has allowed generation of endogenous reporter proteins, and introduction of clinically relevant mutations to better study functions of proteins. An example of this knock-in approach in the host laboratory is the generation of the CWR22Rv1-AR-EK cell line, whereby a donor template was introduced into CWR22Rv1 cells that knocked-in a sequence to introduce a stop codon in exon 5 of the *AR* gene to maintain expression of all endogenous AR-Vs without expression of FL-AR (Kounatidou *et al.*, 2019).

In recent years there have been several adaptations to the CRISPR/Cas9 system including CRISPR activation and CRISPR inhibition, termed CRISPR interference (CRISPRi), by using a catalytically dead version of Cas9 (dCas9) that lacks nuclease activity so does not cleave DNA (Qi *et al.*, 2013). This allows specific upregulation or downregulation of a gene of interest. There are also epigenome Cas9 editors where dCas9 is fused to epigenetic regulators such as a methyltransferase or demethylase enzymes and directed to a specific locus to alter gene expression (Liu *et al.*, 2016).

## CRISPR/Cas9 Gene Editing



**Figure 6.1 Mechanism of CRISPR genome editing**

A sgRNA guides Cas9 to a region of interest in the genome. Cas9 then cleaves the DNA three base pairs upstream from the PAM site, causing a double stranded DNA break. This leads to initiation of the DNA damage response. In the absence of a homologous repair template, the error prone non-homologous end joining pathway causes nucleotide deletions or insertions causing disruption of the gene. This gives rise to successful gene knockouts. In the presence of a homologous repair template, homologous recombination accurately repairs the DNA which can introduce specific gene edits or insertions of sequences at precise loci. This enable precise CRISPR-directed gene knock-in editing. Template taken from Biorender.com

### 6.1.2 Aims

This project aims to provide for the first time an AR-V interactome in the presence and absence of DNA-damaging IR treatment using proximity biotinylation labelling. This objective has the capacity to increase our understanding of the role of AR-Vs in steady-state transcriptional regulation and during the DDR.

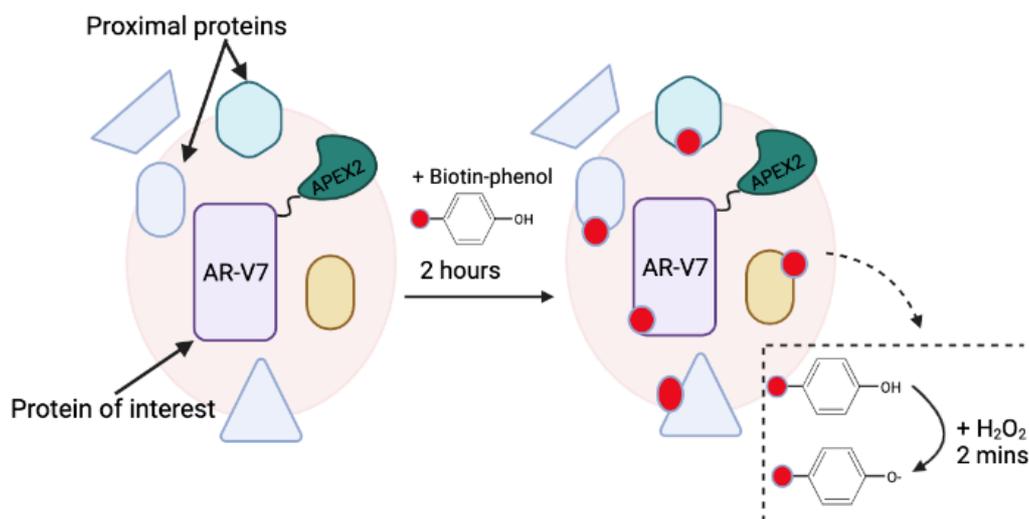
This will be done by:

- i. Attempted generation of an APEX2-AR-V knock-in CWR22Rv1-AR-EK cell line derivative using CRISPR/Cas9 genome editing techniques to provide interactomes of endogenous AR-Vs.
- ii. As a contingency to the above objective, generation and application of an ectopically-expressed APEX2-AR-V7 fusion in CWR22Rv1-AR-EK cells treated with and without IR will provide proximity biotinylation-based interactomes of AR-V7 in steady-state and upon DNA damage.

## 6.2 Specific materials and methods

### 6.2.1 Biotin labelling protocol

$4 \times 10^6$  CWR22Rv1-AR-EK cells were transfected with 10  $\mu\text{g}$  APEX2-AR-V7 plasmid (details of cloning in section 6.3.3) using LT1 transfection reagent in 15 cm dishes and incubated for 48 hours before a second transfection of another 10  $\mu\text{g}$  of plasmid (protocol detailed in section 3.10.4). Cells were incubated for a further 24 hours before being irradiated with 4 Gy radiation. Immediately after irradiation, biotin-phenol (Iris Biotech LS-3500) was added to the media to a final concentration of 500  $\mu\text{M}$  and incubated at 37 °C for 2 hours.  $\text{H}_2\text{O}_2$  (Sigma, 30%) was then added directly to the media to a final concentration of 1 mM, except for the control, and incubated for 2 minutes whilst continuously rotating the dishes to induce the labelling reaction (Figure 6.2). The reaction was then quenched by replacing the media with a quenching buffer containing 100 mM sodium ascorbate, 10 mM TROLOX and 10 mM sodium azide in PBS. This buffer was replaced 4 times and then cells were washed a further 4 times in PBS before being trypsinised, neutralised with media and then transferred to 15 mL falcon tubes before being centrifuged at 500 x g for 10 minutes. The supernatant was then aspirated, and cell pellets were subject to cytoplasmic nuclear extraction (Figure 6.3).



Created in **BioRender.com** 

**Figure 6.2 Diagrammatic representation of proximity labelling reaction.**

APEX2 is fused to a protein of interest (AR-V7) and expressed in cells. Subsequent biotin-labelling occurs following the addition of biotin-phenol and hydrogen-peroxide (H<sub>2</sub>O<sub>2</sub>) to culture media to catalyse the conversion of biotin-phenol to its reactive radical that can biotinylate proximal proteins in a labelling radius of 20 nm.

#### 6.2.2 Cytoplasmic nuclear extraction

Immediately after harvesting cell pellets post-biotin labelling protocol, the cytoplasmic and nuclear cellular fractions were isolated. This was achieved using the NE-PER™ Nuclear and Cytoplasmic Extraction Kit, as per manufacturer's instructions (Thermo Scientific). For cells grown on 15 cm dishes, the reagent amounts for a packed cell volume of 50 μL was applied. The fractionated samples were then stored at – 80 °C until further processing.

#### 6.2.3 Pierce 660nm to determine concentration

Post-labelling and harvesting, either whole cell or cytoplasmic and nuclear lysates were subject to protein quantification using the Pierce™ 660nm Protein Assay (Thermo Scientific), as per the manufacturer's instructions, using a 96-well plate assay format. Whole cell lysates were diluted 1:10 and cytoplasmic and nuclear extracts were quantified without diluting. Absorbance was measured at 660 nm excitation using a microplate reader (Bio-Rad) and the concentrations were extrapolated using a standard curve.

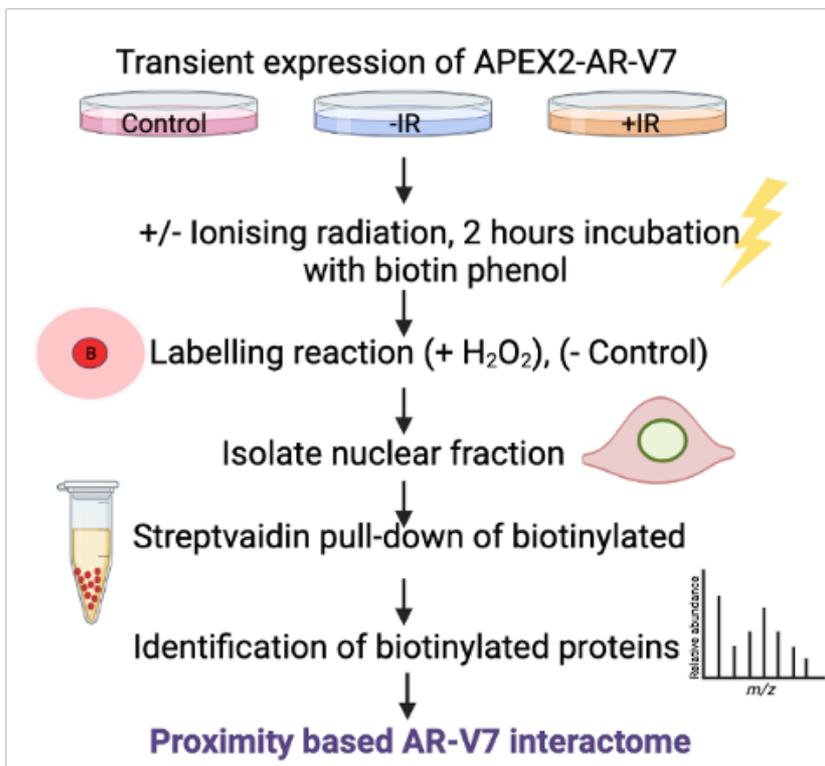


Figure 6.3 Flow chart of proximity labelling and identification protocol.

Cells are transfected with the APEX2-AR-V7 fusion construct, incubated to allow expression of the fusion protein for 72 hours before induction of DNA damage using ionising radiation. Cells are then incubated for 2 hours with the addition of biotin phenol to allow DNA repair to begin and biotin phenol to enter the nucleus. Cells are then harvested before nuclear isolation and enrichment of biotinylated proteins using streptavidin beads.

#### 6.2.4 Streptavidin enrichment and initial analysis of biotinylation by western blot

For initial experiments in HEK293T cells, 360  $\mu\text{g}$  of whole cell lysate was added to 30  $\mu\text{L}$  streptavidin beads (Life Sciences) that had been pre-washed with RIPA buffer (Table 3.9). A 10  $\mu\text{L}$  input was also taken from the initial lysates to confirm the labelling reaction had taken place. The beads and lysates were then incubated overnight at 4  $^{\circ}\text{C}$  on a rotating wheel, before removing the supernatant using magnetic separation, but retaining a small amount as a 'flowthrough' sample. The beads were then washed 7 times with RIPA buffer and then boiled at 100  $^{\circ}\text{C}$  in SDS-sample buffer containing 2 mM biotin and 20 mM DTT for 15 minutes to elute the proteins (Table 3.10).

For samples subject to cytoplasmic and nuclear fractionation, 100-150  $\mu\text{g}$  of nuclear extract was incubated with 30  $\mu\text{L}$  streptavidin beads plus NE-PER™ nuclear extraction reagent to a

final volume of 500  $\mu$ L and incubated overnight at 4 °C on a rotating wheel. The cytoplasmic, nuclear and insoluble fractions were retained for downstream western blot analysis of enriched proteins.

#### 6.2.5 Preparation of samples for mass spectrometry and mass spectrometry data analysis

For samples being sent for MS analysis, beads were washed 7 times with RIPA buffer then 7 times in PBS in a laminar flow hood using filter tips to minimise keratin contamination to samples before being resuspended in 1 mL PBS. Samples were then delivered to the Glasgow MS facility (Polyomics) on dry ice. The samples were then subject to on-bead tryptic digestion followed by acidification with CF<sub>3</sub>COOH and dried down in a vacuum centrifuge ready for MS. Dry peptides residues were solubilized in 20  $\mu$ L 5% acetonitrile with 0.5% formic acid using the auto-sampler of a nanoflow uHPLC system (Thermo Scientific RSLCnano). Online detection of peptide ions was by electrospray ionisation MS/MS with an Orbitrap Elite MS (Thermo Scientific). An injection volume of 5  $\mu$ L of the reconstituted protein digest were desalted and concentrated for 10 min on trap column (0.3  $\times$  5 mm) using a flow rate of 25  $\mu$ L / min with 1% acetonitrile with 0.1% formic acid. Peptide separation was performed on a Pepmap C18 reversed phase column (50 cm  $\times$  75  $\mu$ m, particle size 3  $\mu$ m, pore size 100 Å, Thermo Scientific) using a solvent gradient at a fixed solvent flow rate of 0.3  $\mu$ L / min for the analytical column. The solvent composition was A) 0.1 % formic acid in water B) 0.08 % formic acid in 80% acetonitrile 20% water. The solvent gradient was 4% B for 12 min, 4 to 60% for 90 min, 60 to 99% for 14 min and held at 99% for 5 min. A further 9 minutes at initial conditions for column re-equilibration was used before the next injection.

Protein identifications were assigned using the Mascot search engine (v2.6.2, Matrix Science) to interrogate protein sequences in the Swissprot database using Homo sapiens taxonomy. A mass tolerance of 10 ppm was allowed for the precursor and 0.3 Da for MS/MS matching. Lists were compared between replicates and compared to control samples. Alternative data processing was performed by inputting raw files to MaxQuant to identify and map peptides to proteins and provide intensity based absolute quantification (iBAQ) values. The iBAQ quantification method sums all the peptides intensities and divides by the number of detectable peptides of a protein. An inhouse R Script was then used to filter resulting lists to eliminate contaminants and proteins that were identified with <2 unique peptides. Peptide

intensity values from control samples were subtracted from plus IR and minus IR lists. Relative iBAQ (riBAQ) values were then calculated to provide final lists of proteins. The difference between the mean riBAQ value of the three plus IR samples and mean riBAQ value of the three minus IR samples was calculated to give the FC between the two conditions. Statistical significance was determined using individual paired t-tests using PRISM.

#### 6.2.6 Stable cell line generation

##### mSA insertion to TLCV2 plasmid

Streptavidin-fused Cas9 protein has been shown to increase efficiency of CRISPR/Cas9 knock-in by utilising biotinylated donor templates that will have a high affinity for Cas9-streptavidin (Gu *et al.*, 2018). Therefore, the monomeric streptavidin (mSA) sequence was inserted into the pTLCV2-Cas9 plasmid downstream and in-frame with Cas9-Flag cDNA (Addgene #87360). An mSA fragment was digested from the plasmid pSpCas9-mSA-2A-Puro using a *Bam* HI restriction enzyme (Invitrogen) according to manufacturer's instructions and was gel extracted using a NEB gel purification kit. The pTLCV2-Cas9 plasmid was also digested using *Bam* HI and gel extracted. The donor pTLCV2-Cas9 plasmid and mSA insert were ligated using NEB T4 ligase according to manufacturer's instructions (NEB). The resulting ligation reactions were transformed into *Stb13* cells and plated onto ampicillin-containing LB agar plates to select positive recombinants. After overnight incubation, resulting colonies were screened using colony PCR using Cas9 forward and mSA reverse primers; amplification will only take place if the insert has been successfully ligated into the donor plasmid. A positive clone was sent for Sanger sequencing (GeneWiz) to confirm insertion of the mSA fragment downstream from the Cas9-Flag cDNA. Once confirmed, the plasmid was transfected into CWR22Rv1-AR-EK cells to determine if the Cas9-mSA protein was expressed using western blotting incorporating a Cas9 antibody. The parent pTLCV2-Cas9 plasmid was used as a control to enable comparison of the unmodified Cas9 with the newly generated streptavidin-tagged Cas9 fusion.

##### Inducible Cas9-mSA cell line generation

To generate and inducible CWR22Rv1-AR-EK-Cas-9-mSA cell line,  $4 \times 10^6$  CWR22Rv1-AR-EK cells were seeded into 10 cm dishes and transduced with 1 mL unconcentrated TLCV2-mSA lentivirus 24 hours later (Lentiviral generation and transduction protocol detailed in section 3.14). The cells were then incubated for 7 days to allow integration of the transgene and

resistance marker. 1 µg/mL puromycin was then added to the media to select for cells that have stably integrated the Cas9-mSA/puromycin-resistance cassette. Colonies were allowed to grow and were picked and bulked up to screen for Cas9 expression by western blotting.

#### 6.2.7 dsODN generation

##### Amplification of Puro-T2A-APEX2-GS linker gBlock

A gBlock™ gene fragment (Integrated DNA Technologies) encoding a puromycin resistance cassette, T2A linker and APEX2 sequence was PCR amplified using biotinylated primers (Sigma) that contain 50 bp sequences complementary to the AR sequence adjacent to the Cas9/gRNA cut-site using Phusion PCR Master Mix 2x (NewEngland Biolabs). The following thermal profile was performed on a GeneAMP 2700 thermal cycler (ABI): initial denaturation at 98°C for 30 seconds, 35 cycles of 98°C for 10 seconds, 72°C for 30 seconds , 72°C for 1 min 30 seconds, followed by a final extensions at 72°C for 7 minutes. The PCR amplicons from four 50 µL PCR reactions were combined and purified using a QIAquick PCR clean up kit (Qiagen) and quantified using a nanodrop spectrophotometer.

##### Amplification of RFP-T2A-APEX2-GS linker gBlock

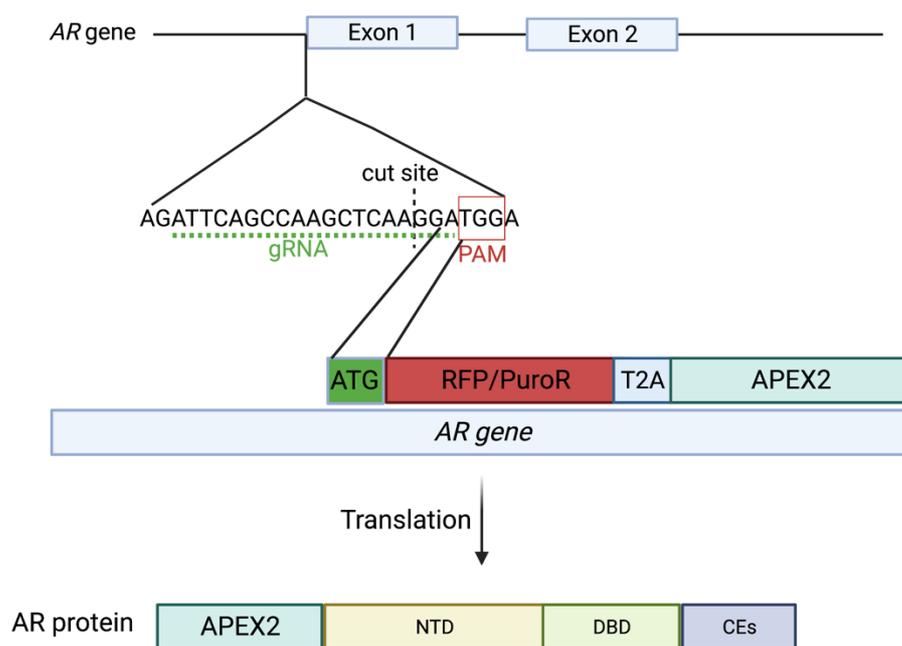
A gBlock™ gene fragment (Integrated DNA Technologies) encoding a Turbo RFP cassette, T2A linker and APEX2 sequence was PCR amplified using biotinylated primers (Sigma) that contain 50 bp sequences complementary to the AR sequence adjacent to the Cas9/gRNA cut-site using Phusion PCR Master Mix 2x (NewEngland Biolabs). The following thermal profile was performed on a GeneAMP 2700 thermal cycler (ABI): initial denaturation at 98 °C for 30 seconds, 20 cycles of 98 °C for 10 seconds, 55 °C for 30 seconds , 72 °C for 1 min 30 seconds, followed by a final extensions at 72 °C for 7 minutes. The PCR amplicons from four 50 µL PCR reactions were combined and purified using a QIAquick PCR clean up kit (Qiagen) and quantified using a nanodrop spectrophotometer.

**Table 6.1 Primer Sequences to generate the dsODN template for the APEX2 knock in.**

Forward (5' – 3'): Homology arm- <b>Puro</b> -APEX gBlock	[Btm]GGAGGCGGGGTAAGGGAAGTAGGTGGAAGATTCAGCCA AGCTCAAGGATGACAGAATATAAACCAACAGT
Reverse (5' – 3'): Homology arm- <b>Puro</b> -APEX gBlock	GTCTTGACGGCGGCCGAGGGTAGACCCTTCCCAGCCCTAACT GCACTTCTCTGAACCTGAACCAGATC
Forward (5' – 3'): Homology arm- <b>RFP</b> -APEX gBlock	[Btm]GGAGGCGGGGTAAGGGAAGTAGGTGGAAGATTCAGCCA AGCTCAAGGATGAGCGAGCTGATCAAGGAGAA
Reverse (5' – 3'): Homology arm- <b>RFP</b> -APEX gBlock	GTCTTGACGGCGGCCGAGGGTAGACCCTTCCCAGCCCTAACT GCACTTCTCTGAACCTGAACCAGATC

### 6.3 Results

To determine the interactome of AR-Vs in steady state and post-IR, introduction of the APEX2 sequence upstream of the *AR* gene in CWR22Rv1-AR-EK cells using CRISPR/Cas9 was attempted (Figure 6.4). Using endogenous AR-V levels to label their interactome is advantageous due to it being in a physiological background. This will also mean 100% of the cells express APEX2-AR-Vs so the quantity of biotinylated proteins that are enriched will be sufficiently robust to enable detection of proteins that have a lower abundance by MS. This novel cell line derivative can also be used to further validate previous AR-V interactome data sets that have been identified using exogenous over-expression of FL-AR or AR-Vs using either antibody-based or RIME detection techniques. Three different approaches for APEX2 gene knock-in were attempted over the course of approximately 16 months, but due to the low knock-in efficiency, possibly as a consequence of the size of the HDR donor template needing to be introduced, this was not successful. Each approach is described below.



**Figure 6.4 AR gene for insertion of APEX2 at the start codon of AR exon 1.**

The cut site for insertion of the RFP/PuroR-T2A-APEX2 sequence at the start codon of AR exon 1 with the sgRNA used, highlighting the PAM site. Subsequent translated AR-V protein will incorporate APEX2-Linker-NTD-DBD-Cryptic exons (Ces).

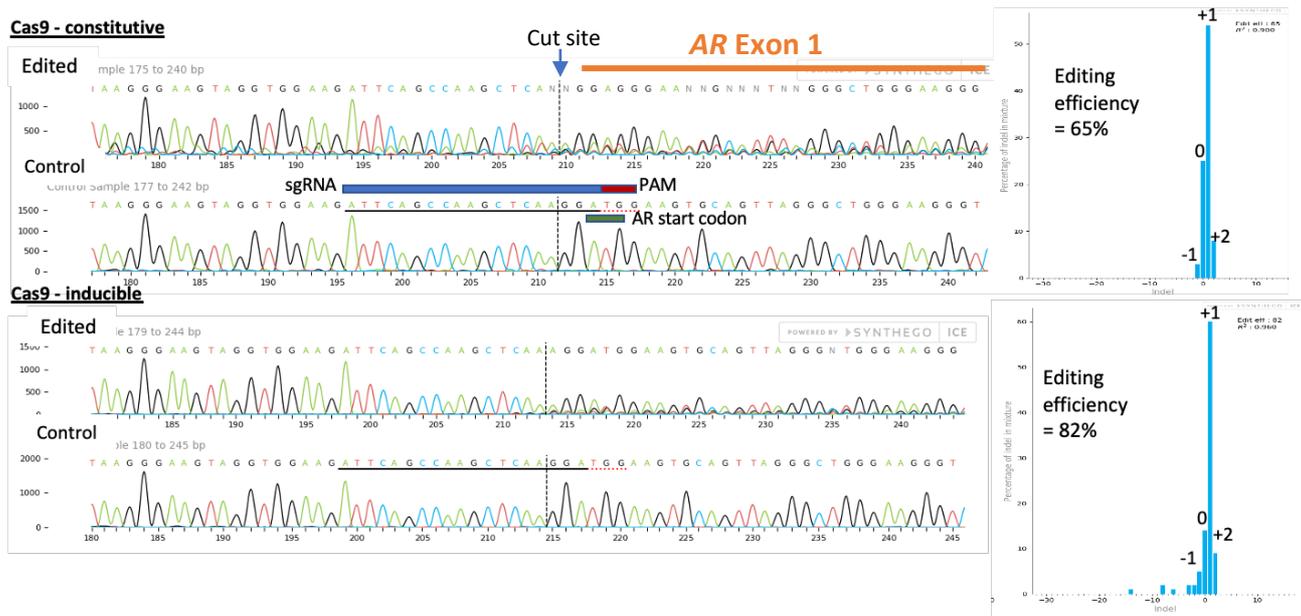
### 6.3.1 Validating sgRNAs

To validate the efficiency of an AR targeting sgRNA (Sigma), both constitutive and inducible Cas9 cell lines generated in house by Laura Walker (fellow PhD student) were used. Briefly, the cell lines were generated using lentiviral transduction of the doxycycline-inducible-LCV2 (TLCV2) and LCV2 constructs into CWR22Rv1-AR-EK cells which were then puromycin selected, and an individual clonal population was validated for Cas9 expression. The inducible cell line was treated with doxycycline for 24 hours then both cell lines were reverse transfected with 25 nM sgRNA and incubated for 72 hours. Cell pellets were then harvested, genomic DNA was isolated and the region encompassing the AR start codon adjacent to the target site of the sgRNA was amplified by PCR using the primers in Table 6.2 and resulting amplicons were then gel extracted and sent for Sanger sequencing. The chromatograms from both the Cas9 only control and the Cas9 plus AR sgRNA were compared to identify the percentage indel efficiency by Cas9 at the specified region. Synthego ICE analysis calculated the indel efficiency and the nature of the introduced indels (for example -1, +1, +2 (base pairs)), which ultimately indicates

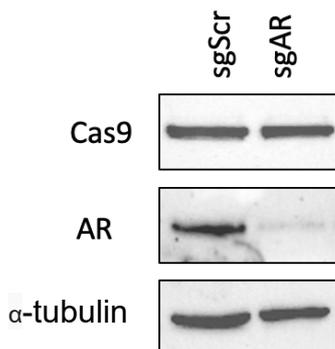
if the indels are in-frame or if they disrupt the reading frame. Importantly, ICE analysis of the amplified AR exon 1 locus post editing showed the editing efficiency for the constitutive and inducible cell lines was 65% and 82%, respectively which was further supported by demonstrating a robust reduction of AR expression (Figure 6.5) in cells transfected with the AR-targeting sgRNA. These validated sgRNAs were then incorporated into the knock-in pipeline to generate the APEX2-AR-V CWR22Rv1-AR-EK cell derivative.

**Table 6.2 Primer sequences used to amplify region around sgRNA mediated cut side of AR gene for TIDE analysis**

<b>Gene</b>	<b>Forward (5'-3')</b>	<b>Reverse (5'-3')</b>
<i>AR</i>	GAGACAGACTGTGAGCCTA	GCTCTGGAACAGATTCTGGAA



**CWR22Rv1-AR-EK-iCas9**



**Figure 6.5 Editing efficiency in CWR22Rv1-AR-EK-Cas9 and CWR22Rv1-AR-EK-iCas9 cell lines using AR Exon 1 targeting sgRNA.**

Cells were seeded in 6-well plates and transfected with 25 nM single guide (sg) scrambled (sgScr) or AR-targeting (sgAR) RNAs. The CWR22Rv1-AR-EK-iCas9 cell line is treated with doxycycline for 24 hours prior to seeding and transfection with sgRNAs. Cells were then harvested, the genomic DNA was isolated and a DNA amplicon incorporating the cut site was amplified using PCR and sequenced for indel analysis using ICE (Synthego). Cells were harvested for western blot analysis 5 days post sgRNA transfection incorporating anti-AR and  $\alpha$ -tubulin antibodies.

Due to the low transfection efficiency of the CWR22Rv1-AR-EK cells, which would reduce the take up of Cas9-expressing plasmids hence compromising the overall knock-in efficiency of the puromycin/RFP-APEX2 cassette, a stable CWR22Rv1-AR-EK-iCas9-mSA derivative cell line was sought that expresses Flag- and streptavidin (mSA)-tagged Cas9 in response to doxycycline. This was generated by firstly inserting the mSA sequence into the pTLCV2-Cas9 vector downstream of the Cas9-Flag cDNA (Figure 6.6) then generation of lentivirus for transduction of CWR22Rv1-AR-EK cells. The cells were then puromycin selected and GFP-positive clone 6 (C6) was validated for mSA-tagged Cas9 expression by inducing the cells with doxycycline then using western blot analysis, incorporating a Flag antibody to detect the Flag-tagged mSA-Cas9 protein (Figure 6.7A), and immunofluorescence to detect GFP which is a surrogate marker of mSA-Cas9 expression (Figure 6.7B)(GFP cDNA is located downstream from the Cas9 cDNA and a T2A sequence; enabling transcription and translation of Cas9 and GFP as a composite which then splits by intracellular cleavage of the T2A peptide). A successful doxycycline-inducible mSA-Cas9 expressing clone was then taken forward. Transfection with the on-target AR-targeting sgRNA was used to validate the functionality of the mSA-Cas9 enzyme and ensure mSA attachment to the Cas9 C-terminus did not impact its editing capability. Using Synthego ICE analysis, this showed the editing efficiency was between 44-53% across two replicates (Figure 6.7C). This suggests mSA slightly impacts the nuclease activity of the Cas9 enzyme, however the efficiency is still sufficient to lead to a significant knockout of AR, as can be seen by assessing AR protein levels post sgRNA-mediated AR knockout using western blot analysis(Figure 6.7A).

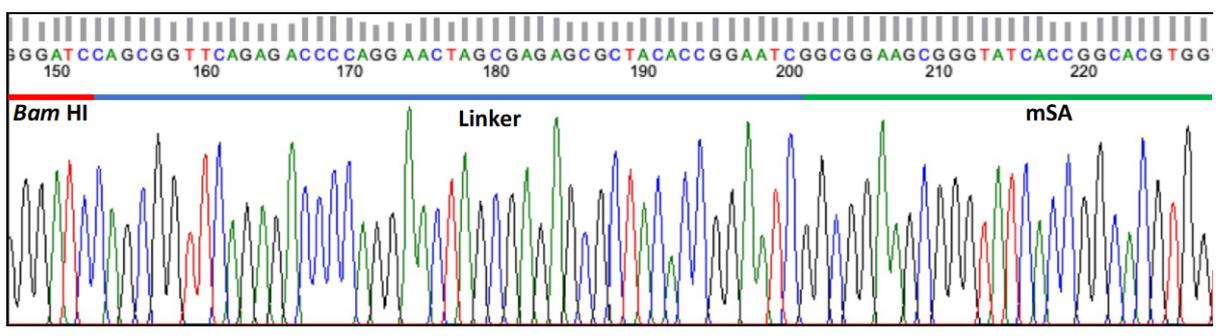
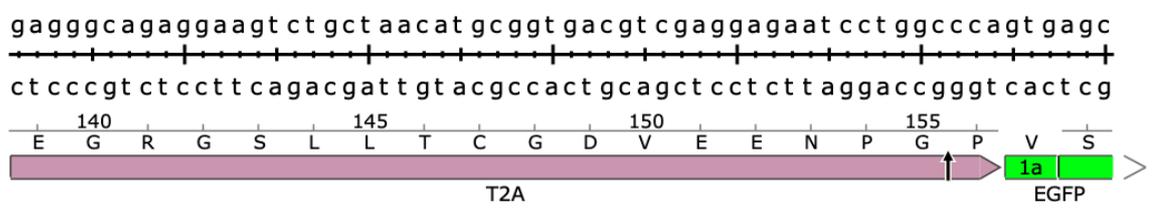
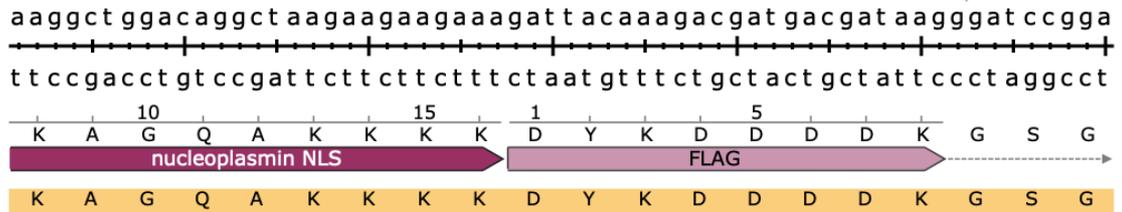
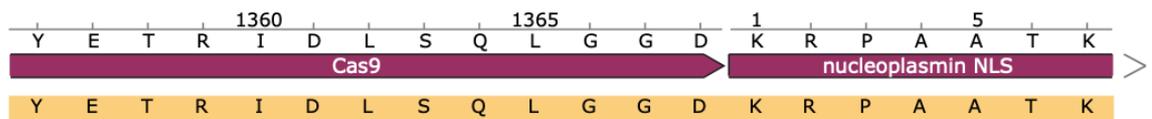
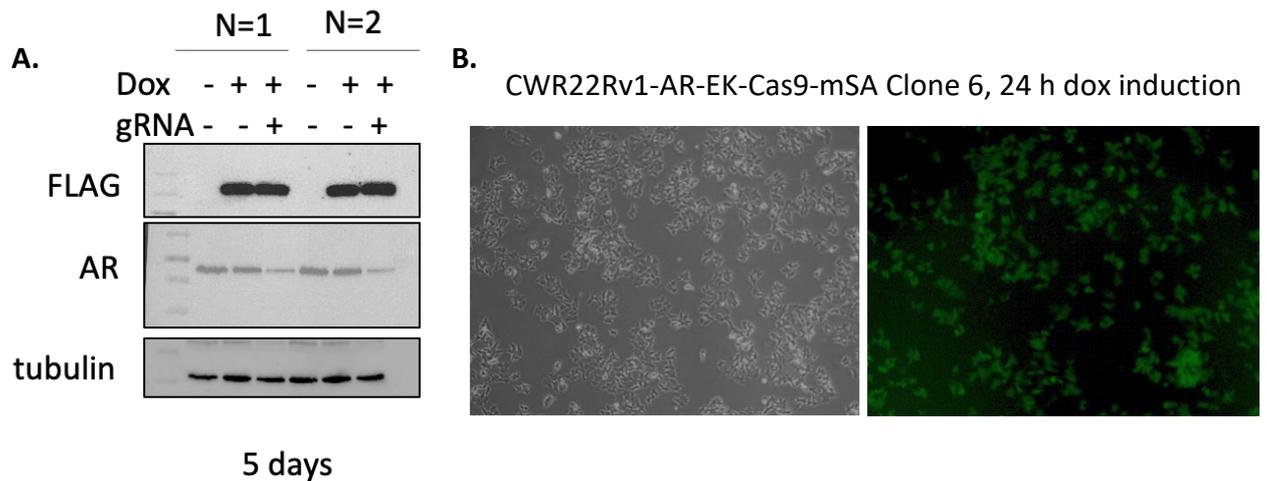
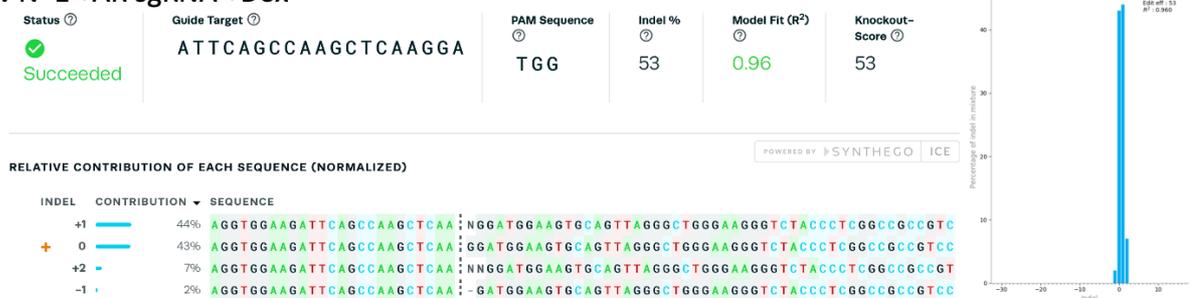


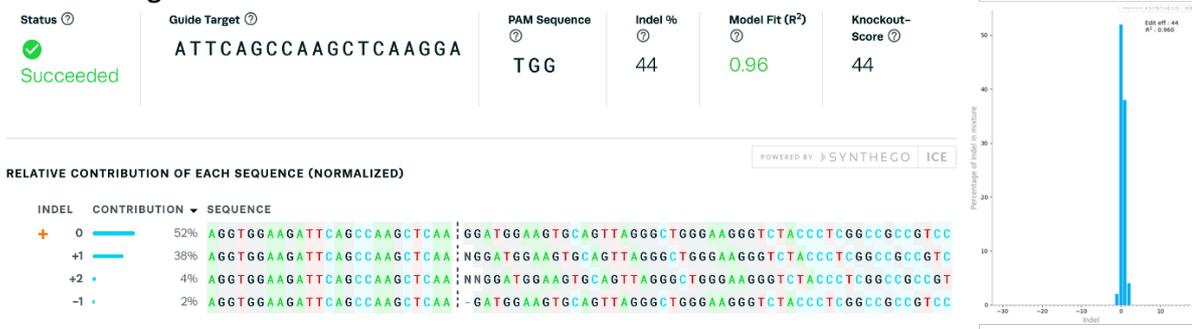
Figure 6.6 TLCV2 sequence showing location of mSA insertion and sequencing showing successful incorporation of mSA sequence.



**C. N=1 +AR sgRNA +Dox**



**N=2 +AR sgRNA +Dox**

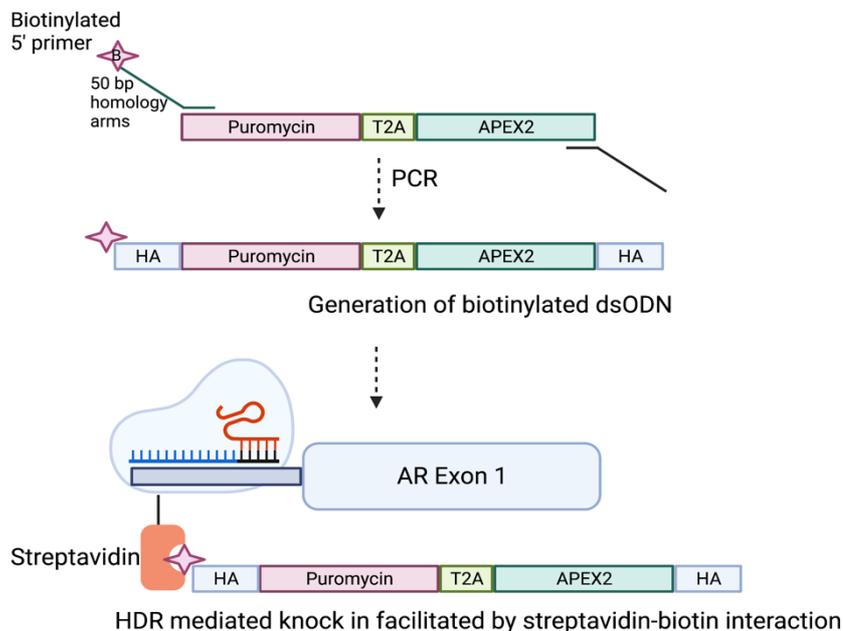


**Figure 6.7 Validation of a CWR22Rv1-AR-EK-iCas9-mSA inducible cell line**

CWR22Rv1-AR-EK-iCas9-mSA cells were treated with doxycycline for 24 hours to induce Cas9-mSA expression. 25 nM sgRNA was then transfected for 96 hours before western analyses (A) to assess AR protein levels. B. Microscopy images shows GFP (a surrogate for Cas9-mSA) expression 24 hours after doxycycline induction (10x magnification). C. 72 hours post sgRNA transfection, cells were harvested for genomic DNA extraction and the region encompassing the sgAR target site was amplified using PCR. Indel Synthego ICE analysis shows the editing efficiency of 2 replicates.

### 6.3.2 Knock in cell line attempts

Initial attempts of the APEX2-AR-V knock-in utilised transient expression of Cas9-mSA, whilst waiting for expansion and validation of clonal populations of the CWR22Rv1-AR-EK-iCas9-mSA inducible cell line. To create the donor template to enable knock-in of the APEX cDNA sequence adjacent to the translational start site of the *AR* gene, a puromycin-T2A-APEX2 gBlock™ (Integrated DNA Technologies) was amplified using PCR with primers that contained 50 bp *AR* gene homology arms that were biotin modified to ensure PCR amplicons were biotinylated and would have an affinity for the streptavidin-tagged Cas9 (Figure 6.8). This generated a Puro-T2A-APEX2 double-stranded donor template (dsODN) that could be incorporated into the downstream knock-in pipeline.

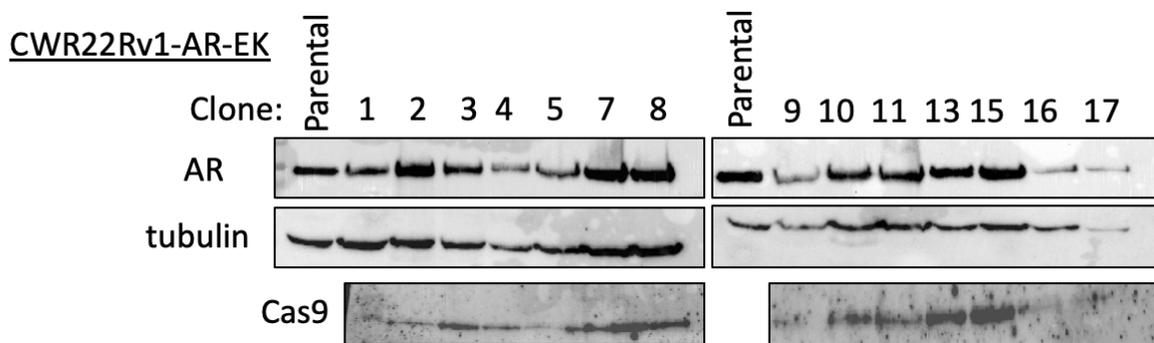


**Figure 6.8 Diagrammatic representation of APEX2 knock-in pipeline**

An APEX2-puromycin cassette gBlock is amplified with 5' biotinylated primers containing complementary sequences (homology arms) to the *AR* gene to generate a donor template. This biotinylated donor interacts with the streptavidin moiety of a Cas9-mSA fusion protein to facilitate APEX2 cDNA knock-in adjacent to the ATG of the *AR* gene.

1  $\mu\text{g}$  pTLCV2-Cas9-mSA was transiently transfected into  $5 \times 10^5$  CWR22Rv1-AR-EK cells alongside 2.5  $\mu\text{g}$  Puro-T2A-APEX2 dsODN and 25 nM on-target sgAR. A second transfection was performed 48 hours later to maximise plasmid take-up and protein expression. Cells were transferred into 10 cm dishes upon reaching 80-90% confluency before treatment with 1  $\mu\text{g}/\text{mL}$  puromycin after 10 days to allow time for expression of Cas9-mSA, and puromycin

resistance from the plasmid to diminish. Plates were then left to incubate until colonies of cells were large enough to be picked and placed into single wells, to enable screening of single clonal populations of cells. Clones were then screened using western blot to determine if there was a size change in the AR protein (expected +27 kDa). No successful knock-in clones were detected. It is likely that puromycin resistance has arisen via stable integration of the pTLCV2-Cas9 plasmid, that itself encodes puromycin resistance, rather than successful knock-in of the puromycin cassette from the donor template. This was shown using western blot of Cas9 in the surviving colonies (Figure 6.9) where a faint band representing Cas9 was detected in all clones.

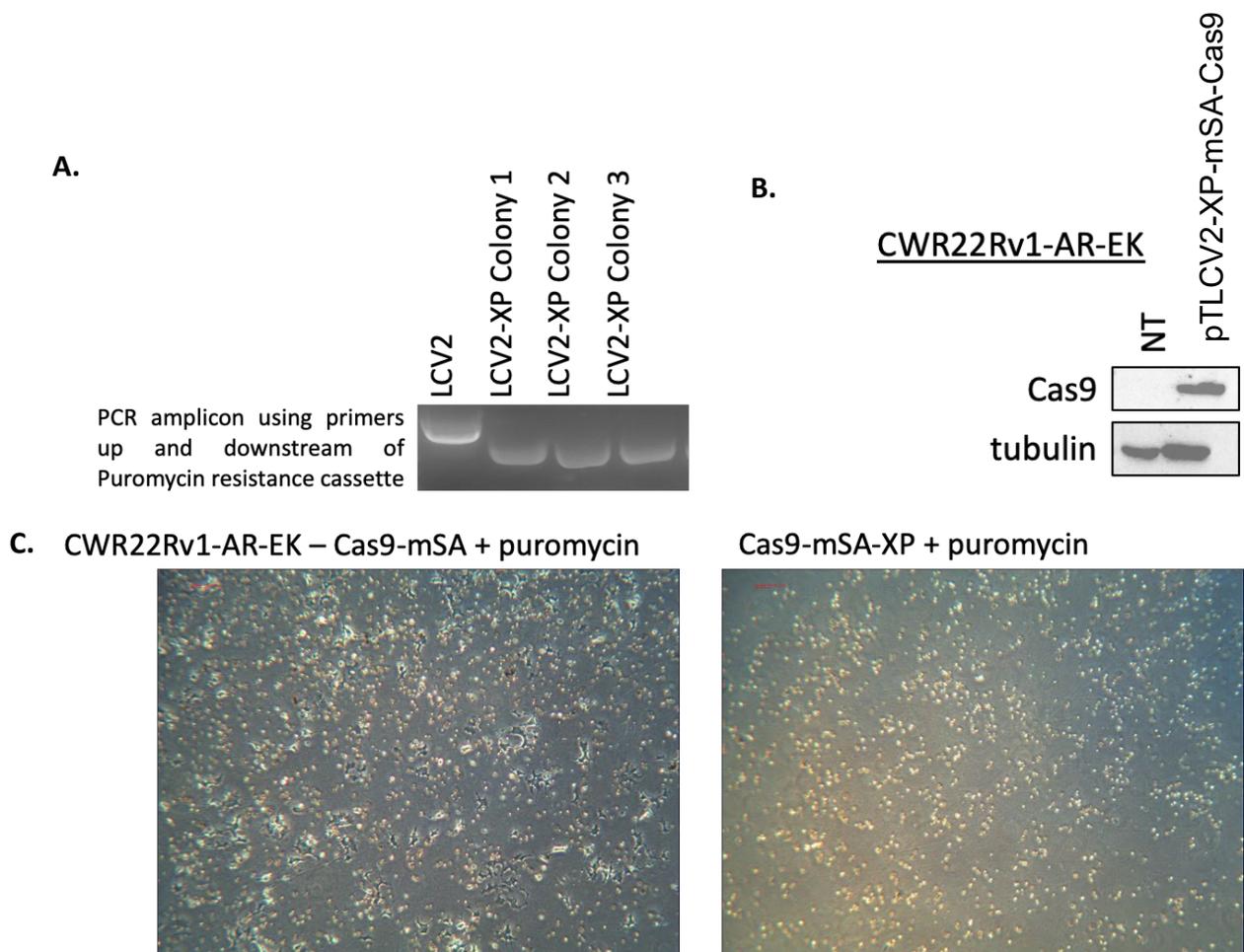


**Figure 6.9 Screening of CWR22Rv1-AR-EK cells after puromycin selection**

CWR22Rv1-AR-EK cells were double transfected with 1 µg pTLCV2-mSA-Cas9, 2.5 µg dsODN template and 25 nM sgAR and puromycin selected after 12 days. Surviving clones were screened using western blot analysis of AR protein compared to a parental CWR22Rv1-AR-EK cell lysate.

To mitigate the problem of cells surviving due to stable integration of the pTLCV2-mSA-Cas9 plasmid, and not because of knock-in of the puromycin HDR template, the puromycin sequence was digested out of the pTLCV2-mSA-Cas9 plasmid using restriction enzymes *Stu* I and *Bsi* WI (New England Biolabs) for 4 hours followed by T4 polymerase (New England Biolabs) treatment to fill in the overhang from the *Bsi* WI digest. The plasmid, now termed pTLCV2-XP-mSA-Cas9, was then ligated with T4 ligase (New England Biolabs) for 2 hours at room temperature then overnight at 4 °C before transformation into *Stb3* cells on ampicillin-containing plates. Colony PCR was used to screen colonies and a successful colony was minipreped and sent for Sanger sequencing (Figure 6.10A). pTLCV2-XP-mSA-Cas9 was transfected into CWR22Rv1-AR-EK cells and Cas9 expression was detected using western blot to confirm removal of the puromycin sequence did not interfere with Cas9 protein expression

(Figure 6.10B). Cells transfected with either pTLCV2-mSA-Cas9 or pTLCV2-XP-mSA-Cas9 were treated with puromycin to validate that LCV2-XP does not provide any resistance to puromycin. Cells were visualised under a microscope after two weeks selection, and this showed the cells transfected with the parental pTLCV2-mSA-Cas9 plasmid enabled cells to survive under selection with puromycin. In contrast, cells transfected pTLCV2-XP-mSA-Cas9 did not survive, so it was concluded that this plasmid will not confer resistance to puromycin, and that any cells that survive after transfection with the Puro-T2A-APEX2 dsODN and puromycin selection, survive because of a successful knock-in event and not stable integration of the pTLCV2-mSA-Cas9 plasmid (Figure 6.10C).



**Figure 6.10 Removal of puromycin resistance cassette from LCV2 plasmid.**

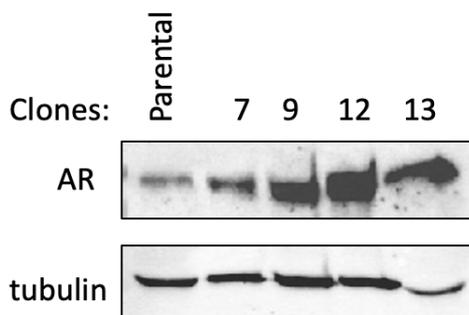
A. Colony PCR of colonies after transformation to detect size decrease in amplicon using primers up and downstream of the puromycin-resistance gene. B. CWR22Rv1-AR-EK cells were reverse transfected with pTLCV2-XP-mSA-Cas9 plasmid to confirm Cas9 expression was not altered in response to removal of the puromycin-resistance gene. C. CWR22Rv1-AR-EK cells transfected with either pTLCV2-mSA-Cas9 or pTLCV2-XP-mSA-Cas9 were treated with puromycin for two weeks to validate removing the puromycin-resistance gene from the pTLCV2-mSA-Cas9 plasmid sensitises cells to

puromycin selection. Images taken using an Olympus CK40 with a Viviscam 5.0 camera (10x magnification).

Having established the newly modified pTLCV2-XP-mSA-Cas9 plasmid does not confer puromycin resistance, CWR22Rv1-AR-EK cells were seeded at a density of  $5 \times 10^5$  in a well of a 6-well plate and were transfected with 1  $\mu\text{g}$  pTLCV2-XP-mSA-Cas9, 2.5  $\mu\text{g}$  dsODN template and 25 nM sgAR. Cells were then transfected 48 hours later with the same reagents after which they were transferred to a 10 cm dishes upon reaching 80-90% confluency. Puromycin was added 24 hours later to begin selection of the potential knock-in clones. Importantly, the control population that were transfected with the pTLCV2-XP-mSA-Cas9 plasmid alone without the donor template all died in response to puromycin. Puromycin-resistant cells from the pTLCV2-XP-mSA-Cas9/dsODN-transfected population were then allowed to form colonies that were later picked, bulked up and screened for APEX2-AR-V expression using western blotting. Unfortunately, none of the clones that successfully grew in the presence of puromycin had the 27 kDa size increase in the AR-V protein band (Figure 6.11A). Clone 13 appeared to have a slight upshift in AR-Vs, so was tested by using the biotinylation protocol and lysates were then ran on a western blot to assess biotin-labelling ability. However, there was no biotinylation seen in the streptavidin enriched or input samples. Consequently, this attempt at APEX2 knock-in was unsuccessful (Figure 6.11B).

**A.**

CWR22Rv1-AR-EK mSA-XP



**B.**

CWR22Rv1-AR-EK mSA-XP Clone 13

	<u>Enriched</u>		<u>Inputs</u>	
Biotin-phenol	-	+	-	+

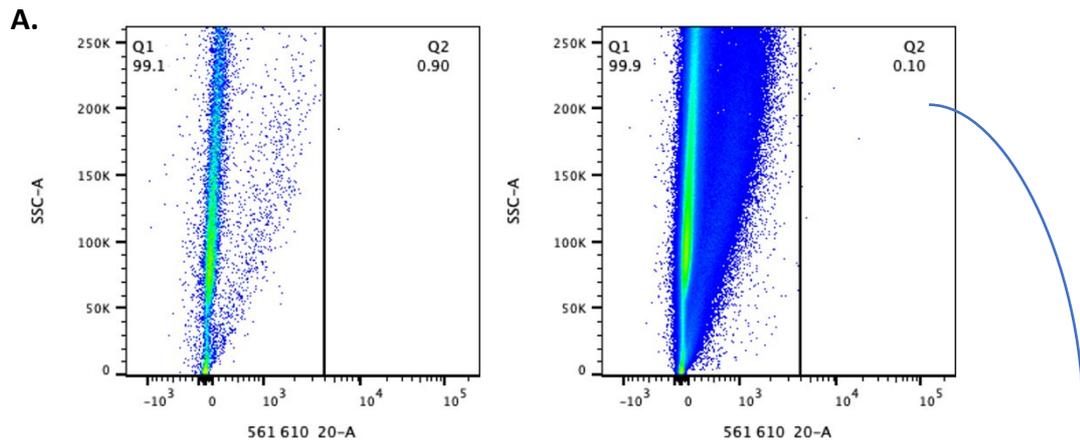


**Figure 6.11 APEX2-AR-V knock-in screening of puromycin-selected clones and biotinylation test of single clone**

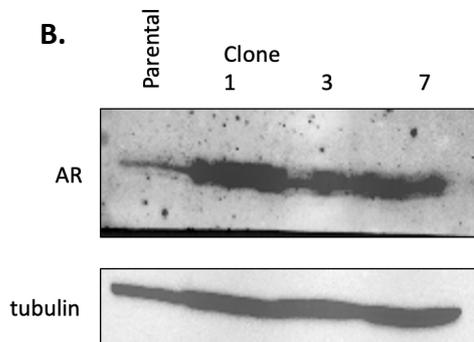
A. CWR22Rv1-AR-EK cells were double transfected with 1  $\mu\text{g}$  pTLCV2-XP-mSA-Cas9, 2.5  $\mu\text{g}$  dsODN template and 25 nM sgAR and puromycin selected. Surviving clones were screened using western blot analysis of AR to detect successful APEX2 integration by band supershift. B. Clone 13 was seeded and treated with biotin-phenol then  $\text{H}_2\text{O}_2$  to induce the biotinylation reaction. Input lysates and streptavidin bead-enriched lysates were analysed using western blot to detect biotinylated proteins using a biotin antibody.

As unsuccessful APEX2 knock-in clones were surviving puromycin selection, an RFP selection approach was next attempted. This would allow for cell sorting of RFP positive cells, and the cells that are RFP positive should undoubtedly be expressing RFP and thus by proxy, the APEX2 AR-V fusion. A new dsODN gBlock™ was designed to encode RFP-T2A-APEX2 in place of the puromycin-resistance gene and this was PCR amplified in-house to add 50 bp homology arms that are biotinylated at the 5' end to enhance efficiency of knock-in. CWR22Rv1-AR-EK cells were seeded into 60 mm dishes at a cell density of  $1 \times 10^6$  and were reverse transfected with 2  $\mu\text{g}$  pTLCV2-mSA-Cas9, 2  $\mu\text{g}$  RFP-T2A-APEX2 and 25 nM sgAR. A control population were transfected with no guide RNA so when selected, cells can be gated to not include this background of cells. Cells were re-transfected with 2  $\mu\text{g}$  pTLCV2-mSA-Cas9, 2  $\mu\text{g}$  RFP-T2A-APEX2 and 25 nM sgAR 48 hours later. Doxycycline was added post transfection to induce Cas9-mSA expression. When cells were 80-90% confluent they were moved up into a 10 cm dish then 3 x T175 culture flasks prior to sorting. Prior to cell sorting, cells were assessed for RFP using a Nikon2000 microscope. However, no RFP could be detected. This could be because if RFP-APEX2 has been inserted into the endogenous AR gene, the expression level may not be high enough for detection. For cell sorting, cells were washed with PBS, trypsinised and counted. ~30 million cells were resuspended in 2 mL FACS buffer (PBS, 1% v/v FBS, 6 mM EDTA) and filtered to remove cell clumps. 13.5 million control cells were used to set the gates for the positive RFP cell population. Around 1000 cells (Q2 in flow cytometry plot in Figure 6.12A) were selected using the FACS Aria cell sorter, which were then seeded into a well of a 6-well plate. It is worth noting there was not a distinct 'RFP-positive' cell population detected by the cell sorter and cells that were being sorted could be debris or auto-fluorescing cells.

The sorted cells were then cultured until they formed colonies that could be picked and bulked up. Clones were then screened using western blot to detect the 27 kDa increase in the AR-V protein (Figure 6.12B). However, no APEX2-AR-V knock-in clones were detected so it was concluded that this approach was not successful. CWR22Rv1-AR-EK-iCas9-mSA cells were also used to attempt the knock-in using the same approach as above. Instead of cell sorting, the FACS Attune machine was used to screen a sub-population of cells to determine if any RFP could be detected in the cells. However, there was no detectable difference between the control cell population and the 'knock-in' cell population, so it was concluded this approach was also unsuccessful.



Name	Statistic	#Cells
161372.fcs		20000
Q1: 561 610_20-A-, SSC-A+	99.1	19820
Q2: 561 610_20-A+, SSC-A+	0.90	180
Q3: 561 610_20-A+, SSC-A-	0	0
Q4: 561 610_20-A-, SSC-A-	0	0
161374.fcs		102000
Q1: 561 610_20-A-, SSC-A+	99.9	1018959
Q2: 561 610_20-A+, SSC-A+	0.10	1041
Q3: 561 610_20-A+, SSC-A-	0	0
Q4: 561 610_20-A-, SSC-A-	9.80E-5	1



**Figure 6.12 Cell sorting of RFP positive clones and western blot screening of sorted clones**

A. 30 million CWR22Rv1-AR-EK cells transfected with pTLCV2-mSA-Cas9, RFP-T2A-APEX2 dsODN and sgAR were harvested into FACS sorting buffer and sorted for RFP positivity using Aria cell sorter, using the control CWR22Rv1-AR-EK cells as a control population to set gates. B. Clones were picked and bulked up for western blot screening of AR protein size change.

Because the CRISPR/Cas9 knock-in technique to generate a stable endogenously-tagged APEX2-AR-V expressing cell line was unsuccessful, and because of time restraints, ectopic expression of APEX2-AR-V7 was used to determine the AR-V7 interactome in steady-state and post-irradiated cells. Although not as physiologically-relevant as an endogenous gene-tagging

approach, I felt that this contingency would still offer an exciting insight into AR-V7 function in steady-state and in response to DNA damage. Furthermore, there are published FL-AR and AR-V interactomes that have used ectopically expressed FL-AR and AR-Vs, so this technique is fitting with what is published and will enable comparison between datasets.

### 6.3.3 Generation of an APEX2-V7 plasmid

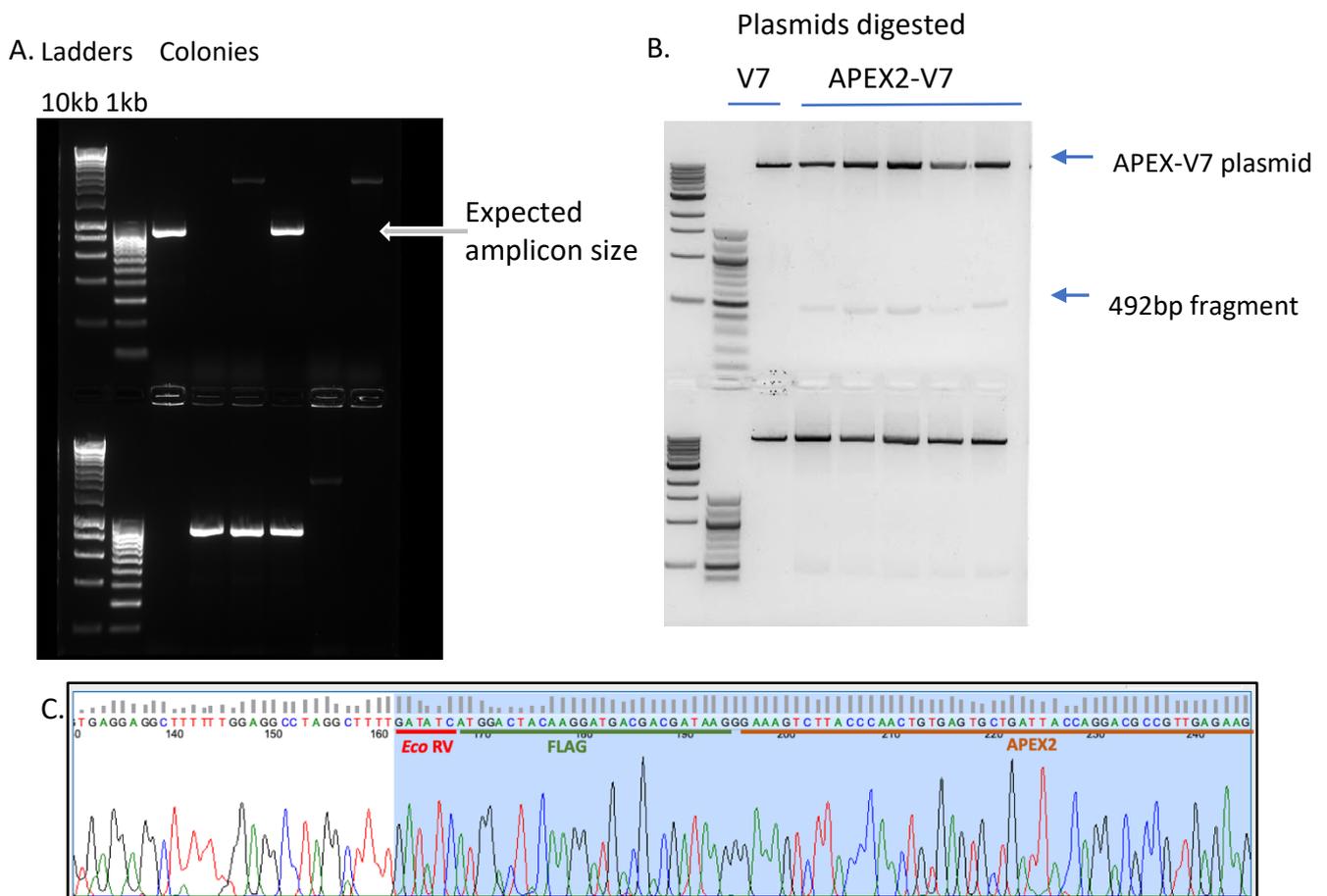
To determine the interactome of AR-V7 an N-terminal APEX2 fusion was generated using a pLV-AR-V7 plasmid. The donor pLV-AR-V7 plasmid was digested adjacent to the ATG start codon of AR-V7 using *Eco* RV (NewEngland Biolabs) according to the manufacturer's instructions for a minimum of 4 hours at 37 °C and then treated with calf intestinal phosphatase (Invitrogen) to prevent re-ligation of the plasmid. This product was then run on a 1% agarose gel to confirm digestion. This was then gel extracted and column purified using a gel extraction kit (NewEngland Biolabs). The APEX2 insert was amplified from an pEJS578\_DD-dSpyCas9-mCherry-APEX2 plasmid (Addgene #108570) using the primers in Table 6.3. The FLAG sequence was included in the forward primer to generate a Flag-tagged APEX2-V7 fusion. The addition of ATC at the start of the APEX2 forward primer retains the *Eco* RV restriction site to enable successful ligation and to ensure the APEX2 sequence was inserted in the correct orientation.

**Table 6.3: Primers used for the amplification of the APEX2 sequence to be inserted into the AR-V7 donor plasmid**

Forward (5'-3'): <u>EcoRV-FLAG-APEX2</u>	<u>ATCATGGACTACAAGGATGACGACGATAAGGGAAAGTCTTACCCAACTGTG</u>
Reverse (5'-3'):	GCATCAGCAAACCCAAGCTCG
AR-V7 Reverse (5'-3'):	GCTCTGGAACAGATTCTGGAA

The PCR fragment was then gel extracted and polynucleotide kinase treated for 30 minutes at 37 °C. The insert and digested plasmid were then ligated, using the suggested ratios of plasmid

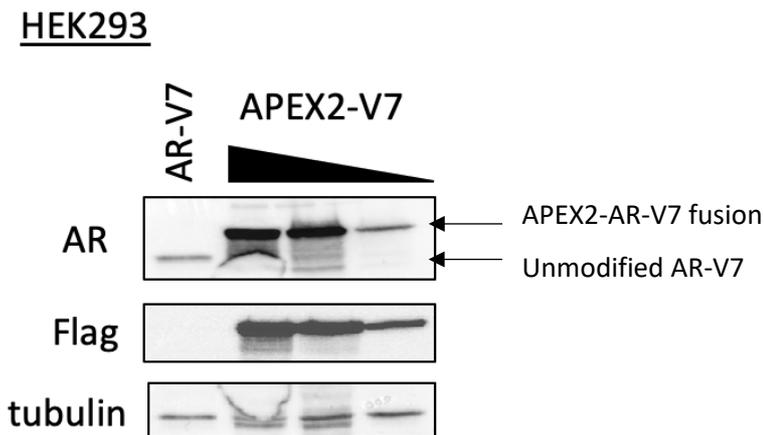
and insert, as predicted by the NEB calculator (1:5 and 1:3), using T4 ligase overnight at 16 °C. Ligation reactions were then transformed into *DH5α* competent cells and resultant colonies screened using colony PCR to identify successful recombinant vectors. Resultant PCR products were subject to electrophoresis to analyse amplicon size (Figure 6.13A). Successful plasmids were then further validated by diagnostic digest incorporating *EcoRV* which cleaves the plasmid twice (once at the restriction site maintained at the 5' end of APEX2 and one in the middle of the APEX2 sequence) to release a 492 bp fragment (Figure 6.13B). 5 successful colonies that released the 492 bp fragment were sent for Sanger sequencing by GeneWiz using an AR reverse primer (Table 6.3) to confirm insertion of the APEX2 fragment (Figure 6.13C).



**Figure 6.13 An APEX2-AR-V7 encoding plasmid was successfully cloned.**

A. Bacterial cell colonies post-transformation were subject to PCR using an APEX2 forward and AR-V7 reverse primer set to confirm the insertion of the APEX2 sequence. The presence of a band at 1 kb indicates successful cloning into the pLV-AR-V7 plasmid with the APEX2 sequence in the correct orientation. B. Diagnostic digest of AR-V7 and APEX2-V7 mini-prepped plasmid with *EcoRV*. Sequencing results confirmed insertion of the FLAG-APEX2 sequence into the pLV-AR-V7 plasmid.

The resultant pLV-FLAG-APEX2-AR-V7 construct or the original pLV-AR-V7 plasmid, was transfected into HEK293 cells for 72 hours prior to western analysis incorporating anti-AR and FLAG antibodies. As shown in Figure 6.14, the expected size increase (+27 kDa) of the FLAG-APEX2-AR-V7 fusion protein was detected in cells transfected with pLV-FLAG-APEX2-AR-V7, but not in control cells transfected with the unmodified pLV-AR-V7 construct.



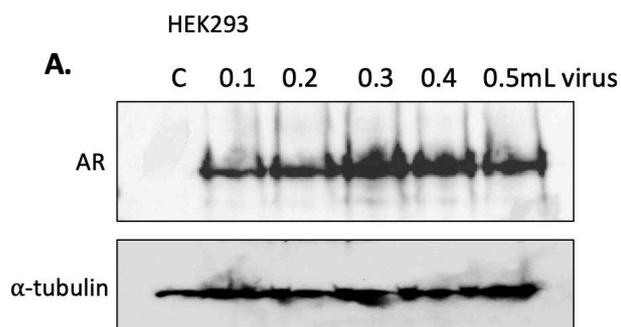
**Figure 6.14 Successful expression of APEX2-AR-V7 fusion protein in HEK293 cells after transfection with APEX2-AR-V7 plasmid.**

HEK293 cells were reverse transfected with either pLV-AR-V7 plasmid, or varying amounts of the pLV-FLAG-APEX2-AR-V7 plasmid for 72 hours. Cells were then harvested and subject to western blot to assess expression of AR-V7 or APEX2-AR-V7 with an AR antibody, and APEX2-AR-V7 expression with a Flag antibody.

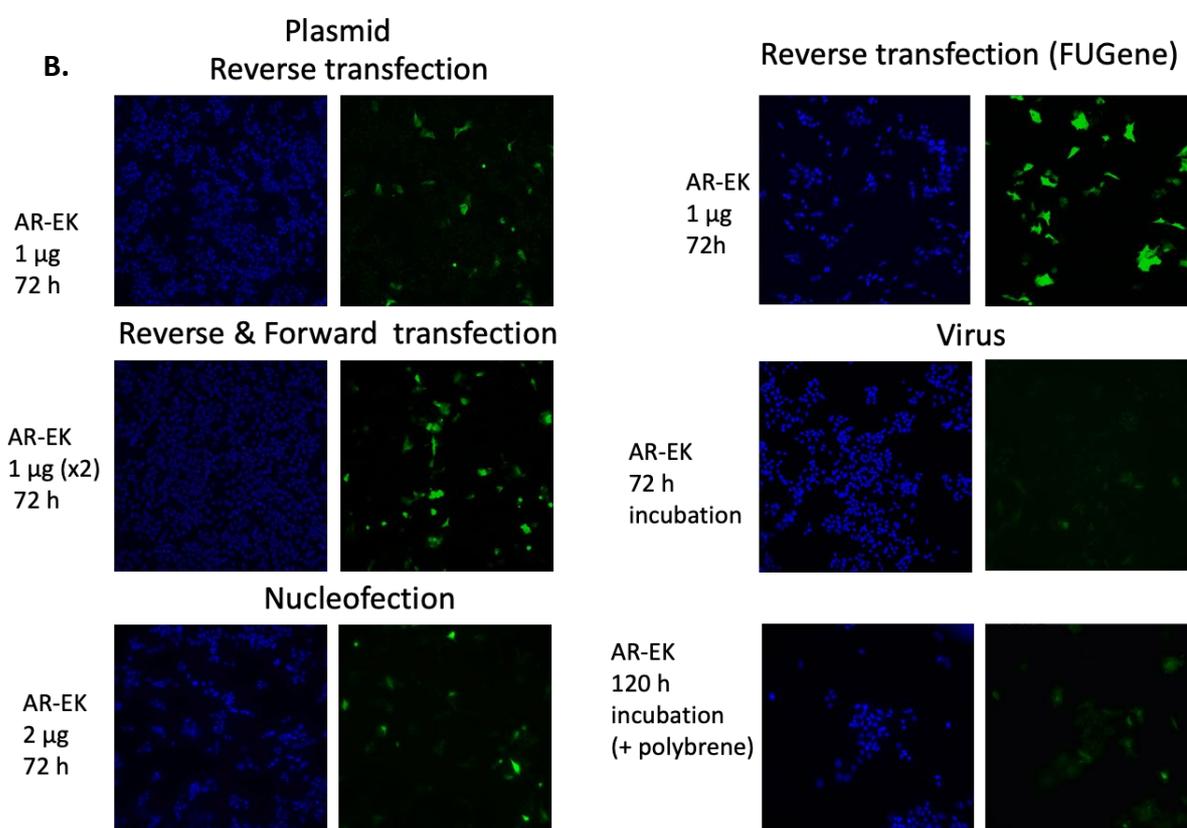
#### 6.3.4 The APEX2-V7 fusion protein is functional and is recruited to androgen response elements on chromatin

A suitable transfection efficiency of the APEX2-AR-V7 plasmid was essential for the proximity labelling technique to enable labelling and enrichment of low abundant proteins for subsequent MS analysis. Furthermore, because it is unknown how much biotinylated protein is being enriched using streptavidin beads, a high transfection efficiency needs to be achieved to increase the chances of detecting and identifying proteins. As the CWR22Rv1-AR-EK cell line is a difficult to transfect cell-line, this required optimisation. Several methods of transfection of the APEX2-AR-V7 plasmid were attempted, including transient transfection using 2 different lipofectamine reagents, LT1 and FuGene® (Promega); as well as nucleofection and lentiviral-based transduction. APEX2-AR-V7 lentivirus was generated and validated for

APEX2-AR-V7 expression by transduction in HEK293 cells (Figure 6.15A). Viral transduction of the APEX2-V7 plasmid was attempted in CWR22Rv1-AR-EK cells as the levels of APEX2-AR-V7 protein may be closer to endogenous levels of AR-V7 protein. However due to the lack of a selection marker in the plasmid, APEX2-AR-V7 expression was not particularly robust in this cell line (Figure 6.15B). As FuGene<sup>®</sup> transfection reagent is optimised for difficult to transfect cells, FuGene was utilised for transfection. In addition, our conventional transfection reagent LT-1 was used, and efficiency of the procedures were assessed using immunofluorescence, incorporating an anti-FLAG antibody, to detect ectopically-expressed FLAG-tagged APEX2-AR-V7 fusions. As shown in Figure 6.15B, a double transfection of APEX2-AR-V7, in which cells are reverse transfected on the day of seeding then again 48-72 hours later, yielded a greater transfection efficiency using LT-1 when compared against CWR22Rv1-AR-EK cells that had been transfected once with FuGene. Therefore this 'double transfection' using LT1 was taken forward to test and optimise the labelling and enrichment protocol.



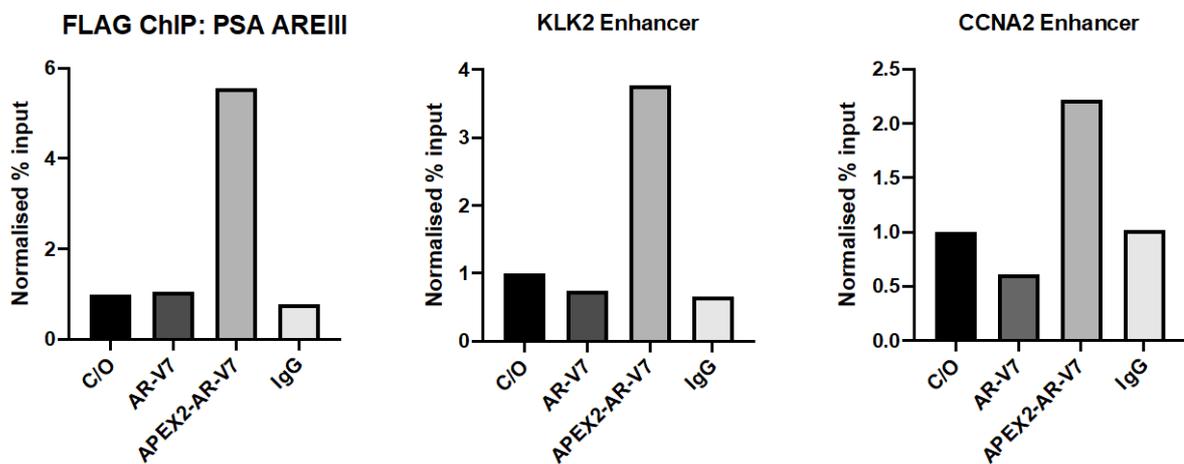
### Optimisation of the transfection efficiency in CWR22Rv1-AR-EK cells



**Figure 6.15 Optimising the transfection efficiency of the APEX2-AR-V7 plasmid**

A. HEK293 cells were transduced with increasing volumes of APEX2-AR-V7 lentivirus to confirm APEX2-AR-V7 expression after 72 hours. B. CWR22Rv1-AR-EK cells were transfected with LT1 transfection reagent (unless stated with FUGene) or subject to nucleofection with APEX2-AR-V7 plasmid or transduced with APEX2-AR-V7 lentivirus and incubated for 72 or 120 hours (as indicated) and then subject to immunofluorescence analysis using a FLAG antibody to assess transfection/transduction efficiency (20x magnification).

Before the proximity biotinylation pipeline could commence, the APEX2-AR-V7 fusion was tested for functionality as there was a possibility that the addition of the APEX2 enzyme would impact the distribution and chromatin binding capacity of the AR-V7 protein. ChIP experiments using either anti-FLAG or isotype control antibodies were performed in CWR22Rv1-AR-EK cells transfected with the AR-V7- or FLAG-APEX2-AR-V7-encoding plasmids, or mock-transfected cells (C/O). As shown in Figure 6.16, the FLAG-APEX2-AR-V7 fusion protein is enriched at *cis*-regulatory elements of canonical AR-V target genes *PSA*, *KLK2* and *CCNA2* so it was assumed that the APEX2 moiety of the fusion does not impact AR-V7 functionality. In hindsight, a more appropriate control for this experiment would have been to directly compare FLAG-tagged AR-V7 chromatin enrichment to that of the FLAG-APEX2-AR-V7, but due to time-constraints, we were unable to generate the FLAG-AR-V7 construct in time to complete these experiments. Critically, the enrichment of the ectopic AR-V7 fusion protein is consistent with the percentage inputs observed for endogenous AR-Vs in CWR22Rv1 and CWR22Rv1-AR-EK cells (Kounatidou *et al.*, 2019).



**Figure 6.16 FLAG antibody-based ChIP shows APEX2-AR-V7 fusion protein is recruited to androgen response elements on chromatin**

CWR22Rv1-AR-EK cells were seeded, and transfected with 5  $\mu$ g pLV-AR-V7 or pLV-FLAG-APEX2-AR-V7 plasmid for 72 hours, or mock for control (C/O), before ChIP incorporating anti-FLAG or isotype control

antibodies. Data shows normalised percentage input to the C/O control sample of FLAG-AR-V7 enrichment at enhancers of *PSA* (ARE III), *KLK2* and *CCNA* and is n=1.

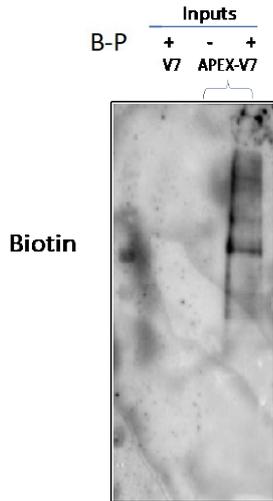
#### 6.3.5 Optimisation of the labelling protocol in HEK293 and CWR22Rv1-AR-EK cells

Now that the APEX2-AR-V7 fusion had been successfully generated, it was critical that the optimisation of the biotin labelling and enrichment protocol was performed. HEK293 cells were chosen for this due to their relative ease of transfection. Cells were transfected with 5  $\mu$ g pLV-FLAG-APEX2-AR-V7 in 10 cm dishes for 72 hours prior to treatment with 500  $\mu$ M biotin-phenol for 1 hour and then 1 mM H<sub>2</sub>O<sub>2</sub> for 30 seconds (refer to Figure 6.2). Cells were then washed twice with quenching solution to stop the labelling reaction and then scraped into PBS and pelleted. Cells were then lysed in RIPA buffer, cleared by centrifugation and then 30  $\mu$ g streptavidin coated DynaBeads was added to the whole cell lysates and incubated at 4 °C overnight. Protein bound streptavidin beads were then subjected to multiple washing steps using RIPA buffer and eluted at 100 °C in SDS supplemented with 20 mM DTT and 2 mM biotin. Inputs were taken from whole cell lysates prior to enrichment. Input and enriched lysates were then subject to western blotting to assess the level of biotinylation using a biotin antibody (Figure 6.17A). Importantly, the western blot showed that proximal protein biotinylation only occurs in the presence of both the APEX2-AR-V7 fusion and the labelling reagents as indicated by a smear of biotinylated proteins in the arm containing the APEX2-AR-V7 fusion and biotin. In contrast, cells expressing APEX2-AR-V7, but not incubated with biotin-phenol or H<sub>2</sub>O<sub>2</sub>, did not demonstrate the smear of biotinylated proteins. Further validation of biotinylation selectivity was conducted by transfecting HEK293 cells with the original pLV-AR-V7 plasmid, lacking the APEX2 fusion, and inducing labelling. As expected, no protein biotinylation was detected which confirmed that biotinylation will only occur in the presence of APEX2-AR-V7 and the labelling reagents (Figure 6.17A).

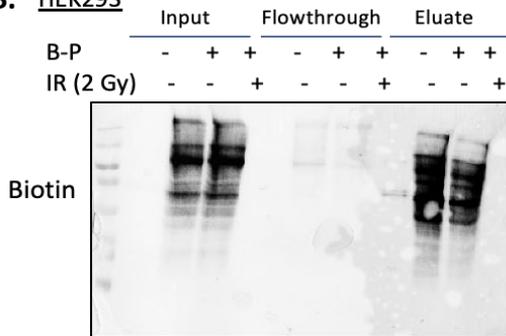
Intriguingly, a recent paper using proximity-based labelling with APEX2-fusion proteins to identify interactomes of DDR-associated proteins, indicated that irradiating cells reduced the number of proteins that were being biotinylated by APEX2 (Gupta *et al.*, 2018). Therefore, it was vital to determine if irradiating cells 2 hours pre-labelling impacted the abundance and magnitude of APEX2-mediated protein biotinylation. Importantly, In the Gupta et al. study, a

dose of 10 Gy irradiation was applied to cells. This is considerably higher than 2 Gy irradiation which was the chosen dose for our experiments as it had been previously demonstrated to be sufficient for inducing significant levels of DNA damage in CWR22Rv1 and CWR22Rv1-AR-EK cells as indicated by  $\gamma$ H2AX foci (Kounatidou *et al.*, 2019). Consequently, 2 and 4 Gy irradiation doses were tested to see if this reduced the number of biotinylated proteins detected by western blot. As shown in Figure 6.17B and C, neither dose of IR impacted the labelling efficiency as evidenced by comparable biotinylation smears in control, 2 and 4 Gy experimental arms. Therefore, the 4 Gy dose was taken forward for subsequent experiments to ensure adequate DNA damage occurs to induce a DDR. Also, as proteins of interest will be located in the nucleus, and as per the protocol in the publication investigating DDR proteins interactomes (Gupta *et al.*, 2018), longer biotin-phenol and H<sub>2</sub>O<sub>2</sub> incubations were adopted, using 2 hours instead of 1 hour pre-incubation with biotin-phenol to allow sufficient time for it to reach the nucleus. Similarly, a 2-minute H<sub>2</sub>O<sub>2</sub> incubation was used for the same reason, as this was also used in the publication by Gupta *et al.* This modified approach was tested in HEK293 and CWR22Rv1-AR-EK to determine if this adaption to the protocol impacted labelling efficiency; as assessed via western blot. These experiments indicated that the amount of biotinylated proteins in both the input and enriched arms was similar between plus and minus 4 Gy irradiation, and a longer incubation with biotin-phenol and H<sub>2</sub>O<sub>2</sub> did not decrease the labelling efficiency of APEX2-AR-V7 (Figure 6.17D-E).

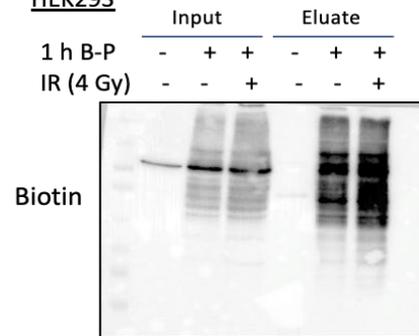
**A. HEK293**



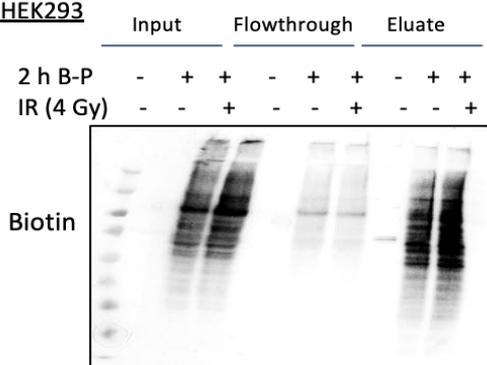
**B. HEK293**



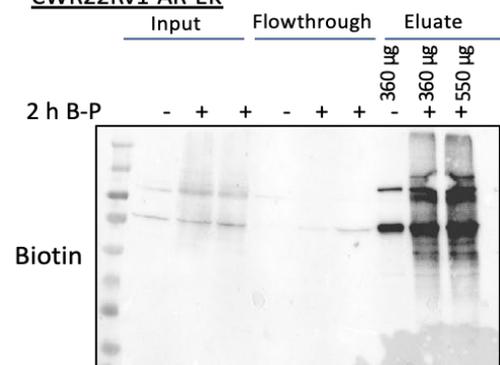
**C. HEK293**



**D. HEK293**



**E. CWR22Rv1-AR-EK**

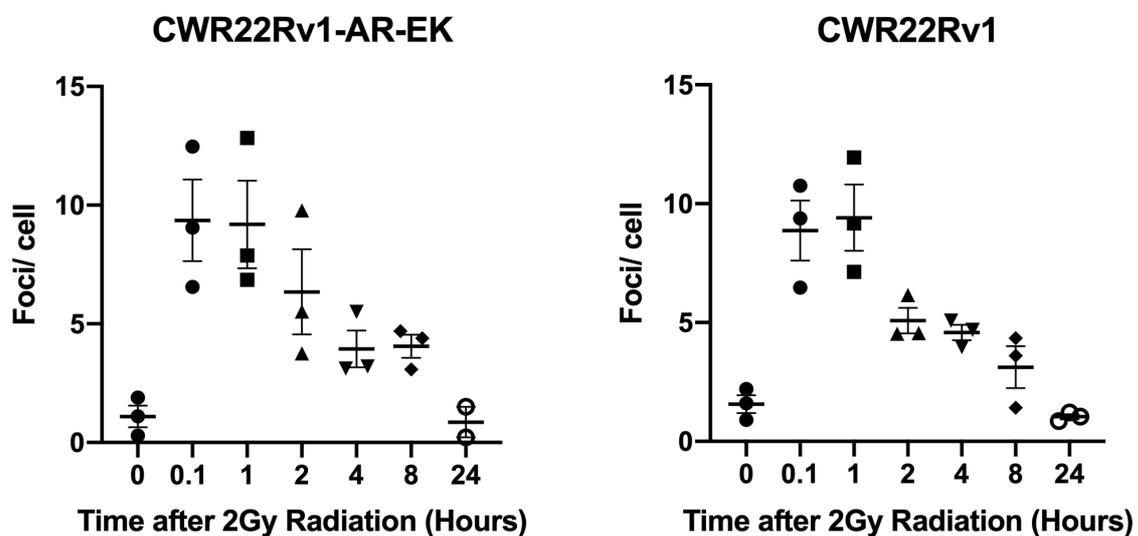


**Figure 6.17 Optimisation of the proximity biotinylation protocol in HEK293 and CWR22Rv1-AR-EK cells.**

A-D. HEK293 cells were reverse transfected with 5 µg AR-V7 or APEX2-AR-V7 encoding plasmids for 72 hours. Cells were then irradiated with 2 or 4 Gy (as stated) with the addition of biotin-phenol for 2 hours. H<sub>2</sub>O<sub>2</sub> was then added to culture media to induce the biotinylation reaction. Cells were then scraped into PBS, pelleted, lysed, enriched with streptavidin beads and biotinylated protein was eluted with SDS sample buffer supplemented with biotin. Lysates were then subject to western blot to analyse biotinylated protein using a biotin antibody. E. CWR22Rv1-AR-EK cells were reverse transfected with

5  $\mu\text{g}$  APEX2-AR-V7 plasmid. Cells were then scraped into PBS, pelleted, lysed, and quantified. 360 or 550  $\mu\text{g}$  lysate was then enriched with streptavidin beads and biotinylated protein was eluted with SDS supplemented with biotin. Lysates were then subject to western blot to analyse biotinylated protein.

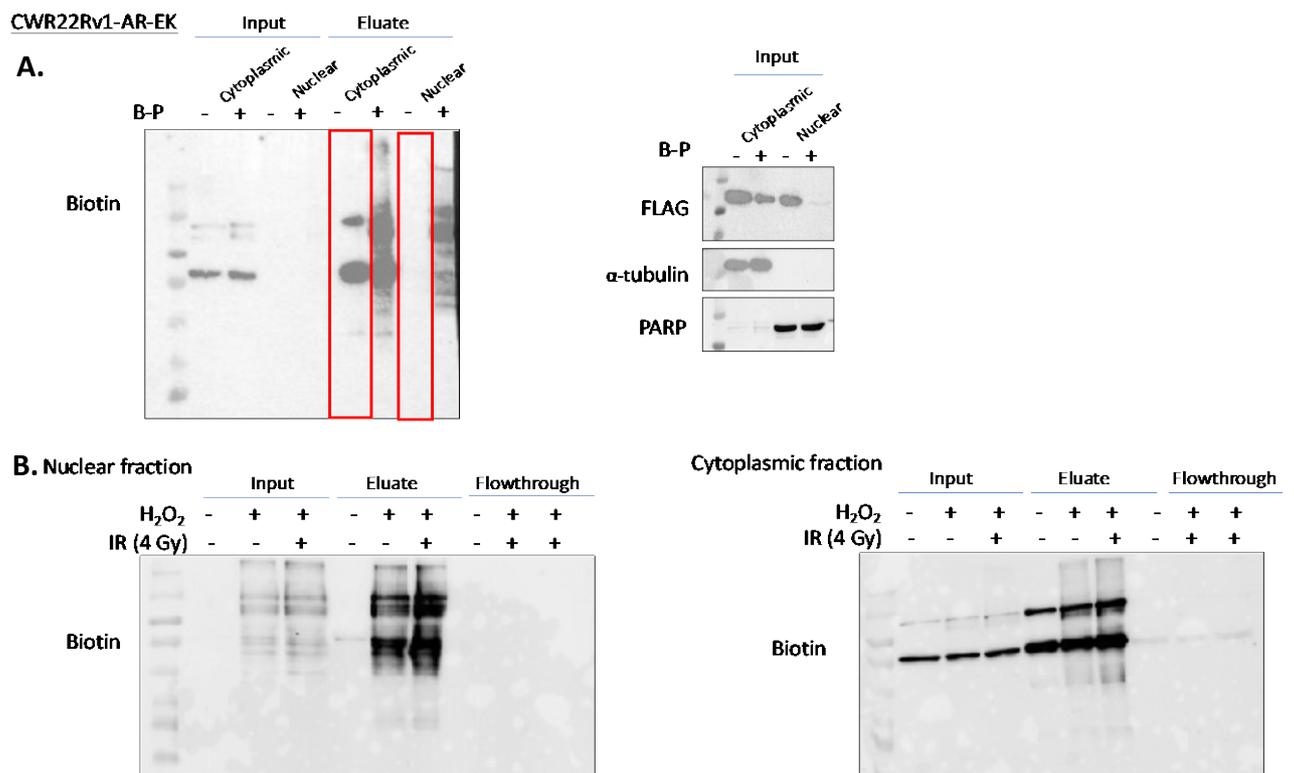
The DNA repair kinetics in CWR22Rv1-AR-EK and CWR22Rv1 cell were also investigated to determine a time point in which the labelling will be induced post-irradiation. This was done by quantifying  $\gamma\text{H2AX}$  foci using immunofluorescence, in cells that were fixed immediately after irradiation, then one, two, four, eight and twenty-four hours post 2 Gy irradiation. Upon DNA double stranded breaks, histone H2AX is rapidly phosphorylated ( $\gamma\text{H2AX}$ ) by ATM and/or DNA-PKcs, and is a common marker used to assess levels or DNA damage and the rate and efficiency of DNA repair (Rogakou *et al.*, 1999). This revealed that after two hours,  $\gamma\text{H2AX}$  foci per cell were decreased, meaning the foci have started to resolve as a consequence of successful DNA repair (Figure 6.18). Therefore, the two-hour time point was chosen for the labelling experiment as one of the aims of the experiment is to determine AR-V7 involvement in the DDR.



**Figure 6.18 Analysis of  $\gamma\text{H2AX}$  foci at several time points post 2 Gy irradiation**

CWR22Rv1 and CWR22Rv1-AR-EK cells were seeded onto glass coverslips for 24 hours. Cells were then irradiated with 2 Gy ionising radiation and harvested for immunofluorescence analysis of  $\gamma\text{H2AX}$  foci immediately after irradiation (0h) then 1, 2, 4, 8 and 24 hours post irradiation.  $\gamma\text{H2AX}$  foci quantification was performed using ImageJ software. Data represents three independent experiments and the mean  $\pm$  SEM is presented.

As shown in Figure 6.15, the APEX2-AR-V7 fusion is expressed in both the cytoplasm and the nucleus. Therefore, to eliminate any background labelling in the cytoplasm, and to maximise detection of nuclear proteins involved in AR-V7-mediated transcription and the DDR, a cytoplasmic nuclear extraction was performed immediately after the labelling protocol so the nuclear fraction could be isolated and further processed by streptavidin-based enrichment. The separation of the cytoplasmic and nuclear compartments also meant that the highly expressed endogenous biotinylated proteins observed in the cytoplasmic fraction (red box; Figure 6.19A), were absent in the nuclear sample and hence unable to interfere with or occupy streptavidin binding sites during enrichment that would ultimately impact protein identification by MS. Importantly, it was also shown in CWR22Rv1-AR-EK cells that the nuclear fractionated biotinylated protein was not affected by irradiation (Figure 6.19B).



**Figure 6.19 Cytoplasmic nuclear fractionation of CWR22Rv1-AR-EK cells enriches for nuclear biotinylated proteins and removes highly expressed endogenously biotinylated protein from eluates.**

A. CWR22Rv1-AR-EK cells were double transfected with pLV-FLAG-APEX2-AR-V7 for 72 hours before induction of the biotin labelling reaction. Post-labelling, cells were subject to cytoplasmic and nuclear fractionation. Resultant input, streptavidin-enriched and flowthrough samples were subject to western blot to analyse biotinylated proteins as well as FLAG, tubulin and PARP levels to assess

efficiency of cellular fractionation. B. CWR22Rv1-AR-EK cells treated as in (A) with the addition of 4 Gy ionising radiation treatment prior to activation of the biotin labelling reaction before sample analysis to detect efficiency of protein labelling.

Now that it has been established that (i) the CWR22Rv1-AR-EK transfection efficiency was sufficient to induce labelling of a detectable ‘fingerprint’ of biotinylated proteins when analysed by western blot; (ii) the APEX2-AR-V7 fusion is functional and is recruited to chromatin; (iii) IR does not impact labelling efficiency of APEX2-AR-V7; (iv) two hour biotin-phenol and two minutes H<sub>2</sub>O<sub>2</sub> incubation are appropriate timings to label nuclear proteins; and (v) cytoplasmic/nuclear fractionation eliminates highly abundant endogenously biotinylated proteins, the next step was to test if there was sufficient material for MS identification of labelled proteins. A schematic of the optimised protocol is presented in Figure 6.20.

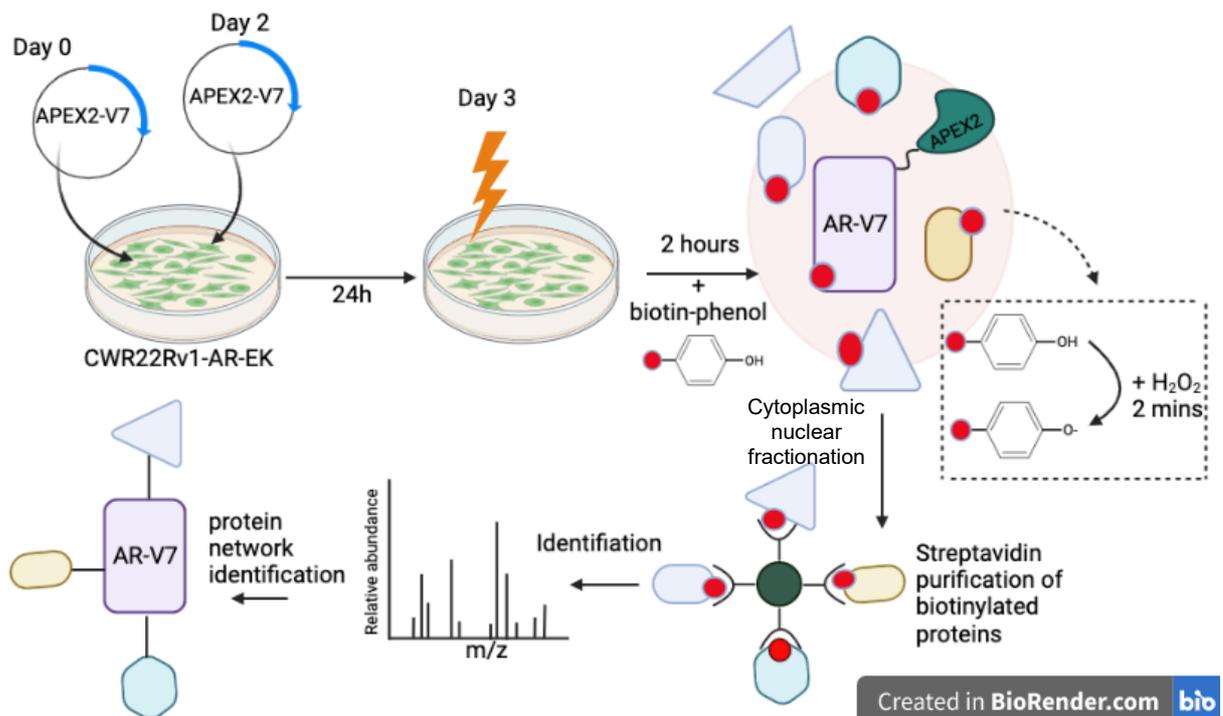


Figure 6.20 Optimised expression/ labelling protocol of AR-V7 interactome experiment.

### 6.3.6 The androgen receptor was identified as the most abundant biotinylated protein in CWR22Rv1-AR-EK cells by mass spectrometry

As it was still unknown whether the streptavidin-enriched protein samples were concentrated enough for identification, preliminary MS experiments were performed. CWR22Rv1-AR-EK cells were seeded into 10 cm dishes and transfected with 5 µg pLV-FLAG-APEX2-AR-V7 plasmid for 48 hours prior to a second transfection for a further 24 hours before conducting the labelling protocol. On the day of labelling, cells were irradiated with 4 Gy irradiation and concurrently incubated with biotin-phenol for 2 hours. H<sub>2</sub>O<sub>2</sub> was added to the plus and minus IR arms but not to the control arm. Subsequently, cytoplasmic and nuclear fractions were isolated, and 60 µg of the nuclear fraction was incubated overnight at 4 °C with 30 µL streptavidin beads. Beads were then washed 7 times with RIPA buffer then 7 times with PBS, to remove residual detergent, before being resuspended in 1 mL PBS and sent to the Newcastle University MS facility for on-bead protein digestion and analysis. This initial experiment was to determine if sufficient material was being enriched by the IP step to permit identification of proteins by MS. Table 6.4 shows that the AR is in the top two abundant proteins in the samples in which the labelling was enabled. This was expected due to APEX2 being fused directly to AR-V7, and therefore the most proximal protein to APEX2 is AR-V7 making it the most readily biotinylated during the reaction. The log(e) value indicates the expectation of finding the protein stochastically. Therefore, proteins that have a value of higher than -2 are removed from the final list. The number of proteins that were identified in each treatment arm were 21 in the control, 37 in the minus IR sample, and 38 in the plus IR sample.

**Table 6.4 Top two proteins identified from first CWR22Rv1-AR-EK labelling experiment in each experimental arm.**

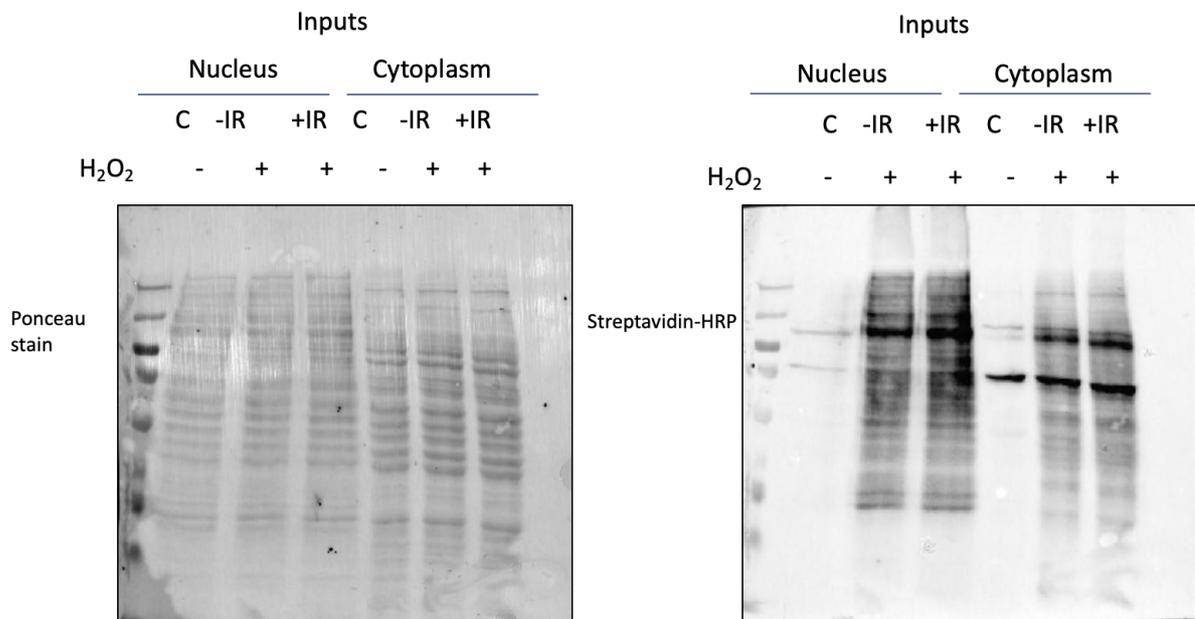
Identifier	log(l)	rl	log(e)	pI	Mr	Description
<b>Control</b>						
ENSP00000344848	3.77	2	-10.2	5.6	515.9	plectin
ENSP00000354486	4.36	2	-9.7	5.3	261	golgin A4
<b>minus IR</b>						
ENSP00000216181	4.47	16	-127	5.5	226.4	myosin heavy chain 9

<u>ENSP00000363822</u>	4.6	9	-67.8	6	99.1	<u>androgen receptor</u>
<i>plus IR</i>						
ENSP00000216181	4.9	26	-157	5.5	226.4	myosin heavy chain 9
<u>ENSP00000363822</u>	4.86	17	-91.4	6	99.1	<u>androgen receptor</u>

Due to the relatively small number of identified proteins compared to previous work performed in the host laboratory that identified 271 proteins interacting with AR-V7 by RIME utilising FLAG antibody IP, the experiment was scaled up from 10 cm dishes to 15 cm dishes and the amount of nuclear extract added to the streptavidin beads was doubled to ~120 µg. The scaled-up experiment validated that the APEX2-AR-V7 fusion was still able to label neighbouring proteins that can be identified using western blot (see Figure 6.21).

For the subsequent experiments, Glasgow Polyomics was used to acquire MS data due to issues with the facility at Newcastle University. Similar to the first experiment, the labelling reaction was performed, the nuclear extract was isolated and quantified, and an input sample was taken for western blot analysis to confirm successful labelling. To validate there is genuinely higher levels of biotinylated proteins in the experimental arms treated with H<sub>2</sub>O<sub>2</sub> than the control arms, lysates were quantified using a Pierce 660nm kit and 10 µg of lysate was run on a western blot and the membrane was stained with Ponceau S solution (Figure 6.21). This confirmed equal amounts of protein was loaded between samples. To validate that biotinylation occurred, either a biotin or streptavidin antibody was used. Figure 6.21 showed the level of biotinylation in the nuclear lysates for the minus and plus IR arms is similar, and there are considerably less biotinylated proteins detected in the control arm. The western blot

also showed the presence of the endogenously biotinylated proteins are mainly in the cytoplasmic fraction.



**Figure 6.21 Biotinylation of proteins was validated in samples that were subsequently analysed by mass spectrometry**

CWR22Rv1-AR-EK cells were transfected with 10 µg pLV-FLAG-APEX2-AR-V7 on the day of seeding and again 48 hours later prior to treatment with biotin-phenol and +/- IR (4 Gy) and the cells were incubated for 2 hours. In the minus and plus IR arms, H<sub>2</sub>O<sub>2</sub> was added to cells to induce the labelling reaction. Cells were then quenched, harvested and the cytoplasmic and nuclear fractions were isolated and quantified. 10 µg of the lysate analysed via western blot which was subject to Ponceau Stain then streptavidin-HRP.

30 µg streptavidin beads were added to the remaining ~120 µg nuclear lysate, this was then incubated overnight at 4 °C. The following day, the beads were washed 7 times with RIPA buffer and 7 times with PBS to minimise contamination of non-labelled proteins and remove detergents. The protein-bound streptavidin beads were then sent to Polyomics where an on-bead trypsin digest and clean-up was performed before injecting the samples into an Orbitrap mass spectrometer.

The first replicate had a high level of keratin contamination. This is problematic because a high concentration of keratins and other contaminants can interfere with MS analysis as they cause

reduced signal and limits protein identification. Therefore, for subsequent experiments, wash steps were performed in a laminar flow hood and filter tips were used so introduction of contaminants into samples was minimised. Subsequent runs identified substantially higher numbers of proteins. The Thermo RAW files were analysed using MaxQuant to carry out label-free identification to provide an iBAQ value that represents protein abundance. The protein lists were then processed to omit common contaminants and to filter for proteins that were identified by >2 unique peptides. Table 6.5 shows there are a higher number of proteins identified in replicates n=2-4 when compared to n=1 where there was a high level of keratin.



**Table 6.5** Number of proteins identified by mass spectrometry for each proximity labelling replicate

<b>CONTROL</b>	<b>No. of proteins</b>	<b>- IR</b>	<b>No. of proteins</b>	<b>+ IR</b>	<b>No. of proteins</b>
iBAQ_n1	65	iBAQ_n1	142	iBAQ_n1	174
iBAQ_n2	33	iBAQ_n2	265	iBAQ_n2	427
iBAQ_n3	64	iBAQ_n3	444	iBAQ_n3	436
iBAQ_n4	166	iBAQ_n4	484	iBAQ_n4	510

Due to keratin contamination causing a lower number of identified proteins, replicate 1 was omitted from subsequent analyses of replicates 2-4, where precautions had been taken to minimise keratin levels. The clustering of each experimental arm was investigated using a heat map and principal component analysis. Expectedly, this showed that the control samples cluster away from the minus and plus IR arms. As seen in the heat map, the control samples cluster to the right, as does the minus IR\_n2, likely due to the smaller number of proteins

present in this experimental arm (Figure 6.22). There was no distinct clustering between the samples that had been irradiated or the samples that had not. This is likely due to the APEX2-AR-V7 fusion not necessarily labelling exclusively at the sites of DNA damage and because there will be endogenous DNA damage occurring in steady state conditions. The PCA plots also show subtle differences between the minus and plus IR arms. This is most prominent in replicate 2, but again this is likely due to the lower number of proteins being identified in the minus IR arm (Figure 6.23).

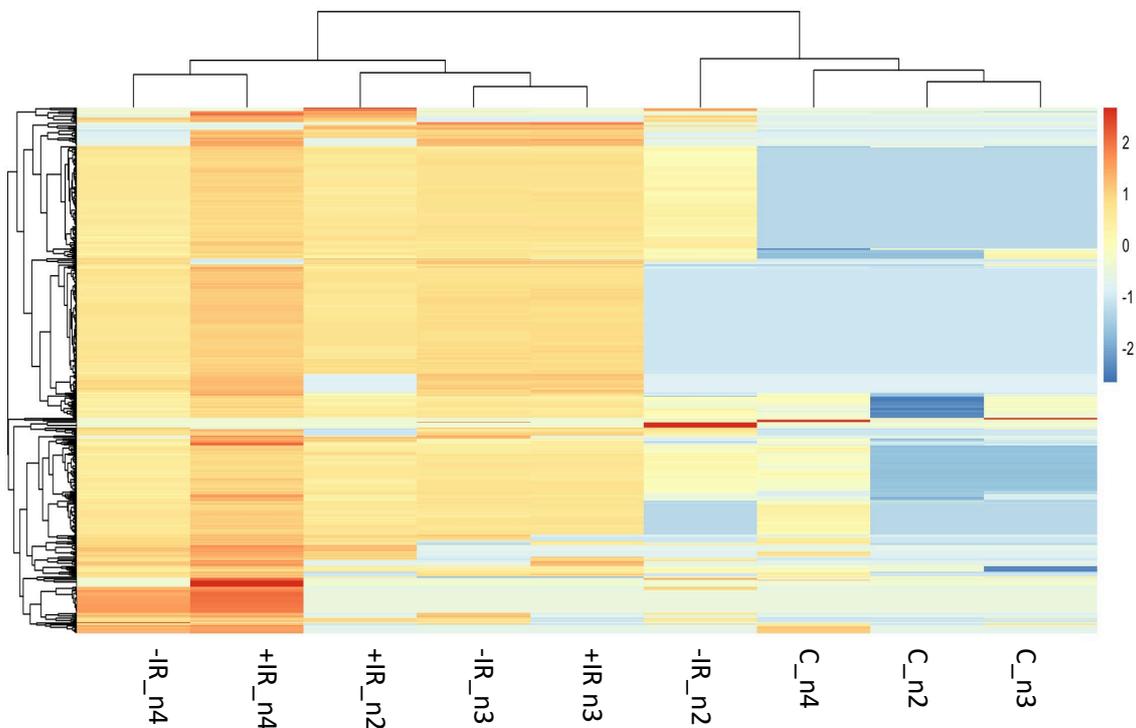
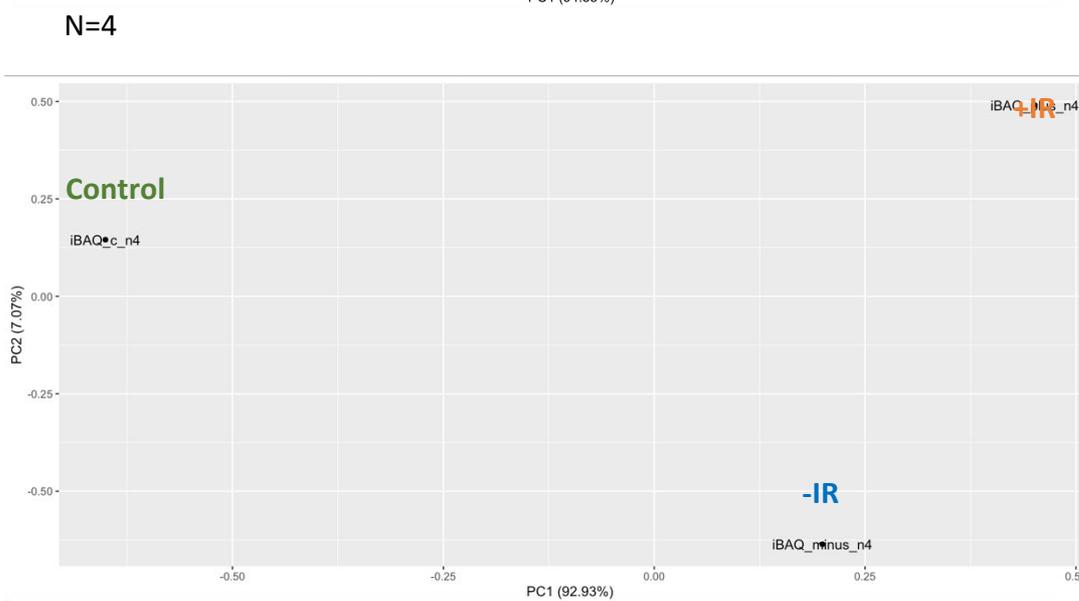
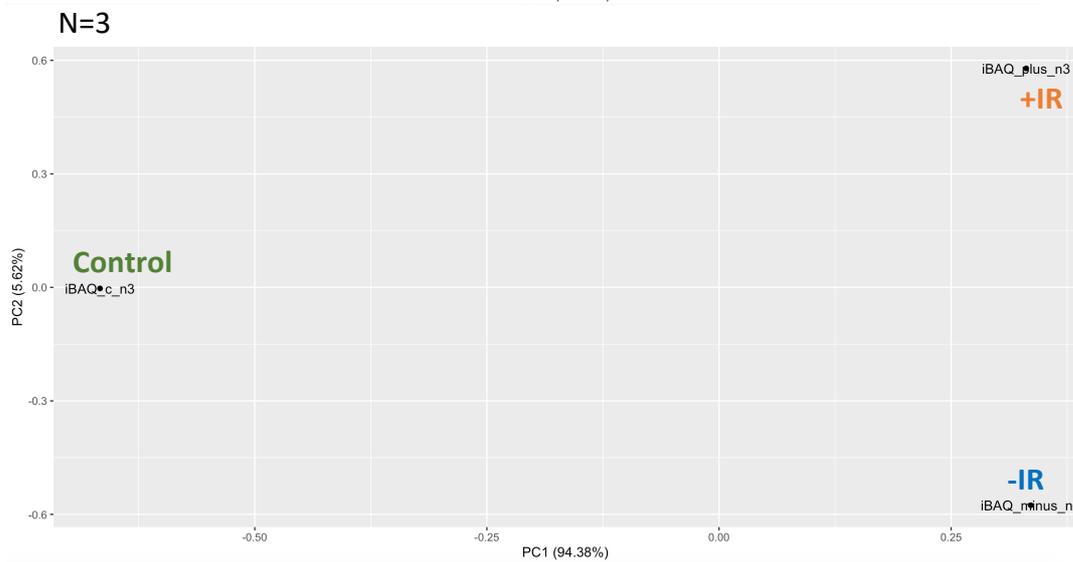
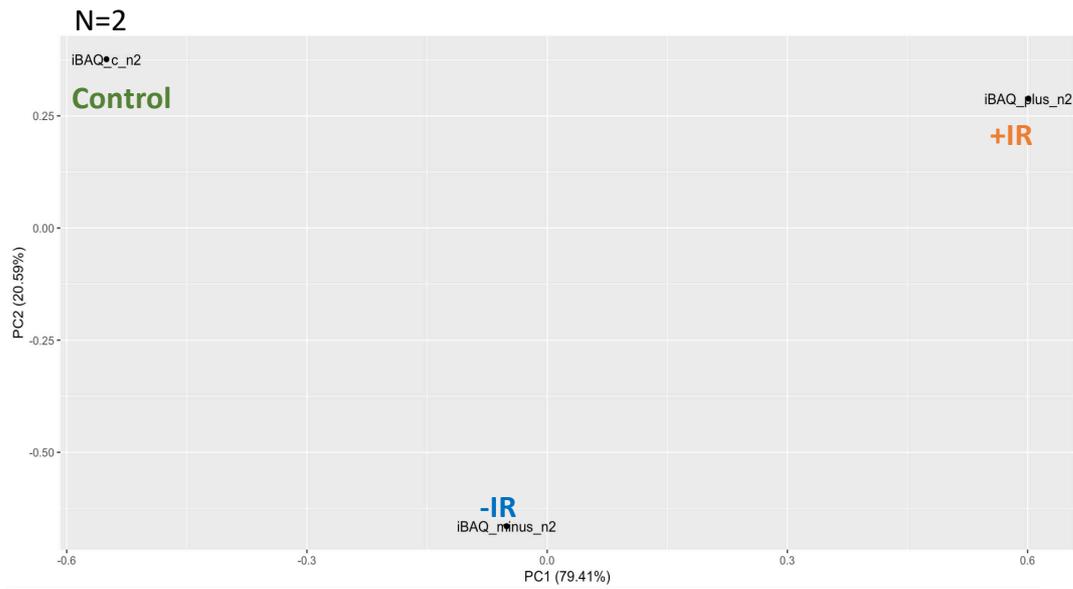


Figure 6.22 Heat map of identified proteins in each experimental arm across three replicates



**Figure 6.23 Principal component analysis of proteomics data of each of three replicates.**

The identified proteins that had >2 unique peptides and were identified in  $\geq 2$  replicates were plotted using their mean riBAQ values (Figure 6.24). The proteins with the highest abundance (riBAQ) are closer to 0 and the least abundant proteins are closer to -20. Reassuringly, AR was the fourth most abundant protein in both the minus- and plus-IR arms and known AR-interacting proteins were identified including PARP1 and FUS (Kounatidou *et al.*, 2019, Haile *et al.*, 2011). Interestingly, DNA-PKcs was identified as an AR-V7 interacting protein both in steady-state and post-irradiation. This confirms what has been shown in Chapter 4 and 5 in which DNA-PKcs was characterised as an AR-V coregulator and supports previous published findings (Yin *et al.*, 2017).

## APEX2-AR-V7 Interacting Proteins

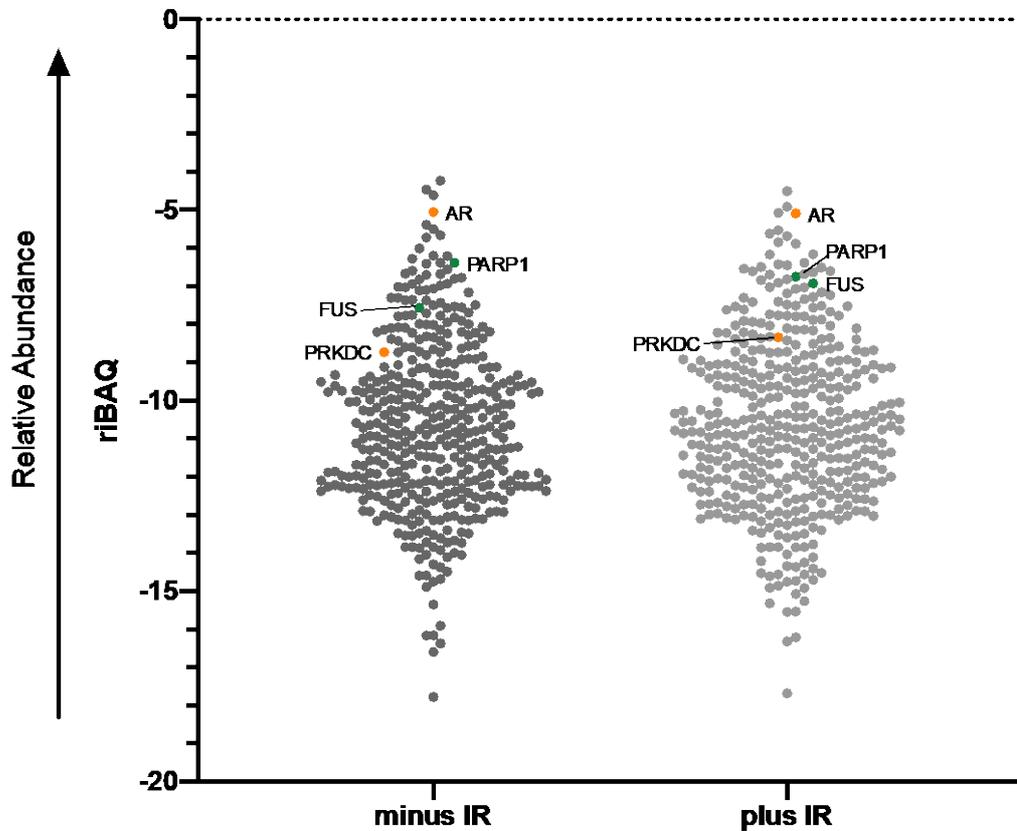


Figure 6.24 APEX2-AR-V7 interacting proteins identified by mass spectrometry in the presence and absence of ionising radiation.

Mean riBAQ scores of all APEX2-AR-V7 interacting proteins identified by mass spectrometry were plotted using PRISM. AR and DNA-PKcs are highlighted in orange and two known AR interactors, PARP1 and FUS are highlighted in green.

The FC of proteins between the two conditions was next determined to assess the interactome changes that occurs upon DNA damage. The riBAQ values were input into PRISM, the mean riBAQ for each protein was determined and the difference between the plus and minus IR arms was calculated. A t-test was also performed to determine if the changes in abundance are statistically significant. The resulting volcano plot is presented in Figure 6.25. Proteins that have a FC of  $>1.5$  are present towards the right of the plot and the proteins that have a FC of  $<-1.5$  are present towards the left of the plot. There are 73 proteins whose abundance increase by 1.5-fold in response to ionising radiation and 62 proteins whose abundance decreased in response to ionising radiation (Appendix 8.1). Out of these proteins, there are

17 proteins that are significantly increased or decreased in response to DNA damage (Table 6.6).

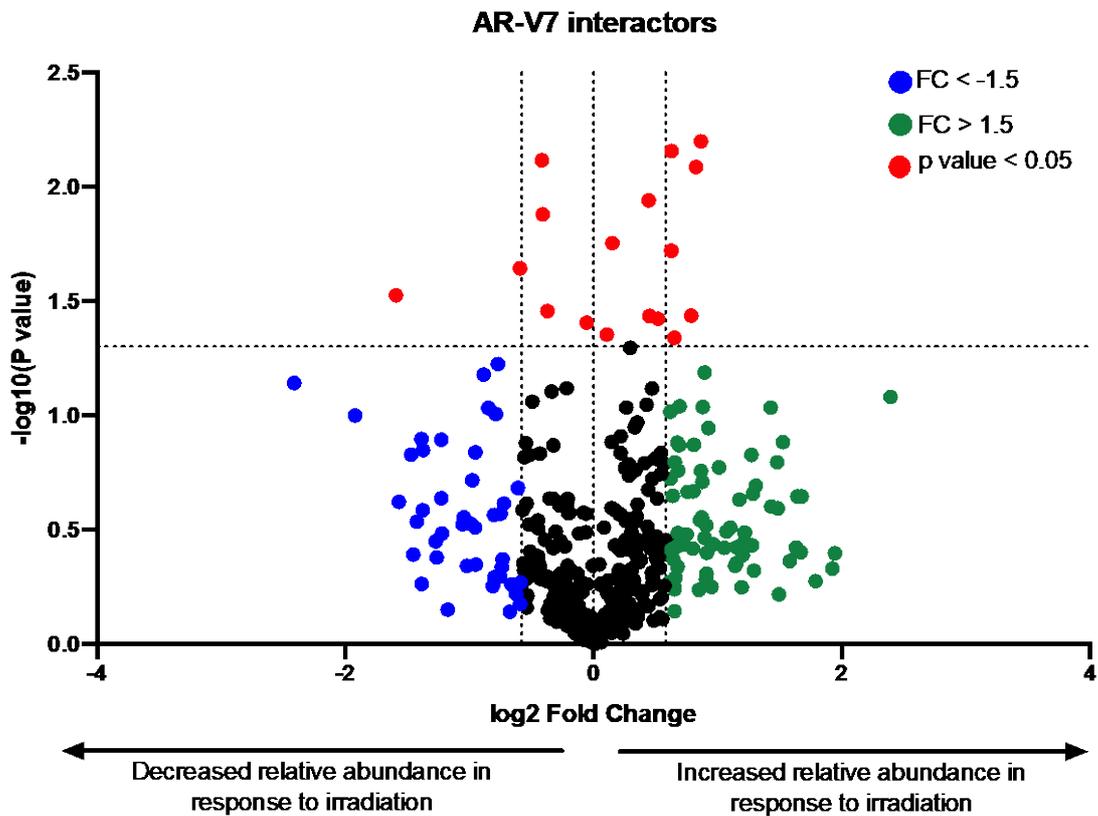


Figure 6.25 APEX2-AR-V7 interacting proteins

riBAQ scores for all interacting proteins for both plus and minus experimental arms were analysed using multiple paired t-tests and the differences between the two conditions were plotted against the  $-\log_{10}$  p value.

Table 6.6 Proteins that abundances are statistically significantly altered in response to irradiation. Proteins in blue are those that have riBAQ scores that are significantly decreased in response to irradiation.

Protein	p value	Protein	p value
TNRC18	0.006336	SEC61A1	0.029899
PABPN1	0.006971	SLTM	0.035009
RBM4	0.007679	PRCC	0.036644
PLEC	0.008211	IK	0.036819
RPS8	0.011480	CDC5L	0.037821
MED17	0.013217	FTSJ3	0.039354

DDX1	0.017675	RUVBL1	0.044374
HSPA9	0.019045	ATP5C1	0.045894
NCOR1	0.022783		

The proteins whose abundance increased by >1.5 fold in response to irradiation are presented in a heat map to visualise the level of change between the two conditions (Figure 6.26). These proteins are also presented in Figure 6.27. The list of proteins was subsequently input into Gene Ontology analysis which indicated that proteins that are more abundant in the plus IR arm, are involved in biological processes such as splicing and mRNA processing, rather than processes related to the DDR.

Interestingly, RBMX and its paralogue RBMXL1 were both identified as members of the AR-V7 interactome. Furthermore, they were both increased by >1.5 fold in response to irradiation. RBMX was identified as an AR-V splicing regulator in Chapter 5. RBMX has also been identified as a regulator of DNA repair. It was reported that RBMX accumulates at DNA lesions in a PARP1 dependent manner which promotes HR (Adamson *et al.*, 2012).

**Proteins with a >1.5-fold increase in relative abundance in response to irradiation**

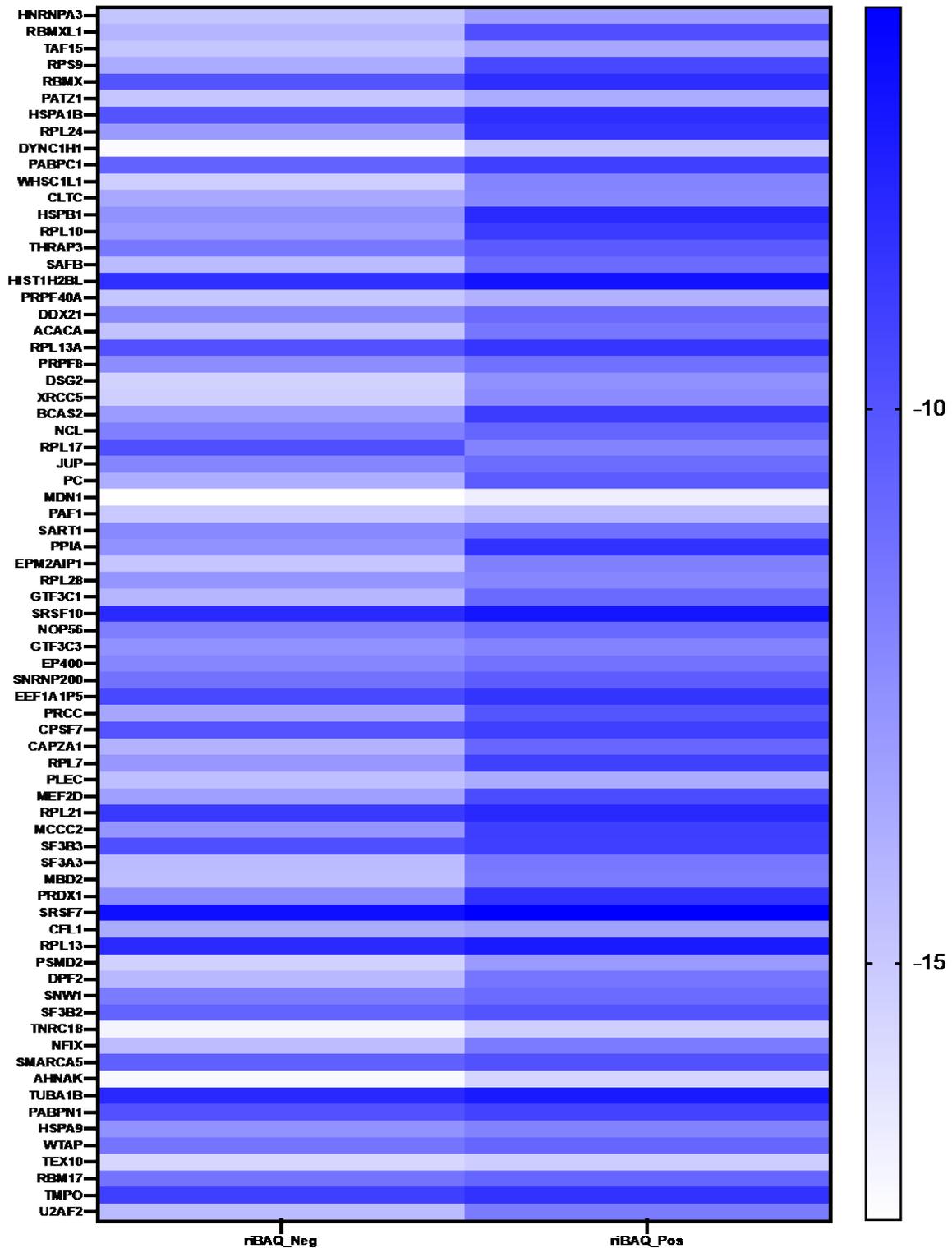
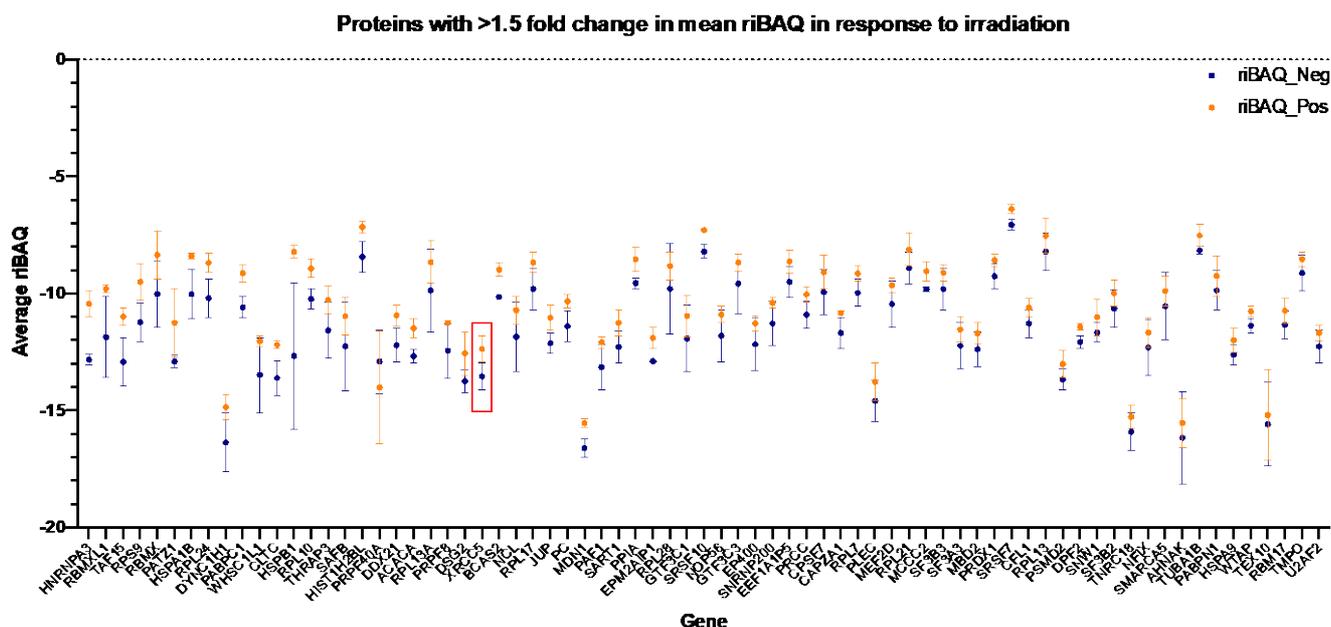


Figure 6.26 Proteins with a riBAQ value >1.5-fold higher in response to irradiation.

APEX2-AR-V7 interacting proteins that have a riBAQ score >1.5-fold higher in response to irradiation. Data points represent the mean of 2-3 replicates (depending on if the protein is identified in two or three replicates).

**Table 6.7 Top 10 biological processes that are enriched in the list of proteins that are more abundant AR-V7 interactors in response to irradiation.**

<b>GO biological process complete</b>	<b>Fold Enrichment</b>	<b>False discovery rate</b>
positive regulation of mRNA splicing, via spliceosome	52.79	4.12E-05
negative regulation of mRNA splicing, via spliceosome	52.29	7.49E-04
mRNA cis splicing, via spliceosome	47.74	9.39E-04
positive regulation of mRNA processing	47.06	3.96E-06
negative regulation of RNA splicing	42.23	1.33E-03
positive regulation of RNA splicing	41.18	7.87E-06
protein refolding	37.43	1.98E-02
negative regulation of mRNA processing	36.6	2.10E-03
ribosomal large subunit assembly	30.5	3.20E-02
positive regulation of transcription by RNA polymerase I	26.57	4.46E-02



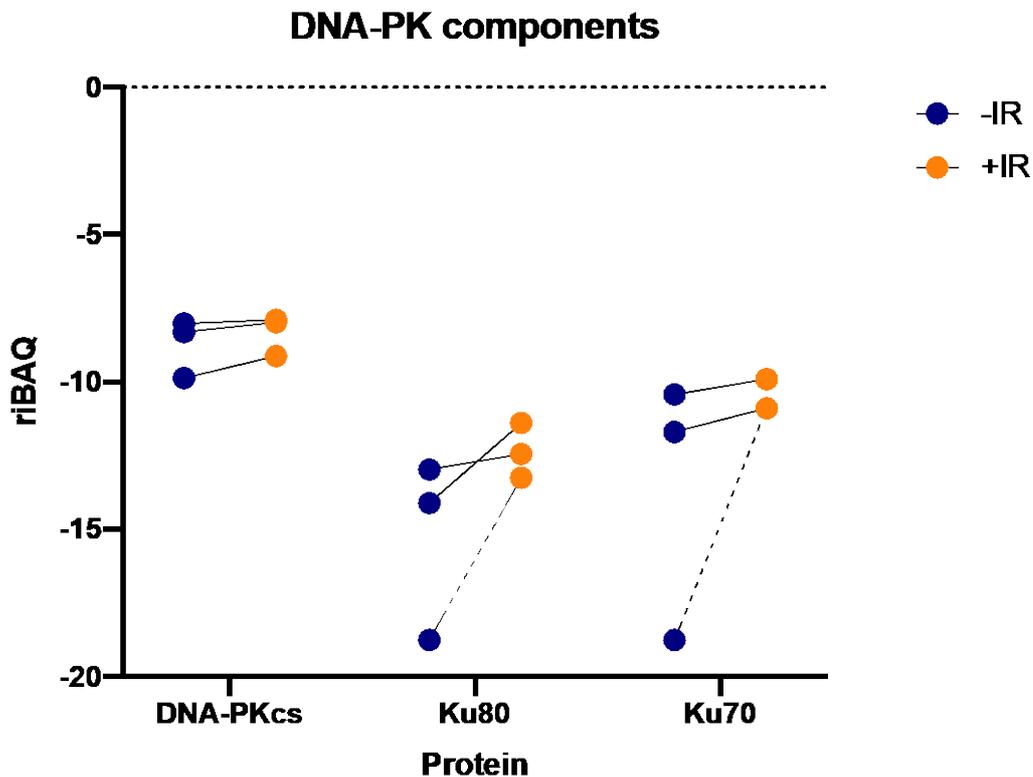
**Figure 6.27 Proteins with a riBAQ >1.5 fold higher in response to irradiation.**

APEX2-AR-V7 interacting proteins that have a riBAQ score >1.5 fold in response to irradiation. Ku80 (XRCC5) is highlighted in a red box. Data points represent the mean of 2-3 replicates (depending on if the protein is identified in two or three replicates)  $\pm$  SEM.

The other components of the DNA-PK holoenzyme were identified in the MS data. The Ku heterodimer proteins, Ku70 (XRCC6) and Ku80 (XRCC5) were identified in all three of the samples that had been treated with 4 Gy irradiation, and they were both identified in two out of the three minus IR arms. This is consistent with what has been reported previously, where tandem MS identified the three members of the DNA-PK trimeric complex as AR associated proteins (Mayeur et al., 2005). Interestingly, this study mapped the Ku70/80/AR interaction to the AR LBD (Mayeur et al., 2005) which is absent in AR-Vs meaning the AR-V-DNA-PKcs holoenzyme interaction is not depended on presence of the LBD. A more recent study using transient transfection of AR-V7 and DNA-PKcs in PC3 cells mapped the WHTLF domain in the NTD of AR-Vs as being critical for DNA-PKcs interaction. Mutant WHTLF AR-V expression vectors displayed reduced interactions as measured using proximity ligation assays (Yin et al., 2017).

Because Ku70 and Ku80 were not identified in one of the minus IR replicates, to plot these proteins to visualise the changes in abundance between the two conditions, the smallest value across all identified proteins was substituted as the value for the replicate where Ku70 and

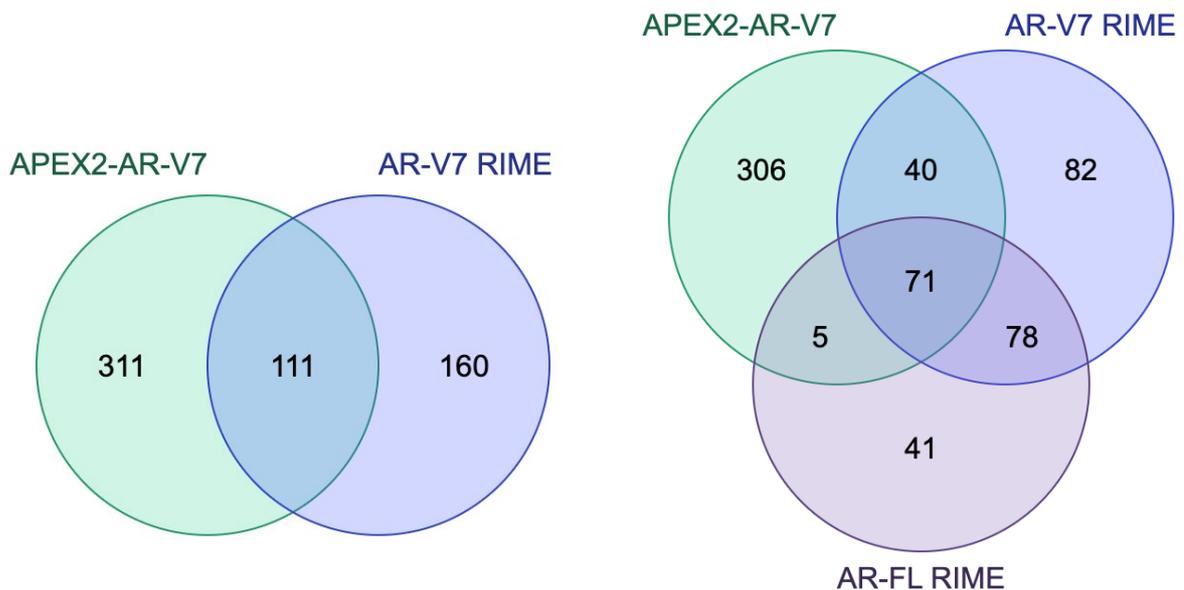
Ku80 were not identified. This showed that in response to irradiation, the abundance of each of the components of the DNA-PK holoenzyme is increased in all three replicates. This could suggest that AR-V7 is being recruited to sites of DNA damage in response to irradiation, which would support what has been shown in the literature whereby AR-V<sup>567es</sup> was shown to interact with  $\gamma$ H2AX and DNA-PKcs in response to DNA damage. It is worth noting that the abundance of immunoprecipitated DNA-PKcs between the minus- and plus-IR experimental arms does not significantly change in response to irradiation. This is likely due to DNA-PKcs being a prominent interactor of AR-Vs in steady-state conditions, primarily acting as a transcriptional co-regulator, so it was not expected that the interaction would be greatly enhanced in the irradiated samples.



**Figure 6.28** The components of the DNA-PK holoenzyme were identified as interactors of AR-V7

The rBAQ scores of each of the three replicates for the members of the DNA-PK complex, DNA-PKcs, Ku70 and Ku80 were individually plotted for -IR (blue circles) and +IR (yellow circles). In the replicates that did not identify Ku70 and Ku80 as interacting proteins, the rBAQ value was substituted for the lowest rBAQ identified in the whole dataset (the dotted line represents the replicate whereby Ku70 and Ku80 was not detected in the -IR arm).

The lists of interacting proteins were compared to the previous AR-V7 and FL-AR interactomes generated by the host laboratory using RIME (unpublished) (Figure 6.29). 41% of proteins detected in the AR-V7 RIME interactome were identified as APEX2-AR-V7 interacting proteins. This is a reasonable overlap considering the variable nature of pull-down experiments, the potential non-specific interactors identified by antibody IP approaches and the difference is data acquisition and downstream analysis. 39% of proteins that were identified as a FL-AR interactor by RIME have been identified as APEX2-AR-V7 interacting protein. The RIME experiments were performed in CWR22Rv1 cells, whereas the current experiment was performed in CWR22Rv1-AR-EK. Therefore, there could be differences in the AR-V7 interactome due to the distinct cellular backgrounds.

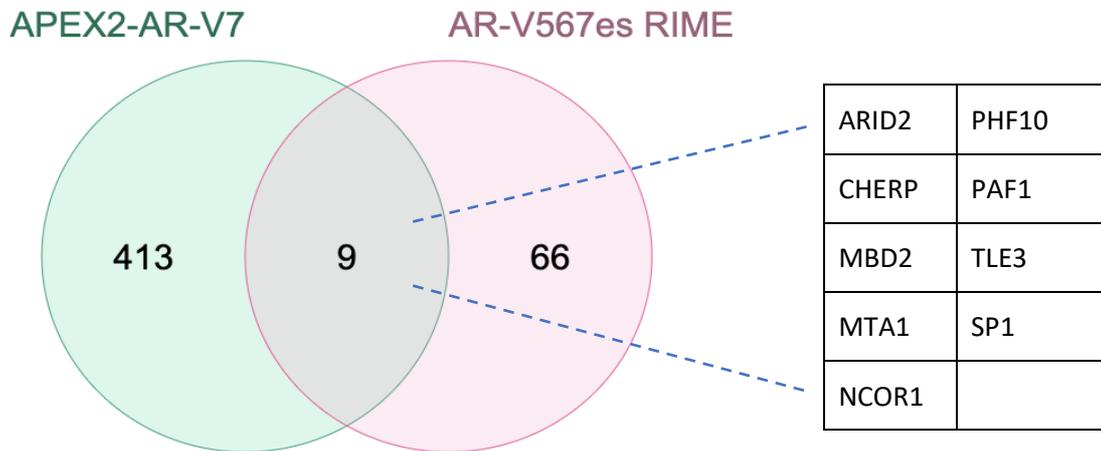


**Figure 6.29 Comparisons of the APEX2-AR-V7 interactome identified using proximity labelling and AR-V7 and FL-AR interactomes identified by RIME.**

The APEX2-AR-V7 interacting proteins were compared to the unpublished AR-V7 and FL-AR RIME interactome generated by the host laboratory (n=1).

The APEX2-AR-V7 interactome was also compared to the published AR-V<sup>567es</sup> interactome identified by RIME (Paltoglou *et al.*, 2017). The list of interacting proteins in the presence and absence of DHT was combined for comparison. This showed that just 9 proteins were shared between the AR-V7 and AR-V<sup>567es</sup> interactomes (Figure 6.30). This could be because the two

different variants have distinct interactomes or because of the different cell lines used to perform the experiments.



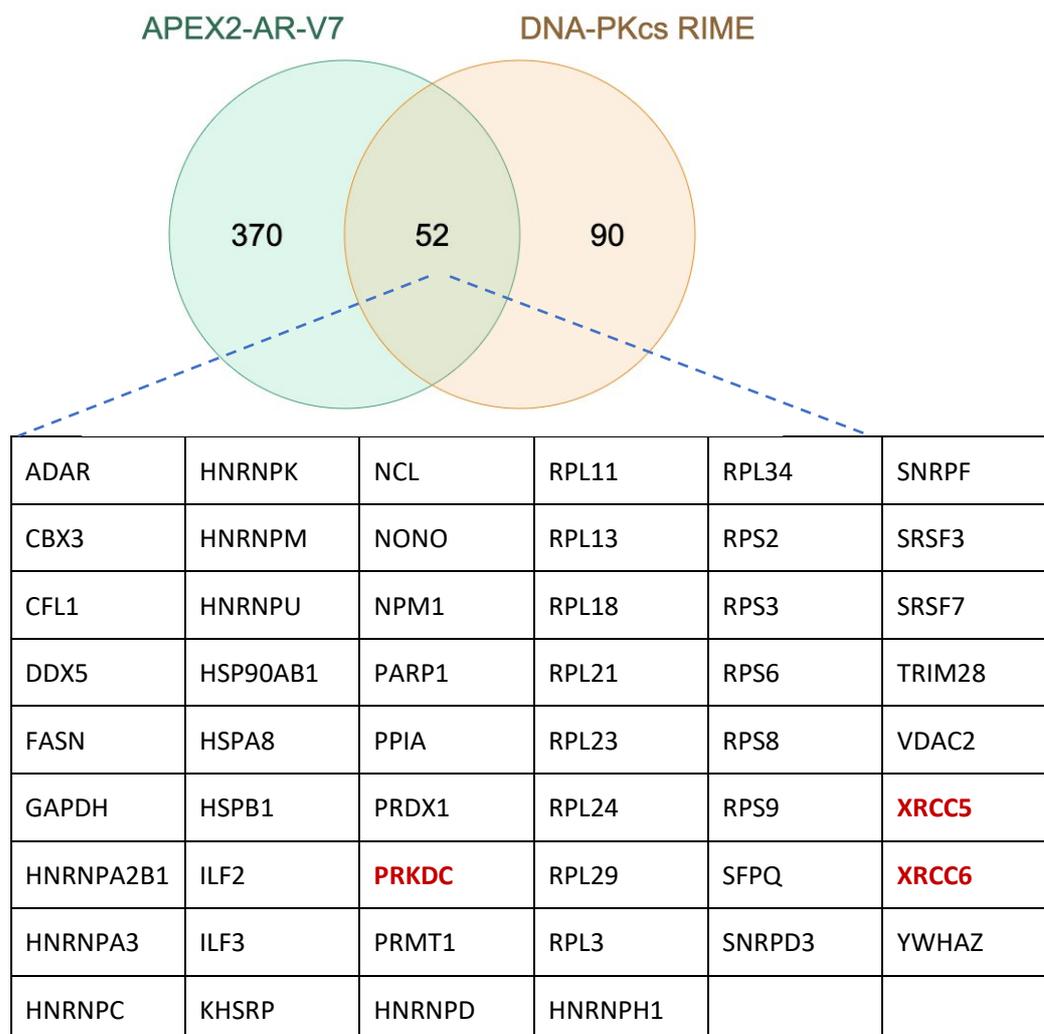
**Figure 6.30 Comparisons of the APEX2-AR-V7 interactome identified using proximity labelling and the published AR-V<sup>567es</sup> interactomes identified by RIME.**

The APEX2-AR-V7 interacting proteins were compared to the AR-V<sup>567es</sup> RIME interactome that was generated using R1-D567 cells (Paltoglou *et al.*, 2017). Final lists of AR-V<sup>567es</sup> interacting proteins were generated by only including proteins that were identified in 3 out of 3 replicates and only if they were not present in 2 or more IgG control samples. The list used to compare with the APEX2-AR-V7 list was the combined proteins identified in either minus or plus DHT experimental arms.

The AR-V<sup>567es</sup> interactome was also identified using IP-MS in R1-D567 cells (Yin *et al.*, 2017). The top 10 interactors in steady-state and in response to 10 Gy irradiation were published, which were compared to the APEX2-AR-V7 interactome. This showed in the steady-state AR-V<sup>567es</sup> interactome, PRKDC, AR, SNRP200 and LMNA were also present in our AR-V7 interactome. In response to irradiation, PRKDC, AR, SF3A1, LMNA and DHX9 were the common interacting proteins of AR-V<sup>567es</sup> and our AR-V7 interactome.

Recently, the DNA-PKcs interactome has been defined for the first time (Dylgjeri *et al.*, 2022) using RIME in the AR-V negative C4-2 PCa cells. AR was not identified in the DNA-PKcs interactome. This could be due to the way the list was filtered; any protein that was identified in the IgG protein list was omitted from the DNA-PKcs interactome. However, because AR is

an abundant protein in C4-2 cells, AR could have been immunoprecipitated by the IgG isotype control by chance, and therefore could have been filtered out of the final DNA-PKcs interactome list. Because of our interest in DNA-PKcs and AR-Vs, the overlap between our AR-V7 interactome and the DNA-PKcs RIME interactome was identified (Figure 6.31). This showed there are 52 shared interacting protein, representing 37% of the DNA-PKcs interactome. Several of these proteins are ribosomal proteins that could be identified due to their high abundances in cells, especially in the APEX2-AR-V7 interactome that was subject to a cytoplasmic nuclear extraction. If these proteins were excluded from the overlapping interactome, there are 38 remaining shared proteins. A large proportion of these are RNA binding/splicing proteins.



**Figure 6.31** Several proteins were identified as part of the AR-V7 and DNA-PKcs interactome

The APEX2-AR-V7 interacting proteins were compared to the DNA-PKcs RIME interactome that was generated using C4-2 PCa cells (Dylgjeri *et al.*, 2022).

## 6.4 Discussion

The DNA damage response is mediated by both FL-AR and AR-V signalling through regulation of several DDR associated genes (Polkinghorn *et al.*, 2013a, Kounatidou *et al.*, 2019). Consequently, attenuating FL-AR with anti-androgens such as enzalutamide provides a synergistic benefit to patients receiving radiotherapy (Warde *et al.*, 2011). It has been reported that AR-Vs may also play a direct role at the sites of DNA damage as the AR-V, AR-V<sup>567es</sup> interacts with  $\gamma$ H2AX in response to IR (Yin *et al.*, 2017). This may provide an additional mechanism of radio-resistance in AR-V positive PCa patients. Therefore, patients that express AR-Vs may not benefit from combined radiotherapy and ADT. Furthermore, because ADT can induce AR-V expression, the timings and ordering of treatment regimens could be critical to patient outcomes (Welti *et al.*, 2016). There are limited reports on how AR-Vs mediate/regulate the DDR. A recent study reported AR-Vs, specifically, AR-V7, mRNA and protein expression, are increased in CWR22Rv1 cell line and xenograft models 12-48 hours post 4 Gy irradiation. Furthermore, knockdown of AR-V7 in radioresistant C4-2 cells was able to significantly radio-sensitise cells while over expression of AR-V7 made parental C4-2 cells more radio-resistant. The study identified a potential mechanism for AR-V7-mediated radio-resistance, whereby AR-V7 promotes expression of the DNA repair protein and DNA-PKcs complex member, Ku80. (Chen *et al.*, 2022). Again, however, this study does not address the direct role of AR-Vs in DNA repair. Another recent study showed AR-V7 is involved in promoting the NHEJ DNA repair pathway by interacting with and potentially activating DNA-PKcs activity as overexpressed AR-V7 led to a greater phospho-DNA-PKcs/DNA-PKcs ratio (Luo *et al.*, 2022). They show knockdown of AR-V7 leads to prolonged  $\gamma$ H2AX foci and increased levels of DNA damage and conversely, over expression of AR-V7 led to a more rapid decline of  $\gamma$ H2AX foci and DNA damage in CWR22Rv1 and C4-2 cell line models (Luo *et al.*, 2022).

Treatments that aim to radio-sensitise PCa patients is an attractive therapeutic strategy, therefore, a better understanding of the role AR-Vs play in response to DNA damaging agents is essential. To investigate this further, an unbiased proteomics-based technique was applied in this project to define the AR-V7 interactome in steady-state and post-irradiated conditions. This was achieved using proximity biotinylation in live cells using the modified APEX2 fused to the most clinically relevant AR-V, AR-V7.

To perform this technique in the most physiologically relevant background, an endogenously tagged AR-V cell line utilising CRISPR/Cas9 technology was attempted to knock-in the APEX2-encoding cDNA immediately upstream of the *AR* gene in the CWR22Rv1-AR-EK cell line; a cell line that expresses several AR-Vs. However, even though multiple attempts using different approaches was performed, this was not successful over the course of the project. Over the past number of years, several knock-in cell lines have been successfully generated in several studies. In the host laboratory, a short sequence was knocked into the *AR* gene in the CWR22Rv1 cell line that causes a premature stop codon to be incorporated into exon 5. This leads to only AR-Vs being expressed at the protein level and no FL-AR. Furthermore, a recent study successfully knocked in the APEX2 sequence into the 53BP1, BRCA2 and MDC1 loci to define their interactomes (Gupta *et al.*, 2018). Despite several adaptations to the CRISPR pipeline, there was no successful APEX2-AR-V fusion cell lines generated. Firstly, a stably expressing doxycycline-inducible Cas9 cell line was generated because CWR22Rv1-AR-EK cells display low transfection efficiencies of plasmids. This would mitigate a low percentage of cells expressing Cas9 when delivering the CRISPR reagents to cells. This was achieved using lentiviral based transduction of the doxycycline-inducible Cas9-mSA transgene and selection of successful clones. A streptavidin-tagged Cas9 was used as it has been shown that utilisation of the strong affinity of streptavidin for biotin can increase the efficiency of knock-ins of large 5' biotinylated donor templates (Pineault *et al.*, 2019). Alongside waiting for selection and validation of the Cas9-mSA expressing cell line, transient Cas9-mSA expression was used with transfection of a biotinylated double-stranded donor template. The dsODN design included a puromycin cassette to allow selection of successful clones. This template was transfected into cells alongside transient expression of Cas9-mSA. Because the Cas9-mSA plasmid contains a puromycin selection cassette, the selection process was delayed until 12 days after Cas9-mSA transfection to hopefully diminish expression of the plasmid and therefore puromycin resistance would be due to successful knock-in of the donor template. However, the selected clones did not contain any APEX2-AR-V fusion protein but had the Cas9 transgene stably integrated and this was conferring the cells puromycin resistance. To improve this pipeline, the puromycin cassette was removed from the plasmid using a double restriction enzyme digest. However, when testing surviving clones, there were no populations of cells that contained the APEX2-AR-V fusion protein. Therefore, the selection marker in the dsODN was

changed to RFP, meaning cell sorting of RFP positive cells could be performed post-delivery of the CRISPR reagents and concurrent bulking up of the cells. However, when selecting the cells, there was a very small number passing the threshold set using a control population, and there did not appear to be a distinct population of RFP positive cells. This could be because RFP levels will be similar to endogenous AR-V levels, which may be too low to detect using the cell sorter. After screening selected clones, however, there were no successful APEX2-AR-V7 expressing clones indicating that the lack of RFP was due to failure to successfully knock-in the RFP-APEX2 cassette into the *AR* gene.

The next steps that will be undertaken to successfully generate a cell line that has endogenous APEX2-tagged AR-Vs is to generate a donor plasmid that will include 500 bp homology arms. As RFP-T2A-APEX2 is a relatively large sequence (~1500 bp) to be knocked into the genome, the presence of longer homology arms could improve the efficiency and chances of the sequence being inserted into the genome by HR. The reason shorter homology arms were used in the aforementioned attempts was because there have been publications that have successfully used donor template with homology arms as short as 35 bp (Paix *et al.*, 2017, Yu *et al.*, 2020). Another approach that could be attempted is utilising Integrated DNA Technologies CRISPR HDR donors that are modified to increase the efficiency of the knock-in. A study by Yu *et al* compared several different modified HDR templates and showed the knock-in efficiency can significantly increase when chemically modified (Yu *et al.*, 2020).

Due to the time constraints of the project, ectopic expression of a FLAG-tagged-APEX2-AR-V7 fusion protein was utilised to identify the AR-V7 interactome. This involved generating an APEX2-AR-V7-encoding construct by cloning APEX2 from the pEJS578\_DD-dSpyCas9-mCherry-APEX2 (Addgene #108570) plasmid and inserting it into the pLV-AR-V7 plasmid. This construct was verified by Sanger sequencing to ensure the sequence had been inserted and was in-frame. The expression of APEX2-AR-V7 with a 28 kDa super shift compared to AR-V7 was validated in cell lines to confirm expression of the APEX2 fusion protein. It was then validated that the presence of APEX2 at the N-terminal of AR-V7 did not impact its functionality. This was achieved by confirming APEX2-AR-V7 was recruited to canonical AR target genes by ChIP utilising a FLAG antibody. The transfection efficiency of the plasmid was then optimised to ensure enough cells expressed the APEX2-AR-V7 fusion that enabled identification of proteins downstream using MS. Several methods of transfection were attempted including using LT1

and FuGene transfection reagents, double LT1 transfection, nucleofection and lentiviral transduction. The transfection method that achieved the highest efficiency was LT1 transfection, this was performed on the day of seeding the cells and then 48 hours later. This led to a detectable smear of biotinylated proteins via western blot analysis of streptavidin enriched protein lysates after induction of the biotinylation reaction. CWR22Rv1-AR-EK cells were chosen for the labelling pipeline as these cells express high levels of AR-V7 endogenously so it was thought this would be the most physiologically relevant model. The biotinylation protocol was optimised in HEK293 cells initially, to validate that APEX2-AR-V7 plus biotin phenol and hydrogen peroxide could biotinylate proteins and that the labelling only took place when all three of these reagents were present in the cells. It was also validated that inducing DNA damage via IR did not decrease the labelling efficiency of APEX2. Two doses of irradiation were tested by comparing the biotinylation smear in the presence and absence of 2 and 4 Gy radiation. It was not apparent that irradiation impacted the labelling efficiency, therefore the 4 Gy radiation dose was taken forward in subsequent experiments in CWR22Rv1-AR-EK cells. A previous report did indicate that irradiation impacted the activity of APEX2 endogenously tagged DNA repair proteins, however, this was using a 10 Gy dose of irradiation which was 2.5-5 times higher than the doses used in this project (Gupta *et al.*, 2018). Additionally, we and others have shown that IR doses as low as 2 Gy leads to a significant DNA damage response in CWR22Rv1-AR-EK cells (Kounatidou *et al.*, 2019). Furthermore, the incubation times with biotin phenol and hydrogen peroxide was prolonged to 2 hours and 2 minutes, respectively. This was in keeping with what was used in the study by Gupta, *et al.* The reason for this was because as AR-V7 is typically present in the nucleus, and because the DNA damage response is of interest, to ensure the reagents reach the nucleus, the incubation periods were increased. In addition, a nuclear extraction was performed after harvesting the cells to concentrate the lysate, so it only contains the organelle of interest. Also, the cell line had abundant endogenously biotinylated proteins that were present in the cytoplasm in the absence of the labelling reagents. Hence, a nuclear extraction was incorporated into the protocol to remove these from the sample that was run on the mass spectrometer to improve identification of AR-V7 interacting proteins. Consequently, the experimental set-up was as follows; a double transfection of the APEX2-AR-V7 construct into CWR22Rv1-AR-EK cells, irradiate the cells with 4 Gy irradiation and incubate the cells for 2 hours before initiating the

labelling reaction, and to incubate the cells for 2 hours with biotin phenol and 2 minutes with H<sub>2</sub>O<sub>2</sub>.

The limitations of this approach are that MS can identify non-specific interactions and false positives. Therefore, stringent washes of streptavidin beads were performed, and three replicates were sent for MS. Then, as part of the MS data analysis, proteins were only included in the final lists if they were identified by more than 2 unique peptides, and if there were identified in 2 or 3 of the replicates. However, there could still be false positives in the final lists of proteins, therefore, to validate if a specific protein interacts with AR-V7, this will need to be experimentally validated using co-IP. Also, because ectopically expressing APEX2-AR-V7 produces AR-V7 protein levels that are greater than endogenous AR-V7 levels, this could also mean non-specific proteins are labelled. However, we provide for the first time, an unbiased AR-V7 interactome in steady-state and post-irradiation, identified using a proximity biotinylation approach. This includes known AR-V interactors such as PARP1 and DNA-PKcs (Luo *et al.*, 2022) and a considerable overlap of 39% with the unpublished AR-V7 RIME dataset acquired in the host laboratory.

When comparing our list of AR-V7 interactors with published AR-V interactomes, it was firstly apparent that our list has a considerably higher number of proteins. For example, the AR-V<sup>567es</sup> identified 75 proteins in the presence and absence of DHT, which included 9 common proteins to our AR-V7 interactome. We identified 422 proteins, potentially due to the capabilities of APEX2 to biotinylate weak and transient interactors, that RIME approaches would fail to detect. Another interesting observation is that some transcription-associated proteins were less abundant in response to irradiation, including SP1, TOP1, SMARCC1, SMARCB1 and FOXA1. This could suggest a shift from AR-V7s transcriptional function to another function that supports DNA repair. Additional comparisons were made between the AR-V7 interactome and the DNA-PKcs interactome identified using RIME. This revealed a considerable percentage of the DNA-PKcs interactome is also part of our AR-V7 interactome. Interestingly, a lot of these proteins were splicing-associated proteins. This relates to what I have shown in the previous chapter, whereby DNA-PKcs inhibition and depletion significantly downregulates the spliceosome gene list. It is not known if the DNA-PKcs-splicing factor interaction is because of transcription-coupled splicing or if this is a separate role of DNA-PKcs and AR-V7.

Taken together, these results suggest AR-V7 could be involved in the DDR, although it is still not known if it is directly involved in DNA repair. This requires identification of AR-V7 specifically at DNA lesions, which this study cannot distinguish. This could be achieved using another proximity labelling approach using the biotin ligase protein, TurboID which can be split and fused to two separate proteins of interest and only when they interact does the biotinylation reaction get activated (Cho *et al.*, 2020). For example, AR-V7 could be fused to the N-terminus of TurboID and a DNA repair protein, such as Ku70 or Ku80, could be fused to the C-terminus of TurboID. This means biotinylation of neighbouring proteins will only occur if AR-V7 and Ku70 or Ku80 interact in response to irradiation and the two inactive fragments of TurboID reconstitute to form an active biotin ligase. This approach is currently underway in the host laboratory and will greatly enhance the data we have already acquired as part of this project.

In summary, this APEX2 proximity-labelling approach has provided a comprehensive list of potential AR-V7 interacting proteins in steady-state and post-irradiated conditions. This provides a base list of proteins to investigate to better understand AR-V7 biology in relation to DNA repair. Also, it could provide novel interacting partners of AR-V7 that may be critical for its regulation and therefore could be targetable to diminish AR-V7 activity.

## Chapter 7 Conclusions and future work

PCa is the second leading cause of cancer related deaths in the UK. Despite advancements in our understanding of the disease, there remains critical gaps in our understanding of advanced disease. Hence, over 11,000 men every year in the UK die from PCa (CRUK). Like most cancers, the earlier stages of disease are relatively well treated with a good survival rate if the disease is diagnosed early. In 1941, pioneering work by Huggins and Hodges discovered PCa was hormone-driven, therefore PCa treatment is centralised around depleting androgens using ADT or castration (Huggins and Hodges, 1972). The AR was identified as the receptor to which androgens bind and was then recognised as a key driver of PCa both in the early stages of disease as well as driving progression to CRPC (Anderson and Liao, 1968, van der Kwast *et al.*, 1991). Anti-androgens that either directly bind and inhibit the AR or that indirectly inhibit the AR by targeting androgen biosynthesis are widely used PCa treatments. Direct inhibitors of the AR include second-generation anti-androgens enzalutamide and are initially successful in treating later-stage PCa (Tran *et al.*, 2009). Unfortunately, patients will become resistant to ADT and anti-androgens as AR signalling becomes androgen-independent and the disease progresses to CRPC (Chen *et al.*, 2004). To date, there are no effective curative treatments available for these patients, so there is a huge unmet clinical need to develop therapies to treat these patients. This will be possible when we better understand the resistance mechanisms of late-stage tumours. We do know that mechanisms, such as AR mutations, allow an antagonist to agonist switch to anti-androgens and AR gene amplification allows AR to remain active in castrate levels of androgen (Taylor *et al.*, 2010). There is also the emergence of AR-Vs that are truncated variants of the FL-AR that lack the LBD so are resistant to all current treatments (Dehm and Tindall, 2011). AR-Vs are constitutively active and do not require a ligand to become active. AR-Vs drive transcription of several canonical FL-AR target genes as well as some unique target genes such as *UBE2C* and *CCNA2*. This highlights AR-Vs distinct transcriptional programme and potentially unique regulatory mechanisms. AR-Vs are generated by alternative splicing of AR pre-mRNA transcripts that incorporate CEs and lead to the translation of truncated AR-V protein.

The aims of this thesis were to (i) investigate if DNA-PKcs is a co-regulator of AR-V transcriptional activity and investigate the mechanism of this regulatory relationship using

RNA-sequencing and (ii) interrogate AR-V7 involvement in the DDR to determine if AR-Vs are directly involved in DNA repair aside from their role of regulating transcription of DNA repair-associated genes. I believe the work presented in this thesis has shown DNA-PKcs is a key regulator of AR-V transcription as well as a regulator of transcription of the splicing factor RBMX, which is likely involved in AR-V synthesis, potentially through splicing. Furthermore, the proteomics data generated by biotin labelling the AR-V7 interactome has provided limited evidence of a shift in AR-V7 function in response to DNA damage, although further optimisation of the technique is required to better define this shift.

DNA-PKcs has been recognised as a therapeutic target in many cancer types because of its role in the DDR, particularly in NHEJ. Inhibition of DNA-PKcs improves the response to radiotherapy and other DNA damaging agents *in vitro* and *in vivo* (Zhao *et al.*, 2006, Willmore *et al.*, 2008, Ciszewski *et al.*, 2014, Dong *et al.*, 2018). Consequently, newer generation DNA-PKcs inhibitors have entered clinical trials, including the compounds M3814 (NCT02516813) and AZD7648 (NCT03907969). Aside from its function in DNA damage repair, it has been implicated in several other biological processes such as in cell cycle (Lee *et al.*, 2011), telomere maintenance (Espejel *et al.*, 2002) and transcriptional regulation (Goodwin *et al.*, 2015).

DNA-PKcs is a transcriptional regulator of FL-AR and has been shown to promote metastasis in PCa (Goodwin *et al.*, 2015). Importantly, in this study, I have shown DNA-PKcs is an important regulator of AR-V transcriptional activity, as well as regulating expression of several splicing-associated genes that are involved in the generation of AR-Vs. We firstly investigated the transcriptional effects of DNA-PKcs manipulation on AR-V target gene expression and AR/DNA-PKcs recruitment to *cis*-regulatory elements on chromatin. Here I show AR/AR-V transcriptional competency is significantly impacted when the catalytic activity of DNA-PKcs is inhibited or if DNA-PKcs is depleted. We showed DNA-PKcs is recruited to *cis*-regulatory elements upstream of AR and AR-V target genes, suggesting DNA-PKcs is directly involved in AR/AR-V transcriptional regulation. RNA-sequencing was performed to assess the global transcriptional impact upon DNA-PKcs inhibition and knockdown. Interestingly, this revealed the spliceosome gene-list was significantly downregulated in response to DNA-PKcs inhibition with NU7441 and NU5455, and depletion. This was particularly interesting as inhibiting

splicing regulators of AR-Vs could be potential therapeutic avenues for CRPC patients. Further interrogation and cross-referencing with a publicly available dataset using VCaP cells treated with the anti-androgen darolutamide provided a list of splicing factors that may be involved in AR-V generation. One of which, RBMX, was investigated as a potential splicing regulator of AR-Vs. This was achieved by using siRNA-mediated depletion of RBMX in several AR-V expressing cell lines followed by determination of AR-V protein and transcript levels. This revealed a significant depletion of several AR-V transcripts levels and AR-V7 protein levels. RNA immunoprecipitation revealed a significant enrichment of RBMX at FL-AR mRNA transcripts and preliminary work suggests enrichment at AR-V7 pre-mRNA transcripts. This suggests AR-V7 splicing regulation by RBMX could be coupled with transcription, aligning with what has been reported in which RBMX has been shown to directly interact with the m6A RNA modification that occurs co-transcriptionally (Zhou *et al.*, 2019). Future work to complete this aspect of the project should include rescue experiments, in which RBMX will be depleted and then re-introduced into cells to determine if this rescues AR-V7 levels, which will help to validate RBMX involvement in AR-V splicing. Also, RIP-sequencing would provide information on what other proteins are regulated by RBMX-associated splicing. Furthermore, DNA-PKcs ChIP sequencing will allow further interrogation of its involvement in AR-V transcriptional regulation by determining the level of overlap between AR-V binding sites on chromatin and DNA-PKcs binding sites. Given the impact of DNA-PKcs blockade on AR-V transcriptional activity, DNA-PKcs inhibition plus enzalutamide or darolutamide could be investigated to determine if DNA-PKcs inhibition can re-sensitise CWR22Rv1, or enzalutamide-resistant VCaP cell lines. Finally, to validate the impact of RBMX-regulated expression of AR-Vs being responsible for the anti-proliferative effects in PCa cell lines, AR-Vs could be reintroduced post RBMX manipulation to determine if this would rescue cells.

In the clinic, DNA-PKcs inhibitors are being investigated in early-stage trials to assess safety and efficacy in several tumour types, including prostate cancer (van Bussel *et al.*, 2021, Dylgjeri *et al.*, 2019). However, it is not yet known if certain patient cohorts may be more responsive to DNA-PKcs blockade. For example, approximately 10-15% of patients have a mutation in the *SPOP* gene that impacts function of the SPOP protein (Barbieri *et al.*, 2012). Because SPOP is involved in the regulation of DNA repair genes that are associated with the HR pathway, SPOP mutant patients may have a defective HR pathway and therefore will rely more heavily on the

NHEJ pathway, rendering them more sensitive to DNA-PKcs inhibitors (Boysen *et al.*, 2015). A better understanding of patient groups that will respond better to DNA-PKcs blockade could greatly increase their chances of success in the clinic. Furthermore, it has been reported that DNA-PKcs inhibition can modulate immune response, in combination with IR (Nakamura *et al.*, 2021). Given that IR-induced DNA damage is likely to induce an adaptive immune response via the cGAS/STING pathway, and that DNA-PKcs has been shown to inhibit this pathway, we hypothesise DNA-PKcs blockade will enhance anti-tumour immunity in PCa (Carr *et al.*, 2022). In addition, we have shown DNA-PKcs significantly alters alternative splicing, suggesting that blockade of DNA-PKcs in PCa may facilitate generation of neoantigens to enhance adaptive immune cell-tumour engagement (Lu *et al.*, 2021). This raises the exciting prospect of utilising DNA-PKcs inhibitors in combination with immunomodulatory drugs to effectively cause tumour regression in advanced PCa.

The second part of the project was to determine the AR-V7 interactome using a novel technique called proximity labelling. This involved optimisation of a CRISPR pipeline to endogenously tag AR-Vs with APEX2 at the N-terminal, adjacent to the start codon in exon 1 of the *AR* gene. This was attempted using numerous pipelines including generation of a doxycycline-inducible Cas9-mSA expressing CWR22Rv1-AR-EK cell line, then attempting to knock in APEX2 by incorporating a puromycin or RPF selection cassette into the donor template. However, due to the time constraints of the project, this was not achieved. Therefore, an APEX2-AR-V7 construct was generated that was used to ectopically express the fusion protein. Development of an APEX2 knock-in cell line will be beneficial in the future, not only to validate what we have shown with the APEX2-AR-V7 construct but to optimise the CRISPR knock-in pipeline to endogenously tag AR or other proteins with APEX2, GFP or other tags to facilitate studying their biology. Future work could involve generation of this cell line to improve the APEX2 proximity labelling pipeline described in this thesis. Using ectopic expression of APEX2-AR-V7 has allowed optimisation of the proximity labelling pipeline and has ultimately provided a comprehensive list of AR-V7 interacting proteins in steady-state and post-irradiated conditions. What remains unknown is what the AR-V7 interactome is specifically at DNA lesions, as this project captured the general AR-V7 interactome in response to DNA damage. Although the labelling reaction was initiated post-irradiation, even if AR-V7

is directly involved in the repair process, APEX2-AR-V7 will not necessarily exclusively be located at sites of DNA damage.

The next steps to progress this project could be to validate some potential AR-V7 interacting proteins identified from the proteomics data. To narrow down current lists, the following comparisons and analyses could be performed, (i) determine which of the identified proteins are upregulated in tumour versus normal samples (ii) identify proteins that are upregulated in metastatic CRPC patients that have a high likely hood of AR-V7 expression, (iii) identify proteins that are upregulated in AR-V7 positive patients and rank them based on which ones are the most significantly upregulated and (iv) compare lists of proteins to potential hits that have been identified by other members of the host laboratory by alternative techniques. This could result in a smaller list of proteins that could be validated as potential AR-V7 interactors and provide evidence for therapeutically targeting them. A potential protein that could be validated is tripartite motif containing 28 (TRIM28). TRIM28 is a member of the TRIM family of proteins and is a transcriptional regulator. It has been reported to interact with the full length AR in PCa and was identified at AR sites on chromatin that was independent of androgen stimulation (Stelloo *et al.*, 2018). TRIM28 is upregulated in CRPC samples and was identified in both the APEX2-AR-V7 interactome and the AR-V7 RIME interactome (unpublished, generated by a member of the host laboratory) and would warrant further investigation.

Taken together, the findings of the work presented provide a rationale for applying DNA-PKcs inhibitors in AR-V positive PCa patients as it has been shown that they partially inhibit AR-V signalling and decrease AR-V synthesis which ultimately leads to a reduction of PCa cell growth. Moreover, the kinase-independent and potential scaffolding role of DNA-PKcs appears to be highly important in AR-V regulation as well as the transcriptional regulation of splicing factors that are correlated with an upregulation of AR-V production. This could be investigated further by generating kinase-dead and knockout cell lines and characterising them to determine what their distinct responses are.

Although a clear positioning of DNA-PKcs inhibitors in the clinical setting has not yet been established, there are current early-stage trials ongoing that aim to sensitise patients to DNA damaging agents. Data from these trials could be utilised to determine which subgroups of

patients respond better to treatment and if there are CRPC patients included, their AR-V status could be determined to see if this aligns with what we predict.

In terms of what this work has provided to the overall field of AR-V biology, I have comprehensively studied DNA-PKcs role in AR-V driven PCa cell lines and provided some key pathways that are significantly altered upon DNA-PKcs inhibition and knockdown. One of which, splicing, has been interrogated in greater detail and I have shown DNA-PKcs regulates transcription of several splicing factors that could contribute to anti-proliferative effects seen in PCa cell line when DNA-PKcs activity is compromised. There are several other avenues worthy of further exploration as a result of the work that will be published from this project, including DNA-PKcs potential involvement in the immune response as we speculate the splicing changes in response to DNA-PKcs blockade may enhance neoantigen presentation to elicit adaptive anti-tumour immune modulation. I have also provided a comprehensive list of possible AR-V7 interacting proteins in steady-state and post DNA damage, which is a valuable contribution that will hopefully be of use to the wider scientific community.

## Chapter 8 Appendix

### 8.1 Splicing genes differentially expressed in response to DNA-PKcs manipulation

<b>NU5455</b>	NES = -1.83	<b>siDNA-PKcs</b>	NES = -2.34	<b>NU7441</b>	NES = -2.3
<b>Gene</b>	<b>Log2FC</b>	<b>Gene</b>	<b>Log2FC</b>	<b>Gene</b>	<b>Log2FC</b>
THOC1	0.203	PRPF40B	0.723	HNRNPA3	0.139
HNRNPA3	0.192	DDX42	0.433	THOC2	0.097
THOC2	0.177	ISY1	0.343	DDX39B	0.076
CDC40	0.158	PHF5A	0.234	LSM8	0.069
SNRNP27	0.154	PRPF31	0.216	CRNKL1	0.068
PRPF18	0.14	ACIN1	0.213	SRSF5	0.05
CRNKL1	0.136	DDX5	0.18	DDX42	0.043
DHX16	0.125	SRSF8	0.178	PRPF18	0.042
SF3A1	0.122	HSPA6	0.154	CDC40	0.028
PRPF8	0.121	PLRG1	0.137	THOC1	0.027
SRSF5	0.116	LSM7	0.097	SLU7	0.021
U2AF1	0.1	USP39	0.088	PRPF40B	0.017
DHX8	0.099	SNRNP70	0.071	PRPF38A	0.016
DDX42	0.098	HNRNPA1	0.014	SNRNP27	0.015
ACIN1	0.082	THOC1	-0.008	U2SURP	0.011
PLRG1	0.078	PPIL1	-0.027	ACIN1	0.011
DDX39B	0.078	SNU13	-0.037	PRPF38B	0.008
SF3B1	0.076	PPIE	-0.043	PRPF40A	0.008
SNRNP70	0.052	THOC2	-0.043	DHX8	0.006
SF3B2	0.048	SRSF5	-0.045	U2AF1	0.004
SLU7	0.046	CWC15	-0.045	SF3B1	0.003
HNRNPU	0.044	PRPF19	-0.066	PRPF8	0.001
SRSF4	0.022	LSM8	-0.074	ZMAT2	0
CDC5L	0.004	SYF2	-0.083	NCBP2	-0.015
PRPF38A	0	PRPF38B	-0.086	SF3A1	-0.016
DDX5	-0.007	ZMAT2	-0.094	DHX16	-0.016
AQR	-0.007	PQBP1	-0.105	SF3B2	-0.016
PRPF38B	-0.01	HSPA2	-0.111	SNRPA	-0.017
SNRNP200	-0.022	SNW1	-0.118	AQR	-0.017
SRSF10	-0.024	SLU7	-0.126	TRA2B	-0.018
PRPF40A	-0.028	SF3B6	-0.129	SNRNP70	-0.019
ZMAT2	-0.033	RBM25	-0.145	HNRNPU	-0.021
CCDC12	-0.04	DHX16	-0.156	PLRG1	-0.025
BUD31	-0.052	SRSF9	-0.157	RBM17	-0.025
PRPF31	-0.053	DDX23	-0.159	DDX5	-0.025
SNW1	-0.061	WBP11	-0.164	HNRNPA1L 2	-0.026
RBM8A	-0.066	SF3A3	-0.182	SRSF6	-0.026

TRA2B	-0.073	TXNL4A	-0.183	SNRNP200	-0.029
U2SURP	-0.075	CRNKL1	-0.183	PRPF19	-0.029
DHX38	-0.076	SF3A1	-0.186	PRPF31	-0.03
PRPF40B	-0.079	CDC40	-0.192	TCERG1	-0.031
PPIE	-0.08	SF3B5	-0.193	SYF2	-0.033
PRPF6	-0.084	SNRPD2	-0.199	CCDC12	-0.033
CTNNBL1	-0.088	SART1	-0.209	HNRNPK	-0.033
NCBP2	-0.089	HSPA8	-0.209	PPIE	-0.035
NCBP1	-0.089	LSM6	-0.212	RBM8A	-0.038
LSM8	-0.096	SF3B1	-0.215	SRSF10	-0.038
RBM25	-0.097	CTNNBL1	-0.221	PHF5A	-0.039
TRA2A	-0.101	U2AF1	-0.223	HSPA2	-0.04
RBM17	-0.105	DHX8	-0.236	NCBP1	-0.044
EFTUD2	-0.107	PRPF38A	-0.237	CDC5L	-0.045
SF3B3	-0.108	HNRNPA3	-0.238	EFTUD2	-0.046
DDX23	-0.114	RBM17	-0.246	CHERP	-0.047
HSPA8	-0.116	CHERP	-0.247	PQBP1	-0.048
PUF60	-0.117	SF3A2	-0.252	SF3B3	-0.048
USP39	-0.122	AQR	-0.256	RBM22	-0.049
CHERP	-0.124	PRPF18	-0.257	RBM25	-0.049
SRSF6	-0.125	DHX38	-0.26	HSPA8	-0.05
SYF2	-0.126	BCAS2	-0.271	PRPF6	-0.055
CWC15	-0.127	U2SURP	-0.272	BUD31	-0.056
SNRPG	-0.127	SNRNP27	-0.274	SRSF4	-0.058
WBP11	-0.132	SF3B2	-0.277	DHX38	-0.061
U2AF2	-0.135	MAGOH	-0.28	LSM4	-0.061
PRPF4	-0.135	EFTUD2	-0.282	PUF60	-0.063
XAB2	-0.143	HNRNPU	-0.289	SNRPD3	-0.065
SF3B4	-0.143	SRSF4	-0.29	SRSF8	-0.065
HSPA1B	-0.146	PRPF6	-0.296	TXNL4A	-0.067
PCBP1	-0.147	LSM5	-0.298	CWC15	-0.07
SRSF9	-0.148	HSPA1L	-0.306	SF3A2	-0.07
SF3A2	-0.152	PRPF40A	-0.309	HNRNPC	-0.071
HNRNPC	-0.155	TCERG1	-0.316	PCBP1	-0.072
HNRNPK	-0.158	TRA2B	-0.321	BCAS2	-0.072
ISY1	-0.16	SRSF6	-0.321	USP39	-0.075
SART1	-0.163	SNRPF	-0.344	SNRPD2	-0.076
HSPA2	-0.164	HNRNPC	-0.346	SMNDC1	-0.078
SNRPA	-0.171	RBMX	-0.35	LSM5	-0.078
PRPF3	-0.173	HSPA1B	-0.357	SNRPF	-0.078
HNRNPA1L 2	-0.174	SNRNP40	-0.364	HSPA1A	-0.081
TCERG1	-0.174	CCDC12	-0.373	HNRNPM	-0.081
THOC3	-0.175	SNRPA	-0.38	SNU13	-0.083

BCAS2	-0.176	SNRPE	-0.38	LSM7	-0.083
PQBP1	-0.184	NCBP1	-0.381	XAB2	-0.084
HSPA1L	-0.186	RBM8A	-0.385	HNRNPA1	-0.085
LSM5	-0.186	DDX39B	-0.385	SF3B5	-0.086
RBM22	-0.189	HNRNPM	-0.401	LSM6	-0.087
TXNL4A	-0.19	XAB2	-0.401	DDX46	-0.088
PRPF19	-0.198	PUF60	-0.402	SNW1	-0.088
SMNDC1	-0.203	HNRNPA1L 2	-0.415	DHX15	-0.091
LSM6	-0.207	SNRNP200	-0.421	SRSF9	-0.091
SRSF2	-0.211	RBM22	-0.437	SF3B4	-0.092
HNRNPM	-0.213	BUD31	-0.437	SF3B6	-0.092
SF3A3	-0.213	SNRPC	-0.447	MAGOH	-0.093
SNRPD2	-0.215	PRPF3	-0.465	WBP11	-0.093
DDX46	-0.215	SRSF2	-0.469	U2AF2	-0.093
EIF4A3	-0.225	SNRPD3	-0.478	TRA2A	-0.096
SNRPD3	-0.241	TRA2A	-0.486	SF3A3	-0.096
SF3B6	-0.242	EIF4A3	-0.491	SART1	-0.098
PHF5A	-0.243	SMNDC1	-0.524	PRPF4	-0.103
DHX15	-0.261	HNRNPK	-0.525	SRSF2	-0.104
LSM3	-0.269	PRPF4	-0.527	EIF4A3	-0.104
SNRNP40	-0.277	THOC3	-0.532	ISY1	-0.108
LSM4	-0.278	PRPF8	-0.547	SRSF7	-0.114
SNRPC	-0.284	SNRPD1	-0.557	CTNBL1	-0.115
SRSF3	-0.297	SNRPB2	-0.561	DDX23	-0.115
RBMX	-0.3	SRSF7	-0.583	SRSF1	-0.118
SNRPB2	-0.317	LSM3	-0.586	PRPF3	-0.122
MAGOH	-0.318	LSM4	-0.602	SNRPG	-0.125
HSPA1A	-0.327	SRSF10	-0.603	PPIH	-0.126
SNRPA1	-0.329	PPIH	-0.613	SNRNP40	-0.128
SRSF7	-0.335	LSM2	-0.62	RBMX	-0.129
SF3B5	-0.341	U2AF2	-0.628	THOC3	-0.131
SNRPB	-0.345	HSPA1A	-0.632	HSPA1B	-0.137
LSM7	-0.348	CDC5L	-0.635	SRSF3	-0.139
PPIH	-0.36	SF3B3	-0.677	SNRPB2	-0.14
SNRPF	-0.364	SRSF3	-0.686	SNRPB	-0.141
SNU13	-0.375	PCBP1	-0.706	SNRPC	-0.145
HNRNPA1	-0.388	NCBP2	-0.734	LSM2	-0.145
SRSF1	-0.41	ALYREF	-0.752	SNRPD1	-0.147
LSM2	-0.416	DHX15	-0.815	PPIL1	-0.149
HSPA6	-0.432	MAGOHB	-0.827	LSM3	-0.164
MAGOHB	-0.463	SF3B4	-0.842	MAGOHB	-0.172
ALYREF	-0.492	SNRPG	-0.843	SNRPA1	-0.233
SNRPD1	-0.496	DDX46	-0.882	ALYREF	-0.25

SRSF8	-0.513	SRSF1	-0.919	SNRPE	-0.274
PPIL1	-0.574	SNRPB	-0.965		
SNRPE	-0.637	SNRPA1	-1.117		

## 8.2 dsODN sequences

### RFP-T2A-APEX2-linker

AGCGAGCTGATCAAGGAGAACATGCACATGAAGCTGTACATGGAGGGCACCGTGAACAACCACCACTTCAAG  
 TGCACCAGCGAGGGCGAGGGCAAGCCCTACGAGGGCACCCAGACCATGAAGATCAAGGTGGTGGAGGGCG  
 GCCCCCTGCCCTTCGCCTTCGACATCCTGGCCACCAGCTTCATGTACGGCAGCAAGGCCTTCATCAACCACACC  
 CAGGGCATCCCCGACTTCTTCAAGCAGAGCTTCCCCGAGGGCTTCACCTGGGAGAGGATCACCACCTACGAGG  
 ACGGCGGCGTGCTGACCGCCACCCAGGACACCAGCTTCCAGAACGGCTGCATCATCTACAACGTGAAGATCAA  
 CGGCGTGAACCTCCCCAGCAACGGCCCCGTGATGCAGAAGAAGACCAGGGGGCTGGGAGGCCAACACCGAGA  
 TGCTGTACCCCGCCGACGGCGCCTGAGGGGCCACAGCCAGATGGCCCTGAAGCTGGTGGGCGGCGGCTACC  
 TGCACTGCAGCTTCAAGACCACCTACAGGAGCAAGAAGCCCGCCAAGAACCTGAAGATGCCCGGCTTCCACTT  
 CGTGGACCACAGGCTGGAGAGGATCAAGGAGGCCGACAAGGAGACCTACGTGGAGCAGCAGAGATGGCC  
 GTGGCCAAGTACTGCGACCTGCCAGCAAGCTGGGCCACAGG **GAGGGCAGAGGAAGTCTGCTAACATGCGG**  
**TGACGTCGAGGAGAATCCTGGCCA** **GGAAAGTCTTACCCA**ACTGTGAGTGCTGATTACCAGGACGCCGTTGAG  
 AAGGCGAAGAAGAAGCTCAGAGGCTTCATCGCTGAGAAGAGATGCGCTCCTCTAATGCTCCGTTTGGCATTCC  
 ACTCTGCTGGAACCTTTGACAAGGGCACGAAGACCGGTGGACCCTTCGGAACCATCAAGCACCTGCCGAAC  
 GGCTCACAGCGTAACAACGGTCTTGACATCGCTGTTAGGCTTTGGAGCCACTCAAGGCGGAGTTCCCTATTT  
 TGAGCTACGCCGATTTCTACCAGTTGGCTGGCGTTGTTGCCGTTGAGGTCACGGGTGGACCTAAGGTTCCATT  
 CACCCTGGAAGAGAGGACAAGCCTGAGCCACCACCAGAGGGTGCCTTGGCCGATCCACTAAGGGTTCTGAC  
 CATTGAGAGATGTGTTTGGCAAAGCTATGGGGCTTACTGACCAAGATATCGTTGCTCTATCTGGGGGTCACA  
 CTATTGGAGCTGCACACAAGGAGCGTTCTGGATTGAGGGTCCCTGGACCTAATCCTCTTATTTTCGACAAC  
 TCATACTTCACGGAGTTGTTGAGTGGTGAAGAAGGTTCTCCTCAGCTACCTTCTGACAAGGCTCTTTTGT  
 TGACCCTGTATTCCGCCCTCTCGTTGACAAATATGCAGCGGACGAAGATGCCTTCTTTGCTGATTACGCTGAGG  
 CTCACAAAAGCTTTCGAGCTTGGGTTTGTGATGCC **GGATCTGGTTCAGGTTCAGGA**

**Red** – Turbo RFP <https://www.fpbase.org/protein/turborfp/>

**Green** – T2A

**Blue** – APEX2

**Yellow** – linker

### Puro-T2A-APEX2-linker

ACAGAATATAAACCAACAGTACGGCTAGCTACACGCGACGACGTCCCCAGGGCCGTACGCACCCTCG  
 CCGCCGCGTTCGCCGACTACCCCGCCACGCGCCACACCGTCGATCCGGACC GCCACATCGAGCGGGT  
 CACCGAGCTGCAAGAACTCTCCTCACGCGCGTGGGCTCGACATCGGTAAGTATGGGTTGCAGAT  
 GATGGTGCCGCGGTGGCGGTCTGGACCACGCCGAGAGCGTGAAGCGGGGGCGGTGTTCCGCCGA  
 GATCGGCCCGCGCATGGCCGAGTTGAGCGGTTCCCGGCTGGCCGCGCAGCAACAGATGGAAGGCCT  
 CCTGGCGCCGACCCGGCCAAAGGAGCCCGCGTGGTTCCTGGCAACAGTAGGAGTATCGCCAGATCA  
 TCAAGGTAAAGGTCTAGGTAGTGCTGTTGACTTCTGGAGTGGAGGCGGCCGAGCGCGCCGGGGT  
 GCCCGCCTTCTGGAGACCTCCGCGCCCCGCAACCTCCCCTTCTACGAGCGGCTCGGCTTACCCTCA  
 CCGCCGACGTCGAGGTGCCGAAGGACCGCGCACCTGGTGCATGACCCGCAAGCCCGGTGCC **GAGG**

GCAGAGGAAGTCTGCTAACATGCGGTGACGTCGAGGAGAATCCTGGCCAGGAAAGTCTTACCCAA  
 CTGTGAGTGCTGATTACCAGGACGCCGTTGAGAAGGCGAAGAAGAAGCTCAGAGGCTTCATCGCTG  
 AGAAGAGATGCGCTCCTCTAATGCTCCGTTTGGCATTCCACTCTGCTGGAACCTTTGACAAGGGCAC  
 GAAGACCGGTGGACCCTTCGGAACCATCAAGCACCTGCCGAAGTGGCTCACAGCGCTAACAAACGG  
 TCTTGACATCGCTGTTAGGCTTTTGGAGCCACTCAAGGCGGAGTTCCTATTTTGGAGCTACGCCGATT  
 TCTACCAGTTGGCTGGCGTTGTTGCCGTTGAGGTCACGGGTGGACCTAAGGTTCCATTCCACCCTGG  
 AAGAGAGGACAAGCCTGAGCCACCACCAGAGGGTCGCTTGCCCGATCCACTAAGGGTTCTGACCA  
 TTTGAGAGATGTGTTTGGCAAAGCTATGGGGCTTACTGACCAAGATATCGTTGCTCTATCTGGGGGT  
 CACACTATTGGAGCTGCACACAAGGAGCGTTCTGGATTTGAGGGTCCCTGGACCTCTAATCCTCTTAT  
 TTTCGACAACTCATACTTCACGGAGTTGTTGAGTGGTGAGAAGGAAGTCTCCTTCAGCTACCTTCTG  
 ACAAGGCTCTTTTGTCTGACCCTGTATTCCGCCCTCTCGTTGACAAATATGCAGCGGACGAAGATGCC  
 TTCTTTGCTGATTACGCTGAGGCTCACAAAAGCTTTCGAGCTTGGGTTTGTCTGATGCCGGATCTGG  
 TTCAGGTTTCAGGA

Purple –Puromycin

Green – T2A

Blue – APEX2

Yellow – linker

### 8.3 Proteins that are increased and decreased in response to ionising radiation

Gene name	Mean difference (mean riBAQ (IR) – mean riBAQ (+IR))	Gene name	Mean difference (mean riBAQ (IR) – mean riBAQ (+IR))
HNRNPA3	2.386315578	XRN2	-0.587225057716326
RBMXL1	2.069686665	HNRNPF	-0.591503615210041
TAF15	1.936051608	LYAR	-0.599986789309114
RPS9	1.725722793	NCOR1	-0.600981442031969
RBMX	1.667921337	SMARCB1	-0.601263745121731
PATZ1	1.659192323	WDR18	-0.616770679
HSPA1B.HSPA1A	1.633119588	KHDRBS1	-0.617710591154719
RPL24	1.516638658	MYO1D	-0.623371510681075
DYNC1H1	1.507086282	H3F3B.H3F3A.HIST2H3 A.HIST3H3	-0.630251076388146
PABPC1	1.469423109	SMARCC1	-0.634557335791117
WHSC1L1	1.439068665	GPATCH4	-0.656766336078592
CLTC	1.422075093	SRSF3	-0.672151434475913
HSPB1	1.407225137	LBR	-0.673958396108462
RPL10	1.305802741	RPS27L	-0.692133594223247
THRAP3	1.300629505	AKAP8	-0.693273422020004
SAFB	1.281479201	RFX1	-0.701007537812655

HIST1H2BL	1.275806854	RPS15A	-0.704390156833519
PRPF40A	1.268014597	RNF40	-0.706314111
DDX21	1.263066162	ZNF592	-0.729725754440068
ACACA	1.208054907	SNRNP70	-0.730377947043705
RPL13A	1.202160945	RPS3	-0.744258963544826
PRPF8	1.191884801	SMC3	-0.753285073935833
DSG2	1.181773544	ZNF638	-0.758350651237599
XRCC5	1.177583221	RAC2.RAC1.RAC3	-0.802883388
BCAS2	1.165274519	POP1	-0.808511693581368
NCL	1.145071995	GATAD2B	-0.809169934803274
RPL17	1.130978678	FBL	-0.818927389498311
JUP	1.093874066	DDX54	-0.838350235018844
PC	1.068685647	EWSR1	-0.843280647388339
MDN1	1.061921972	RPL19	-0.853070250488789
PAF1	1.061867199	RBM14	-0.856507481905043
SART1	1.043393793	GTF3C4	-0.857339347190745
PPIA	1.017788858	RUVBL2	-0.88240215
EPM2AIP1	0.999394624248778	RRP12	-0.883839489663945
RPL28	0.971790408172794	ZFR	-0.894756980730774
GTF3C1	0.960402353259415	RBM12B	-0.958073215945001
SRSF10	0.916979019413215	ZC3H14	-0.961529681937312
NOP56	0.905197107297823	SLC25A5	-0.962858121724992
GTF3C3	0.902217650494993	TRIP12	-0.986010866493469
EP400	0.898549786214016	KIF4A.KIF4B	-0.986247087561388
SNRNP200	0.886667607108224	DMAP1	-0.998133311224828
EEF1A1P5.EEF1A2	0.871628596020519	MTA1	-1.053641791
PRCC	0.869451314252997	PHC3	-1.05557302
CPSF7	0.857990458524567	ZNF326	-1.064538008
CAPZA1	0.855125803535316	BUB3	-1.113970783
RPL7	0.828448724189565	RBM4.RBM4B	-1.177187841
PLEC	0.815470858849842	PPP1CC	-1.184232126
MEF2D	0.801414203906946	RAVER1	-1.230379046
RPL21	0.799382256263073	ZNF687	-1.235588699

MCCC2	0.767503523784345	BRD8	-1.254858617
SF3B3	0.700505898694058	SRRT	-1.280717472
SF3A3	0.697901313565387	NOC3L	-1.307547739
MBD2	0.689727438451104	SP1	-1.393600039
PRDX1	0.688327432935035	WAPAL	-1.395462977
SRSF7	0.682478348851637	TOP1	-1.427256652
CFL1	0.673203798261834	FOXA1.FOXA3	-1.432542138
RPL13	0.668090740542978	GTPBP4	-1.462239869
PSMD2	0.662901229943257	SMC1A	-1.478976391
DPF2	0.656670282366441	EIF4A3	-1.577810551
SNW1	0.651066471	SEC61A1.SEC61A2	-1.600823911
SF3B2	0.647355688186712	ARID2	-1.931666764
TNRC18	0.647169513314157	NUP205	-2.662041783
NFIX	0.644419934125882		
SMARCA5	0.643620774471124		
AHNAK	0.636910771798274		
TUBA1B.TUBA1A.TUBA1C.TUBA4A	0.629036646		
PABPN1	0.619762079305312		
HSPA9	0.616737532152447		
WTAP	0.611684903380663		
TEX10	0.597576411302883		
RBM17	0.596598074434089		
TMPO	0.592480021656263		

## 8.4 British association for cancer research 60<sup>th</sup> anniversary meeting abstract

### Abstract

#### Session 3 - Resistance to Therapy

##### The role of DNA-PKcs in androgen receptor variant regulation

Beth Adamson<sup>1</sup>, Luke Gaughan<sup>1</sup>, Stuart McCracken<sup>1</sup>, Elaine Willmore<sup>1</sup>, Ian Hickson<sup>1</sup>

<sup>1</sup>Translational and Clinical Research Institute/ Newcastle University Centre for Cancer

The androgen receptor (AR) is a master regulator of prostate cancer (PC) development and progression, hence current therapies target the AR signalling pathway to inhibit tumour growth. The generation of alternatively spliced AR variants (AR-Vs) are a major resistance mechanism observed in patients who progress to the advanced castrate-resistant PC (CRPC) stage of disease. AR-Vs are constitutively active and drive the growth of PC in the absence of ligands. Furthermore, AR-Vs are refractory to the current repertoire of AR-targeting therapies hence there is a major drive to develop treatments that can inhibit these aberrantly functioning receptors. Targeting AR-V co-regulatory proteins that are required for enabling their function represents a tractable means for inactivating AR-Vs in advanced disease.

DNA-PKcs, a key kinase in the DNA damage response, has been shown to regulate full-length AR transcriptional activity and is upregulated in both PC and CRPC. Due to AR-Vs being previously shown to regulate DNA damage also, we hypothesised that DNA-PKcs may influence AR-V activity as a co-regulator. Here, by proximity biotinylation we show that DNAPKcs is a prominent AR-V interacting protein in the presence and absence of DNA damage and we have shown DNA-PKcs regulates AR-V transcriptional activity in CRPC cell lines. Furthermore, DNA-PKcs inhibition and depletion has anti-proliferative effects in several CRPC cell lines in the absence of DNA damage. Transcriptomic analysis of RNA-Sequencing data has revealed a potential role for DNA-PKcs in the regulation of splicing and could be important in the generation of AR-Vs. Additionally, scaffolding roles of DNA-PKcs may play a more influential role than kinase function in modulating key *cis*- regulatory elements of AR target genes. Inhibition of these roles could improve patient outcome as combining this with hormone therapy may have synergistic effects.

## Chapter 9 References

- Abida, W., Armenia, J., Gopalan, A., Brennan, R., Walsh, M., Barron, D., Danila, D., Rathkopf, D., Morris, M., Slovin, S., McLaughlin, B., Curtis, K., Hyman, D. M., Durack, J. C., Solomon, S. B., Arcila, M. E., Zehir, A., Syed, A., Gao, J., Chakravarty, D., Vargas, H. A., Robson, M. E., Joseph, V., Offit, K., Donoghue, M. T. A., Abeshouse, A. A., Kundra, R., Heins, Z. J., Penson, A. V., Harris, C., Taylor, B. S., Ladanyi, M., Mandelker, D., Zhang, L., Reuter, V. E., Kantoff, P. W., Solit, D. B., Berger, M. F., Sawyers, C. L., Schultz, N. & Scher, H. I. 2017. Prospective Genomic Profiling of Prostate Cancer Across Disease States Reveals Germline and Somatic Alterations That May Affect Clinical Decision Making. *JCO Precis Oncol*, 2017.
- Abida, W., Cyrta, J., Heller, G., Prandi, D., Armenia, J., Coleman, I., Cieslik, M., Benelli, M., Robinson, D., Van Allen, E. M., Sboner, A., Fedrizzi, T., Mosquera, J. M., Robinson, B. D., De Sarkar, N., Kunju, L. P., Tomlins, S., Wu, Y. M., Nava Rodrigues, D., Loda, M., Gopalan, A., Reuter, V. E., Pritchard, C. C., Mateo, J., Bianchini, D., Miranda, S., Carreira, S., Rescigno, P., Filipenko, J., Vinson, J., Montgomery, R. B., Beltran, H., Heath, E. I., Scher, H. I., Kantoff, P. W., Taplin, M. E., Schultz, N., deBono, J. S., Demichelis, F., Nelson, P. S., Rubin, M. A., Chinnaiyan, A. M. & Sawyers, C. L. 2019. Genomic correlates of clinical outcome in advanced prostate cancer. *Proc Natl Acad Sci U S A*, 116, 11428-11436.
- Adamson, B., Smogorzewska, A., Sigoillot, F. D., King, R. W. & Elledge, S. J. 2012. A genome-wide homologous recombination screen identifies the RNA-binding protein RBMX as a component of the DNA-damage response. *Nat Cell Biol*, 14, 318-28.
- Agarwal, N., Hutson, T. E., Vogelzang, N. J. & Sonpavde, G. 2010. Abiraterone acetate: a promising drug for the treatment of castration-resistant prostate cancer. *Future Oncol*, 6, 665-79.
- Aggarwal, R. R., Schweizer, M. T., Nanus, D. M., Pantuck, A. J., Heath, E. I., Campeau, E., Attwell, S., Norek, K., Snyder, M., Bauman, L., Lakhotia, S., Feng, F. Y., Small, E. J., Abida, W. & Alumkal, J. J. 2020. A Phase Ib/IIa Study of the Pan-BET Inhibitor ZEN-3694 in

- Combination with Enzalutamide in Patients with Metastatic Castration-resistant Prostate Cancer. *Clin Cancer Res*, 26, 5338-5347.
- Allen, B. L. & Taatjes, D. J. 2015. The Mediator complex: a central integrator of transcription. *Nat Rev Mol Cell Biol*, 16, 155-66.
- An, J., Yang, D. Y., Xu, Q. Z., Zhang, S. M., Huo, Y. Y., Shang, Z. F., Wang, Y., Wu, D. C. & Zhou, P. K. 2008. DNA-dependent protein kinase catalytic subunit modulates the stability of c-Myc oncoprotein. *Mol Cancer*, 7, 32.
- Anderson, K. M. & Liao, S. 1968. Selective retention of dihydrotestosterone by prostatic nuclei. *Nature*, 219, 277-9.
- Andrews, S. 2010. *FastQC: A Quality Control Tool for High Throughput Sequence Data [Online]* [Online]. Available: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/> [Accessed 2022].
- Antonarakis, E. S., Chandhasin, C., Osbourne, E., Luo, J., Sadar, M. D. & Perabo, F. 2016. Targeting the N-Terminal Domain of the Androgen Receptor: A New Approach for the Treatment of Advanced Prostate Cancer. *Oncologist*, 21, 1427-1435.
- Antonarakis, E. S., Feng, Z., Trock, B. J., Humphreys, E. B., Carducci, M. A., Partin, A. W., Walsh, P. C. & Eisenberger, M. A. 2012. The natural history of metastatic progression in men with prostate-specific antigen recurrence after radical prostatectomy: long-term follow-up. *BJU Int*, 109, 32-9.
- Arora, V. K., Schenkein, E., Murali, R., Subudhi, S. K., Wongvipat, J., Balbas, M. D., Shah, N., Cai, L., Efstathiou, E., Logothetis, C., Zheng, D. & Sawyers, C. L. 2013. Glucocorticoid receptor confers resistance to antiandrogens by bypassing androgen receptor blockade. *Cell*, 155, 1309-22.
- Asangani, I. A., Dommeti, V. L., Wang, X., Malik, R., Cieslik, M., Yang, R., Escara-Wilke, J., Wilder-Romans, K., Dhanireddy, S., Engelke, C., Iyer, M. K., Jing, X., Wu, Y. M., Cao, X., Qin, Z. S., Wang, S., Feng, F. Y. & Chinnaiyan, A. M. 2014. Therapeutic targeting of BET bromodomain proteins in castration-resistant prostate cancer. *Nature*, 510, 278-82.
- Atas, E., Oberhuber, M. & Kenner, L. 2020. The Implications of PDK1-4 on Tumor Energy Metabolism, Aggressiveness and Therapy Resistance. *Front Oncol*, 10, 583217.

- Balbas, M. D., Evans, M. J., Hosfield, D. J., Wongvipat, J., Arora, V. K., Watson, P. A., Chen, Y., Greene, G. L., Shen, Y. & Sawyers, C. L. 2013. Overcoming mutation-based resistance to antiandrogens with rational drug design. *Elife*, 2, e00499.
- Barbieri, C. E., Baca, S. C., Lawrence, M. S., Demichelis, F., Blattner, M., Theurillat, J. P., White, T. A., Stojanov, P., Van Allen, E., Stransky, N., Nickerson, E., Chae, S. S., Boysen, G., Auclair, D., Onofrio, R. C., Park, K., Kitabayashi, N., MacDonald, T. Y., Sheikh, K., Vuong, T., Guiducci, C., Cibulskis, K., Sivachenko, A., Carter, S. L., Saksena, G., Voet, D., Hussain, W. M., Ramos, A. H., Winckler, W., Redman, M. C., Ardlie, K., Tewari, A. K., Mosquera, J. M., Rupp, N., Wild, P. J., Moch, H., Morrissey, C., Nelson, P. S., Kantoff, P. W., Gabriel, S. B., Golub, T. R., Meyerson, M., Lander, E. S., Getz, G., Rubin, M. A. & Garraway, L. A. 2012. Exome sequencing identifies recurrent SPOP, FOXA1 and MED12 mutations in prostate cancer. *Nat Genet*, 44, 685-9.
- Basil, P., Robertson, M. J., Bingman, W. E., 3rd, Dash, A. K., Krause, W. C., Shafi, A. A., Piyarathna, B., Coarfa, C. & Weigel, N. L. 2022. Cistrome and transcriptome analysis identifies unique androgen receptor (AR) and AR-V7 splice variant chromatin binding and transcriptional activities. *Sci Rep*, 12, 5351.
- Baumgart, S. J., Nevedomskaya, E., Lesche, R., Newman, R., Mumberg, D. & Haendler, B. 2020. Darolutamide antagonizes androgen signaling by blocking enhancer and super-enhancer activation. *Mol Oncol*, 14, 2022-2039.
- Ben-Shlomo, Y., Evans, S., Ibrahim, F., Patel, B., Anson, K., Chinegwundoh, F., Corbishley, C., Dorling, D., Thomas, B., Gillatt, D., Kirby, R., Muir, G., Nargund, V., Popert, R., Metcalfe, C., Persad, R. & group, P. s. 2008. The risk of prostate cancer amongst black men in the United Kingdom: the PROCESS cohort study. *Eur Urol*, 53, 99-105.
- Bentley, D. 1999. Coupling RNA polymerase II transcription with pre-mRNA processing. *Curr Opin Cell Biol*, 11, 347-51.
- Berthold, D. R., Pond, G. R., Soban, F., de Wit, R., Eisenberger, M. & Tannock, I. F. 2008. Docetaxel plus prednisone or mitoxantrone plus prednisone for advanced prostate cancer: updated survival in the TAX 327 study. *J Clin Oncol*, 26, 242-5.

- Blackford, A. N. & Jackson, S. P. 2017. ATM, ATR, and DNA-PK: The Trinity at the Heart of the DNA Damage Response. *Mol Cell*, 66, 801-817.
- Bolla, M., Van Tienhoven, G., Warde, P., Dubois, J. B., Mirimanoff, R. O., Storme, G., Bernier, J., Kuten, A., Sternberg, C., Billiet, I., Torecilla, J. L., Pfeffer, R., Cutajar, C. L., Van der Kwast, T. & Collette, L. 2010. External irradiation with or without long-term androgen suppression for prostate cancer with high metastatic risk: 10-year results of an EORTC randomised study. *Lancet Oncol*, 11, 1066-73.
- Bonnet, S., Archer, S. L., Allalunis-Turner, J., Haromy, A., Beaulieu, C., Thompson, R., Lee, C. T., Lopaschuk, G. D., Puttagunta, L., Bonnet, S., Harry, G., Hashimoto, K., Porter, C. J., Andrade, M. A., Thebaud, B. & Michelakis, E. D. 2007. A mitochondria-K<sup>+</sup> channel axis is suppressed in cancer and its normalization promotes apoptosis and inhibits cancer growth. *Cancer Cell*, 11, 37-51.
- Boysen, G., Barbieri, C. E., Prandi, D., Blattner, M., Chae, S. S., Dahija, A., Nataraj, S., Huang, D., Marotz, C., Xu, L., Huang, J., Lecca, P., Chhangawala, S., Liu, D., Zhou, P., Sboner, A., de Bono, J. S., Demichelis, F., Houvras, Y. & Rubin, M. A. 2015. SPOP mutation leads to genomic instability in prostate cancer. *Elife*, 4.
- Bria, E., Cuppone, F., Giannarelli, D., Milella, M., Ruggeri, E. M., Sperduti, I., Pinnaro, P., Terzoli, E., Cognetti, F. & Carlini, P. 2009. Does hormone treatment added to radiotherapy improve outcome in locally advanced prostate cancer?: meta-analysis of randomized trials. *Cancer*, 115, 3446-56.
- Bunch, H., Lawney, B. P., Lin, Y. F., Asaithamby, A., Murshid, A., Wang, Y. E., Chen, B. P. & Calderwood, S. K. 2015. Transcriptional elongation requires DNA break-induced signalling. *Nat Commun*, 6, 10191.
- Butcher, S. E. & Brow, D. A. 2005. Towards understanding the catalytic core structure of the spliceosome. *Biochem Soc Trans*, 33, 447-9.
- Cai, C., He, H. H., Chen, S., Coleman, I., Wang, H., Fang, Z., Chen, S., Nelson, P. S., Liu, X. S., Brown, M. & Balk, S. P. 2011. Androgen receptor gene expression in prostate cancer is directly suppressed by the androgen receptor through recruitment of lysine-specific demethylase 1. *Cancer Cell*, 20, 457-71.

- Calsou, P., Frit, P., Humbert, O., Muller, C., Chen, D. J. & Salles, B. 1999. The DNA-dependent protein kinase catalytic activity regulates DNA end processing by means of Ku entry into DNA. *J Biol Chem*, 274, 7848-56.
- Cancer Genome Atlas Research, N. 2015. The Molecular Taxonomy of Primary Prostate Cancer. *Cell*, 163, 1011-25.
- Carr, M. I., Chiu, L. Y., Guo, Y., Xu, C., Lazorchak, A. S., Yu, H., Qin, G., Qi, J., Marelli, B., Lan, Y., Sun, Q., Czauderna, F., Zenke, F. T., Blaukat, A. & Vassilev, L. T. 2022. DNA-PK Inhibitor Peposertib Amplifies Radiation-Induced Inflammatory Micronucleation and Enhances TGFbeta/PD-L1 Targeted Cancer Immunotherapy. *Mol Cancer Res*, 20, 568-582.
- Carrie, C., Magne, N., Burban-Provost, P., Sargos, P., Latorzeff, I., Lagrange, J. L., Supiot, S., Belkacemi, Y., Peiffert, D., Allouache, N., Dubray, B. M., Servagi-Vernat, S., Suchaud, J. P., Crehange, G., Guerif, S., Brihoum, M., Barbier, N., Graff-Cailleaud, P., Ruffion, A., Dussart, S., Ferlay, C. & Chabaud, S. 2019. Short-term androgen deprivation therapy combined with radiotherapy as salvage treatment after radical prostatectomy for prostate cancer (GETUG-AFU 16): a 112-month follow-up of a phase 3, randomised trial. *Lancet Oncol*, 20, 1740-1749.
- Castro, E. & Eeles, R. 2012. The role of BRCA1 and BRCA2 in prostate cancer. *Asian J Androl*, 14, 409-14.
- Cato, L., de Tribolet-Hardy, J., Lee, I., Rottenberg, J. T., Coleman, I., Melchers, D., Houtman, R., Xiao, T., Li, W., Uo, T., Sun, S., Kuznik, N. C., Goppert, B., Ozgun, F., van Royen, M. E., Houtsmuller, A. B., Vadhi, R., Rao, P. K., Li, L., Balk, S. P., Den, R. B., Trock, B. J., Karnes, R. J., Jenkins, R. B., Klein, E. A., Davicioni, E., Gruhl, F. J., Long, H. W., Liu, X. S., Cato, A. C. B., Lack, N. A., Nelson, P. S., Plymate, S. R., Groner, A. C. & Brown, M. 2019. ARv7 Represses Tumor-Suppressor Genes in Castration-Resistant Prostate Cancer. *Cancer Cell*, 35, 401-413 e6.
- Ceraline, J., Erdmann, E., Erbs, P., Deslandres-Cruchant, M., Jacqmin, D., Duclos, B., Klein-Soyer, C., Dufour, P. & Bergerat, J. P. 2003. A yeast-based functional assay for the detection of the mutant androgen receptor in prostate cancer. *Eur J Endocrinol*, 148, 99-110.

- Cerami, E., Gao, J., Dogrusoz, U., Gross, B. E., Sumer, S. O., Aksoy, B. A., Jacobsen, A., Byrne, C. J., Heuer, M. L., Larsson, E., Antipin, Y., Reva, B., Goldberg, A. P., Sander, C. & Schultz, N. 2012. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov*, 2, 401-4.
- Chan, D. W., Chen, B. P., Prithivirajasingh, S., Kurimasa, A., Story, M. D., Qin, J. & Chen, D. J. 2002. Autophosphorylation of the DNA-dependent protein kinase catalytic subunit is required for rejoining of DNA double-strand breaks. *Genes Dev*, 16, 2333-8.
- Chan, S. C., Selth, L. A., Li, Y., Nyquist, M. D., Miao, L., Bradner, J. E., Raj, G. V., Tilley, W. D. & Dehm, S. M. 2015. Targeting chromatin binding regulation of constitutively active AR variants to overcome prostate cancer resistance to endocrine-based therapies. *Nucleic Acids Res*, 43, 5880-97.
- Chang, A. J., Autio, K. A., Roach, M., 3rd & Scher, H. I. 2014. High-risk prostate cancer-classification and therapy. *Nat Rev Clin Oncol*, 11, 308-23.
- Chaytor, L., Simcock, M., Nakjang, S., Heath, R., Walker, L., Robson, C., Jones, D. & Gaughan, L. 2019. The Pioneering Role of GATA2 in Androgen Receptor Variant Regulation Is Controlled by Bromodomain and Extraterminal Proteins in Castrate-Resistant Prostate Cancer. *Mol Cancer Res*.
- Chen, C. D., Welsbie, D. S., Tran, C., Baek, S. H., Chen, R., Vessella, R., Rosenfeld, M. G. & Sawyers, C. L. 2004. Molecular determinants of resistance to antiandrogen therapy. *Nat Med*, 10, 33-9.
- Chen, D., Chou, F. J., Chen, Y., Huang, C. P., Tian, H., Wang, Y., Niu, Y., You, B., Yeh, S., Xing, N. & Chang, C. 2022. Targeting the radiation-induced ARv7-mediated circNHS/miR-512-5p/XRCC5 signaling with Quercetin increases prostate cancer radiosensitivity. *J Exp Clin Cancer Res*, 41, 235.
- Chen, Y., Sawyers, C. L. & Scher, H. I. 2008a. Targeting the androgen receptor pathway in prostate cancer. *Curr Opin Pharmacol*, 8, 440-8.
- Chen, Y. C., Page, J. H., Chen, R. & Giovannucci, E. 2008b. Family history of prostate and breast cancer and the risk of prostate cancer in the PSA era. *Prostate*, 68, 1582-91.

- Cher, M. L., Bova, G. S., Moore, D. H., Small, E. J., Carroll, P. R., Pin, S. S., Epstein, J. I., Isaacs, W. B. & Jensen, R. H. 1996. Genetic alterations in untreated metastases and androgen-independent prostate cancer detected by comparative genomic hybridization and allelotyping. *Cancer Res*, 56, 3091-102.
- Chi, K. N., Agarwal, N., Bjartell, A., Chung, B. H., Pereira de Santana Gomes, A. J., Given, R., Juarez Soto, A., Merseburger, A. S., Ozguroglu, M., Uemura, H., Ye, D., Deprince, K., Naini, V., Li, J., Cheng, S., Yu, M. K., Zhang, K., Larsen, J. S., McCarthy, S., Chowdhury, S. & Investigators, T. 2019. Apalutamide for Metastatic, Castration-Sensitive Prostate Cancer. *N Engl J Med*, 381, 13-24.
- Cho, K. F., Branon, T. C., Udeshi, N. D., Myers, S. A., Carr, S. A. & Ting, A. Y. 2020. Proximity labeling in mammalian cells with TurboID and split-TurboID. *Nat Protoc*, 15, 3971-3999.
- Ciszewski, W. M., Tavecchio, M., Dastyh, J. & Curtin, N. J. 2014. DNA-PK inhibition by NU7441 sensitizes breast cancer cells to ionizing radiation and doxorubicin. *Breast Cancer Res Treat*, 143, 47-55.
- Claessens, F., Alen, P., Devos, A., Peeters, B., Verhoeven, G. & Rombauts, W. 1996. The androgen-specific probasin response element 2 interacts differentially with androgen and glucocorticoid receptors. *J Biol Chem*, 271, 19013-6.
- Claessens, F., Denayer, S., Van Tilborgh, N., Kerkhofs, S., Helsen, C. & Haelens, A. 2008. Diverse roles of androgen receptor (AR) domains in AR-mediated signaling. *Nucl Recept Signal*, 6, e008.
- Clegg, N. J., Wongvipat, J., Joseph, J. D., Tran, C., Ouk, S., Dilhas, A., Chen, Y., Grillot, K., Bischoff, E. D., Cai, L., Aparicio, A., Dorow, S., Arora, V., Shao, G., Qian, J., Zhao, H., Yang, G., Cao, C., Sensintaffar, J., Wasielewska, T., Herbert, M. R., Bonnefous, C., Darimont, B., Scher, H. I., Smith-Jones, P., Klang, M., Smith, N. D., De Stanchina, E., Wu, N., Ouerfelli, O., Rix, P. J., Heyman, R. A., Jung, M. E., Sawyers, C. L. & Hager, J. H. 2012. ARN-509: a novel antiandrogen for prostate cancer treatment. *Cancer Res*, 72, 1494-503.

- Coffey, K. & Robson, C. N. 2012. Regulation of the androgen receptor by post-translational modifications. *J Endocrinol*, 215, 221-37.
- Cohen, R. J., Shannon, B. A., Phillips, M., Moorin, R. E., Wheeler, T. M. & Garrett, K. L. 2008. Central zone carcinoma of the prostate gland: a distinct tumor type with poor prognostic features. *J Urol*, 179, 1762-7; discussion 1767.
- Colombel, M., Symmans, F., Gil, S., O'Toole, K. M., Chopin, D., Benson, M., Olsson, C. A., Korsmeyer, S. & Buttyan, R. 1993. Detection of the apoptosis-suppressing oncoprotein bc1-2 in hormone-refractory human prostate cancers. *Am J Pathol*, 143, 390-400.
- Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P. D., Wu, X., Jiang, W., Marraffini, L. A. & Zhang, F. 2013. Multiplex genome engineering using CRISPR/Cas systems. *Science*, 339, 819-23.
- CRUK. *Prostate Cancer Statistics* [Online]. Available: <https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/prostate-cancer#heading-Zero> [Accessed 1/5 2019].
- Culig, Z., Hobisch, A., Cronauer, M. V., Radmayr, C., Trapman, J., Hittmair, A., Bartsch, G. & Klocker, H. 1994. Androgen receptor activation in prostatic tumor cell lines by insulin-like growth factor-I, keratinocyte growth factor, and epidermal growth factor. *Cancer Res*, 54, 5474-8.
- D'Amico, A. V., Whittington, R., Malkowicz, S. B., Schultz, D., Blank, K., Broderick, G. A., Tomaszewski, J. E., Renshaw, A. A., Kaplan, I., Beard, C. J. & Wein, A. 1998. Biochemical outcome after radical prostatectomy, external beam radiation therapy, or interstitial radiation therapy for clinically localized prostate cancer. *JAMA*, 280, 969-74.
- Dalal, K., Roshan-Moniri, M., Sharma, A., Li, H., Ban, F., Hessein, M., Hsing, M., Singh, K., LeBlanc, E., Dehm, S., Tomlinson, E. S., Cherkasov, A. & Rennie, P. S. 2014. Selectively targeting the DNA-binding domain of the androgen receptor as a prospective therapy for prostate cancer. *J Biol Chem*, 289, 26417-26429.
- Davis, A. J., Chen, B. P. & Chen, D. J. 2014. DNA-PK: a dynamic enzyme in a versatile DSB repair pathway. *DNA Repair (Amst)*, 17, 21-9.

- Davis, A. J. & Chen, D. J. 2013. DNA double strand break repair via non-homologous end-joining. *Transl Cancer Res*, 2, 130-143.
- de Bono, J. S., Oudard, S., Ozguroglu, M., Hansen, S., Machiels, J. P., Kocak, I., Gravis, G., Bodrogi, I., Mackenzie, M. J., Shen, L., Roessner, M., Gupta, S., Sartor, A. O. & Investigators, T. 2010. Prednisone plus cabazitaxel or mitoxantrone for metastatic castration-resistant prostate cancer progressing after docetaxel treatment: a randomised open-label trial. *Lancet*, 376, 1147-54.
- De Laere, B., van Dam, P. J., Whittington, T., Mayrhofer, M., Diaz, E. H., Van den Eynden, G., Vandebroek, J., Del-Favero, J., Van Laere, S., Dirix, L., Gronberg, H. & Lindberg, J. 2017. Comprehensive Profiling of the Androgen Receptor in Liquid Biopsies from Castration-resistant Prostate Cancer Reveals Novel Intra-AR Structural Variation and Splice Variant Expression Patterns. *Eur Urol*, 72, 192-200.
- Debes, J. D., Sebo, T. J., Lohse, C. M., Murphy, L. M., Haugen, D. A. & Tindall, D. J. 2003. p300 in prostate cancer proliferation and progression. *Cancer Res*, 63, 7638-40.
- Dehm, S. M., Schmidt, L. J., Heemers, H. V., Vessella, R. L. & Tindall, D. J. 2008. Splicing of a novel androgen receptor exon generates a constitutively active androgen receptor that mediates prostate cancer therapy resistance. *Cancer Res*, 68, 5469-77.
- Dehm, S. M. & Tindall, D. J. 2011. Alternatively spliced androgen receptor variants. *Endocr Relat Cancer*, 18, R183-96.
- Ding, Q., Reddy, Y. V., Wang, W., Woods, T., Douglas, P., Ramsden, D. A., Lees-Miller, S. P. & Meek, K. 2003. Autophosphorylation of the catalytic subunit of the DNA-dependent protein kinase is required for efficient end processing during DNA double-strand break repair. *Mol Cell Biol*, 23, 5836-48.
- Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M. & Gingeras, T. R. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, 29, 15-21.
- Dong, J., Ren, Y., Zhang, T., Wang, Z., Ling, C. C., Li, G. C., He, F., Wang, C. & Wen, B. 2018. Inactivation of DNA-PK by knockdown DNA-PKcs or NU7441 impairs non-homologous end-joining of radiation-induced double strand break repair. *Oncol Rep*, 39, 912-920.

- Douglas, P., Ye, R., Trinkle-Mulcahy, L., Neal, J. A., De Wever, V., Morrice, N. A., Meek, K. & Lees-Miller, S. P. 2014. Polo-like kinase 1 (PLK1) and protein phosphatase 6 (PP6) regulate DNA-dependent protein kinase catalytic subunit (DNA-PKcs) phosphorylation in mitosis. *Biosci Rep*, 34.
- Du, H. & Rosbash, M. 2002. The U1 snRNP protein U1C recognizes the 5' splice site in the absence of base pairing. *Nature*, 419, 86-90.
- Dvir, A., Peterson, S. R., Knuth, M. W., Lu, H. & Dynan, W. S. 1992. Ku autoantigen is the regulatory component of a template-associated protein kinase that phosphorylates RNA polymerase II. *Proc Natl Acad Sci U S A*, 89, 11920-4.
- Dylgjeri, E., Kothari, V., Shafi, A. A., Semenova, G., Gallagher, P. T., Guan, Y. F., Pang, A., Goodwin, J. F., Irani, S., McCann, J. J., Mandigo, A. C., Chand, S., McNair, C. M., Vasilevskaya, I., Schiewer, M. J., Lallas, C. D., McCue, P. A., Gomella, L. G., Seifert, E. L., Carroll, J. S., Butler, L. M., Holst, J., Kelly, W. K. & Knudsen, K. E. 2022. A Novel Role for DNA-PK in Metabolism by Regulating Glycolysis in Castration-Resistant Prostate Cancer. *Clin Cancer Res*, 28, 1446-1459.
- Dylgjeri, E., McNair, C., Goodwin, J. F., Raymon, H. K., McCue, P. A., Shafi, A. A., Leiby, B. E., de Leeuw, R., Kothari, V., McCann, J. J., Mandigo, A. C., Chand, S. N., Schiewer, M. J., Brand, L. J., Vasilevskaya, I., Gordon, N., Laufer, T. S., Gomella, L. G., Lallas, C. D., Trabulsi, E. J., Feng, F. Y., Filvaroff, E. H., Hege, K., Rathkopf, D. & Knudsen, K. E. 2019. Pleiotropic Impact of DNA-PK in Cancer and Implications for Therapeutic Strategies. *Clin Cancer Res*, 25, 5623-5637.
- Eide, T., Ramberg, H., Glackin, C., Tindall, D. & Tasken, K. A. 2013. TWIST1, A novel androgen-regulated gene, is a target for NKX3-1 in prostate cancer cells. *Cancer Cell Int*, 13, 4.
- Espejel, S., Franco, S., Sgura, A., Gae, D., Bailey, S. M., Taccioli, G. E. & Blasco, M. A. 2002. Functional interaction between DNA-PKcs and telomerase in telomere length maintenance. *EMBO J*, 21, 6275-87.
- Fan, L., Zhang, F., Xu, S., Cui, X., Hussain, A., Fazli, L., Gleave, M., Dong, X. & Qi, J. 2018. Histone demethylase JMJD1A promotes alternative splicing of AR variant 7 (AR-V7) in prostate cancer cells. *Proc Natl Acad Sci U S A*, 115, E4584-E4593.

- Feldman, B. J. & Feldman, D. 2001. The development of androgen-independent prostate cancer. *Nat Rev Cancer*, 1, 34-45.
- Ferraldeschi, R., Welte, J., Powers, M. V., Yuan, W., Smyth, T., Seed, G., Riisnaes, R., Hedayat, S., Wang, H., Crespo, M., Nava Rodrigues, D., Figueiredo, I., Miranda, S., Carreira, S., Lyons, J. F., Sharp, S., Plymate, S. R., Attard, G., Wallis, N., Workman, P. & de Bono, J. S. 2016. Second-Generation HSP90 Inhibitor Onalespib Blocks mRNA Splicing of Androgen Receptor Variant 7 in Prostate Cancer Cells. *Cancer Res*, 76, 2731-42.
- Fizazi, K., Shore, N., Tammela, T. L., Ulys, A., Vjaters, E., Polyakov, S., Jievaltas, M., Luz, M., Alekseev, B., Kuss, I., Kappeler, C., Snapir, A., Sarapohja, T., Smith, M. R. & Investigators, A. 2019. Darolutamide in Nonmetastatic, Castration-Resistant Prostate Cancer. *N Engl J Med*, 380, 1235-1246.
- Fok, J. H. L., Ramos-Montoya, A., Vazquez-Chantada, M., Wijnhoven, P. W. G., Follia, V., James, N., Farrington, P. M., Karmokar, A., Willis, S. E., Cairns, J., Nikkila, J., Beattie, D., Lamont, G. M., Finlay, M. R. V., Wilson, J., Smith, A., O'Connor, L. O., Ling, S., Fawell, S. E., O'Connor, M. J., Hollingsworth, S. J., Dean, E., Goldberg, F. W., Davies, B. R. & Cadogan, E. B. 2019. AZD7648 is a potent and selective DNA-PK inhibitor that enhances radiation, chemotherapy and olaparib activity. *Nat Commun*, 10, 5065.
- Gaddipati, J. P., McLeod, D. G., Heidenberg, H. B., Sesterhenn, I. A., Finger, M. J., Moul, J. W. & Srivastava, S. 1994. Frequent detection of codon 877 mutation in the androgen receptor gene in advanced prostate cancers. *Cancer Res*, 54, 2861-4.
- Gao, J., Aksoy, B. A., Dogrusoz, U., Dresdner, G., Gross, B., Sumer, S. O., Sun, Y., Jacobsen, A., Sinha, R., Larsson, E., Cerami, E., Sander, C. & Schultz, N. 2013. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal*, 6, p11.
- Gaughan, L., Stockley, J., Wang, N., McCracken, S. R., Treumann, A., Armstrong, K., Shaheen, F., Watt, K., McEwan, I. J., Wang, C., Pestell, R. G. & Robson, C. N. 2011. Regulation of the androgen receptor by SET9-mediated methylation. *Nucleic Acids Res*, 39, 1266-79.
- Geuens, T., Bouhy, D. & Timmerman, V. 2016. The hnRNP family: insights into their role in health and disease. *Hum Genet*, 135, 851-67.

- Giffin, W., Kwast-Welfeld, J., Rodda, D. J., Prefontaine, G. G., Traykova-Andonova, M., Zhang, Y., Weigel, N. L., Lefebvre, Y. A. & Hache, R. J. 1997. Sequence-specific DNA binding and transcription factor phosphorylation by Ku Autoantigen/DNA-dependent protein kinase. Phosphorylation of Ser-527 of the rat glucocorticoid receptor. *J Biol Chem*, 272, 5647-58.
- Gillis, J. L., Selth, L. A., Centenera, M. M., Townley, S. L., Sun, S., Plymate, S. R., Tilley, W. D. & Butler, L. M. 2013. Constitutively-active androgen receptor variants function independently of the HSP90 chaperone but do not confer resistance to HSP90 inhibitors. *Oncotarget*, 4, 691-704.
- Gleason, D. F. 1988. Histologic grade, clinical stage, and patient age in prostate cancer. *NCI Monogr*, 15-8.
- Gong, Y., Wang, D., Dar, J. A., Singh, P., Graham, L., Liu, W., Ai, J., Xin, Z., Guo, Y. & Wang, Z. 2012. Nuclear export signal of androgen receptor (NESAR) regulation of androgen receptor level in human prostate cell lines via ubiquitination and proteasome-dependent degradation. *Endocrinology*, 153, 5716-25.
- Goodwin, J. F. & Knudsen, K. E. 2014. Beyond DNA repair: DNA-PK function in cancer. *Cancer Discov*, 4, 1126-39.
- Goodwin, J. F., Kothari, V., Drake, J. M., Zhao, S., Dylgjeri, E., Dean, J. L., Schiewer, M. J., McNair, C., Jones, J. K., Aytes, A., Magee, M. S., Snook, A. E., Zhu, Z., Den, R. B., Birbe, R. C., Gomella, L. G., Graham, N. A., Vashisht, A. A., Wohlschlegel, J. A., Graeber, T. G., Karnes, R. J., Takhar, M., Davicioni, E., Tomlins, S. A., Abate-Shen, C., Sharifi, N., Witte, O. N., Feng, F. Y. & Knudsen, K. E. 2015. DNA-PKcs-Mediated Transcriptional Regulation Drives Prostate Cancer Progression and Metastasis. *Cancer Cell*, 28, 97-113.
- Goodwin, J. F., Schiewer, M. J., Dean, J. L., Schrecengost, R. S., de Leeuw, R., Han, S., Ma, T., Den, R. B., Dicker, A. P., Feng, F. Y. & Knudsen, K. E. 2013. A hormone-DNA repair circuit governs the response to genotoxic insult. *Cancer Discov*, 3, 1254-71.
- Gordon, V., Bhadel, S., Wunderlich, W., Zhang, J., Ficarro, S. B., Mollah, S. A., Shabanowitz, J., Hunt, D. F., Xenarios, I., Hahn, W. C., Conaway, M., Carey, M. F. & Gioeli, D. 2010. CDK9

- regulates AR promoter selectivity and cell growth through serine 81 phosphorylation. *Mol Endocrinol*, 24, 2267-80.
- Grasso, C. S., Wu, Y. M., Robinson, D. R., Cao, X., Dhanasekaran, S. M., Khan, A. P., Quist, M. J., Jing, X., Lonigro, R. J., Brenner, J. C., Asangani, I. A., Ateeq, B., Chun, S. Y., Siddiqui, J., Sam, L., Anstett, M., Mehra, R., Prensner, J. R., Palanisamy, N., Ryslik, G. A., Vandin, F., Raphael, B. J., Kunju, L. P., Rhodes, D. R., Pienta, K. J., Chinnaiyan, A. M. & Tomlins, S. A. 2012. The mutational landscape of lethal castration-resistant prostate cancer. *Nature*, 487, 239-43.
- Gu, B., Posfai, E. & Rossant, J. 2018. Efficient generation of targeted large insertions by microinjection into two-cell-stage mouse embryos. *Nat Biotechnol*, 36, 632-637.
- Guo, Z., Yang, X., Sun, F., Jiang, R., Linn, D. E., Chen, H., Chen, H., Kong, X., Melamed, J., Tepper, C. G., Kung, H. J., Brodie, A. M., Edwards, J. & Qiu, Y. 2009. A novel androgen receptor splice variant is up-regulated during prostate cancer progression and promotes androgen depletion-resistant growth. *Cancer Res*, 69, 2305-13.
- Gupta, R., Somyajit, K., Narita, T., Maskey, E., Stanlie, A., Kremer, M., Typas, D., Lammers, M., Mailand, N., Nussenzweig, A., Lukas, J. & Choudhary, C. 2018. DNA Repair Network Analysis Reveals Shieldin as a Key Regulator of NHEJ and PARP Inhibitor Sensitivity. *Cell*, 173, 972-988 e23.
- Gupta, S., Li, J., Kemeny, G., Bitting, R. L., Beaver, J., Somarelli, J. A., Ware, K. E., Gregory, S. & Armstrong, A. J. 2017. Whole Genomic Copy Number Alterations in Circulating Tumor Cells from Men with Abiraterone or Enzalutamide-Resistant Metastatic Castration-Resistant Prostate Cancer. *Clin Cancer Res*, 23, 1346-1357.
- Haelens, A., Tanner, T., Denayer, S., Callewaert, L. & Claessens, F. 2007. The hinge region regulates DNA binding, nuclear translocation, and transactivation of the androgen receptor. *Cancer Res*, 67, 4514-23.
- Haile, S., Lal, A., Myung, J. K. & Sadar, M. D. 2011. FUS/TLS is a co-activator of androgen receptor in prostate cancer cells. *PLoS One*, 6, e24197.

- Han, S., Udeshi, N. D., Deerinck, T. J., Svinkina, T., Ellisman, M. H., Carr, S. A. & Ting, A. Y. 2017. Proximity Biotinylation as a Method for Mapping Proteins Associated with mtDNA in Living Cells. *Cell Chem Biol*, 24, 404-414.
- Hara, T., Miyazaki, J., Araki, H., Yamaoka, M., Kanzaki, N., Kusaka, M. & Miyamoto, M. 2003. Novel mutations of androgen receptor: a possible mechanism of bicalutamide withdrawal syndrome. *Cancer Res*, 63, 149-53.
- Hartley, K. O., Gell, D., Smith, G. C., Zhang, H., Divecha, N., Connelly, M. A., Admon, A., Lees-Miller, S. P., Anderson, C. W. & Jackson, S. P. 1995. DNA-dependent protein kinase catalytic subunit: a relative of phosphatidylinositol 3-kinase and the ataxia telangiectasia gene product. *Cell*, 82, 849-56.
- He, S., Zhang, C., Shafi, A. A., Sequeira, M., Acquaviva, J., Friedland, J. C., Sang, J., Smith, D. L., Weigel, N. L., Wada, Y. & Proia, D. A. 2013. Potent activity of the Hsp90 inhibitor ganetespib in prostate cancer cells irrespective of androgen receptor status or variant receptor expression. *Int J Oncol*, 42, 35-43.
- He, Y., Lu, J., Ye, Z., Hao, S., Wang, L., Kohli, M., Tindall, D. J., Li, B., Zhu, R., Wang, L. & Huang, H. 2018. Androgen receptor splice variants bind to constitutively open chromatin and promote abiraterone-resistant growth of prostate cancer. *Nucleic Acids Res*, 46, 1895-1911.
- Heery, D. M., Kalkhoven, E., Hoare, S. & Parker, M. G. 1997. A signature motif in transcriptional co-activators mediates binding to nuclear receptors. *Nature*, 387, 733-6.
- Heinrich, B., Zhang, Z., Raitskin, O., Hiller, M., Benderska, N., Hartmann, A. M., Bracco, L., Elliott, D., Ben-Ari, S., Soreq, H., Sperling, J., Sperling, R. & Stamm, S. 2009. Heterogeneous nuclear ribonucleoprotein G regulates splice site selection by binding to CC(A/C)-rich regions in pre-mRNA. *J Biol Chem*, 284, 14303-15.
- Henzler, C., Li, Y., Yang, R., McBride, T., Ho, Y., Sprenger, C., Liu, G., Coleman, I., Lakely, B., Li, R., Ma, S., Landman, S. R., Kumar, V., Hwang, T. H., Raj, G. V., Higano, C. S., Morrissey, C., Nelson, P. S., Plymate, S. R. & Dehm, S. M. 2016. Truncation and constitutive activation of the androgen receptor by diverse genomic rearrangements in prostate cancer. *Nat Commun*, 7, 13668.

- Hickson, I., Zhao, Y., Richardson, C. J., Green, S. J., Martin, N. M., Orr, A. I., Reaper, P. M., Jackson, S. P., Curtin, N. J. & Smith, G. C. 2004. Identification and characterization of a novel and specific inhibitor of the ataxia-telangiectasia mutated kinase ATM. *Cancer Res*, 64, 9152-9.
- Hosoi, Y., Watanabe, T., Nakagawa, K., Matsumoto, Y., Enomoto, A., Morita, A., Nagawa, H. & Suzuki, N. 2004. Up-regulation of DNA-dependent protein kinase activity and Sp1 in colorectal cancer. *Int J Oncol*, 25, 461-8.
- Hsu, F. N., Chen, M. C., Chiang, M. C., Lin, E., Lee, Y. T., Huang, P. H., Lee, G. S. & Lin, H. 2011. Regulation of androgen receptor and prostate cancer growth by cyclin-dependent kinase 5. *J Biol Chem*, 286, 33141-9.
- Hu, R., Dunn, T. A., Wei, S., Isharwal, S., Veltri, R. W., Humphreys, E., Han, M., Partin, A. W., Vessella, R. L., Isaacs, W. B., Bova, G. S. & Luo, J. 2009. Ligand-independent androgen receptor variants derived from splicing of cryptic exons signify hormone-refractory prostate cancer. *Cancer Res*, 69, 16-22.
- Hu, R., Lu, C., Mostaghel, E. A., Yegnasubramanian, S., Gurel, M., Tannahill, C., Edwards, J., Isaacs, W. B., Nelson, P. S., Bluemn, E., Plymate, S. R. & Luo, J. 2012. Distinct transcriptional programs mediated by the ligand-dependent full-length androgen receptor and its splice variants in castration-resistant prostate cancer. *Cancer Res*, 72, 3457-62.
- Huggins, C. & Hodges, C. V. 1972. Studies on prostatic cancer. I. The effect of castration, of estrogen and androgen injection on serum phosphatases in metastatic carcinoma of the prostate. *CA Cancer J Clin*, 22, 232-40.
- Hung, V., Udeshi, N. D., Lam, S. S., Loh, K. H., Cox, K. J., Pedram, K., Carr, S. A. & Ting, A. Y. 2016. Spatially resolved proteomic mapping in living cells with the engineered peroxidase APEX2. *Nat Protoc*, 11, 456-75.
- Hussain, M., Fizazi, K., Saad, F., Rathenborg, P., Shore, N., Ferreira, U., Ivashchenko, P., Demirhan, E., Modelski, K., Phung, K., Krivosik, A. & Sternberg, C. N. 2018. Enzalutamide in Men with Nonmetastatic, Castration-Resistant Prostate Cancer. *N Engl J Med*, 378, 2465-2474.

- Hussain, M., Mateo, J., Fizazi, K., Saad, F., Shore, N., Sandhu, S., Chi, K. N., Sartor, O., Agarwal, N., Olmos, D., Thiery-Vuillemin, A., Twardowski, P., Roubaud, G., Ozguroglu, M., Kang, J., Burgents, J., Gresty, C., Corcoran, C., Adelman, C. A., de Bono, J. & Investigators, P. R. T. 2020. Survival with Olaparib in Metastatic Castration-Resistant Prostate Cancer. *N Engl J Med*, 383, 2345-2357.
- Iijima, S., Teraoka, H., Date, T. & Tsukada, K. 1992. DNA-activated protein kinase in Raji Burkitt's lymphoma cells. Phosphorylation of c-Myc oncoprotein. *Eur J Biochem*, 206, 595-603.
- Jackson, S. P., MacDonald, J. J., Lees-Miller, S. & Tjian, R. 1990. GC box binding induces phosphorylation of Sp1 by a DNA-dependent protein kinase. *Cell*, 63, 155-65.
- Jenster, G., van der Korput, H. A., Trapman, J. & Brinkmann, A. O. 1995. Identification of two transcription activation units in the N-terminal domain of the human androgen receptor. *J Biol Chem*, 270, 7341-6.
- Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A. & Charpentier, E. 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*, 337, 816-21.
- Jinek, M., East, A., Cheng, A., Lin, S., Ma, E. & Doudna, J. 2013. RNA-programmed genome editing in human cells. *Elife*, 2, e00471.
- Jones, D., Noble, M., Wedge, S. R., Robson, C. N. & Gaughan, L. 2017. Aurora A regulates expression of AR-V7 in models of castrate resistant prostate cancer. *Sci Rep*, 7, 40957.
- Jones, D., Wade, M., Nakjang, S., Chaytor, L., Grey, J., Robson, C. N. & Gaughan, L. 2015. FOXA1 regulates androgen receptor variant activity in models of castrate-resistant prostate cancer. *Oncotarget*, 6, 29782-94.
- Kang, Y. & Massague, J. 2004. Epithelial-mesenchymal transitions: twist in development and metastasis. *Cell*, 118, 277-9.
- Karantanos, T., Corn, P. G. & Thompson, T. C. 2013. Prostate cancer progression after androgen deprivation therapy: mechanisms of castrate resistance and novel therapeutic approaches. *Oncogene*, 32, 5501-11.

- Kicinski, M., Vangronsveld, J. & Nawrot, T. S. 2011. An epidemiological reappraisal of the familial aggregation of prostate cancer: a meta-analysis. *PLoS One*, 6, e27130.
- Kim, D. I., Birendra, K. C., Zhu, W., Motamedchaboki, K., Doye, V. & Roux, K. J. 2014. Probing nuclear pore complex architecture with proximity-dependent biotinylation. *Proc Natl Acad Sci U S A*, 111, E2453-61.
- Kirby, M., Hirst, C. & Crawford, E. D. 2011. Characterising the castration-resistant prostate cancer population: a systematic review. *Int J Clin Pract*, 65, 1180-92.
- Kivinummi, K., Urbanucci, A., Leinonen, K., Tammela, T. L. J., Annala, M., Isaacs, W. B., Bova, G. S., Nykter, M. & Visakorpi, T. 2017. The expression of AURKA is androgen regulated in castration-resistant prostate cancer. *Sci Rep*, 7, 17978.
- Klotz, L., Boccon-Gibod, L., Shore, N. D., Andreou, C., Persson, B. E., Cantor, P., Jensen, J. K., Olesen, T. K. & Schroder, F. H. 2008. The efficacy and safety of degarelix: a 12-month, comparative, randomized, open-label, parallel-group phase III study in patients with prostate cancer. *BJU Int*, 102, 1531-8.
- Klotz, L., Zhang, L., Lam, A., Nam, R., Mamedov, A. & Loblaw, A. 2010. Clinical results of long-term follow-up of a large, active surveillance cohort with localized prostate cancer. *J Clin Oncol*, 28, 126-31.
- Kote-Jarai, Z., Leongamornlert, D., Saunders, E., Tymrakiewicz, M., Castro, E., Mahmud, N., Guy, M., Edwards, S., O'Brien, L., Sawyer, E., Hall, A., Wilkinson, R., Dadaev, T., Goh, C., Easton, D., Collaborators, U., Goldgar, D. & Eeles, R. 2011. BRCA2 is a moderate penetrance gene contributing to young-onset prostate cancer: implications for genetic testing in prostate cancer patients. *Br J Cancer*, 105, 1230-4.
- Kothari, V., Goodwin, J. F., Zhao, S. G., Drake, J. M., Yin, Y., Chang, S. L., Evans, J. R., Wilder-Romans, K., Gabbara, K., Dylgjeri, E., Chou, J., Sun, G., Tomlins, S. A., Mehra, R., Hege, K., Filvaroff, E. H., Schaeffer, E. M., Karnes, R. J., Quigley, D. A., Rathkopf, D. E., He, H. H., Speers, C., Spratt, D. E., Gilbert, L. A., Ashworth, A., Chinnaiyan, A. M., Raj, G. V., Knudsen, K. E. & Feng, F. Y. 2019. DNA-Dependent Protein Kinase Drives Prostate Cancer Progression through Transcriptional Regulation of the Wnt Signaling Pathway. *Clin Cancer Res*, 25, 5608-5622.

- Kounatidou, E., Nakjang, S., McCracken, S. R. C., Dehm, S. M., Robson, C. N., Jones, D. & Gaughan, L. 2019. A novel CRISPR-engineered prostate cancer cell line defines the AR-V transcriptome and identifies PARP inhibitor sensitivities. *Nucleic Acids Res.*
- Kwok, W. K., Ling, M. T., Lee, T. W., Lau, T. C., Zhou, C., Zhang, X., Chua, C. W., Chan, K. W., Chan, F. L., Glackin, C., Wong, Y. C. & Wang, X. 2005. Up-regulation of TWIST in prostate cancer and its implication as a therapeutic target. *Cancer Res*, 65, 5153-62.
- Laczko, I., Hudson, D. L., Freeman, A., Feneley, M. R. & Masters, J. R. 2005. Comparison of the zones of the human prostate with the seminal vesicle: morphology, immunohistochemistry, and cell kinetics. *Prostate*, 62, 260-6.
- Lang, S. H., Swift, S. L., White, H., Misso, K., Kleijnen, J. & Quek, R. G. W. 2019. A systematic review of the prevalence of DNA damage response gene mutations in prostate cancer. *Int J Oncol*, 55, 597-616.
- Lapointe, J., Li, C., Higgins, J. P., van de Rijn, M., Bair, E., Montgomery, K., Ferrari, M., Egevad, L., Rayford, W., Bergerheim, U., Ekman, P., DeMarzo, A. M., Tibshirani, R., Botstein, D., Brown, P. O., Brooks, J. D. & Pollack, J. R. 2004. Gene expression profiling identifies clinically relevant subtypes of prostate cancer. *Proc Natl Acad Sci U S A*, 101, 811-6.
- Leahy, J. J., Golding, B. T., Griffin, R. J., Hardcastle, I. R., Richardson, C., Rigoreau, L. & Smith, G. C. 2004. Identification of a highly potent and selective DNA-dependent protein kinase (DNA-PK) inhibitor (NU7441) by screening of chromenone libraries. *Bioorg Med Chem Lett*, 14, 6083-7.
- Leclerc, D., Pham, D. N., Levesque, N., Truongcao, M., Foulkes, W. D., Sapienza, C. & Rozen, R. 2017. Oncogenic role of PDK4 in human colon cancer cells. *Br J Cancer*, 116, 930-936.
- Lee, K. J., Lin, Y. F., Chou, H. Y., Yajima, H., Fattah, K. R., Lee, S. C. & Chen, B. P. 2011. Involvement of DNA-dependent protein kinase in normal cell cycle progression through mitosis. *J Biol Chem*, 286, 12796-802.
- Lempiainen, J. K., Niskanen, E. A., Vuoti, K. M., Lampinen, R. E., Goos, H., Varjosalo, M. & Palvimo, J. J. 2017. Agonist-specific Protein Interactomes of Glucocorticoid and Androgen Receptor as Revealed by Proximity Mapping. *Mol Cell Proteomics*, 16, 1462-1474.

- Leuprolide Study, G. 1984. Leuprolide versus diethylstilbestrol for metastatic prostate cancer. *N Engl J Med*, 311, 1281-6.
- Li, H., Wang, Z., Tang, K., Zhou, H., Liu, H., Yan, L., Guan, W., Chen, K., Xu, H. & Ye, Z. 2018. Prognostic Value of Androgen Receptor Splice Variant 7 in the Treatment of Castration-resistant Prostate Cancer with Next generation Androgen Receptor Signal Inhibition: A Systematic Review and Meta-analysis. *Eur Urol Focus*, 4, 529-539.
- Li, L., Karanika, S., Yang, G., Wang, J., Park, S., Broom, B. M., Manyam, G. C., Wu, W., Luo, Y., Basourakos, S., Song, J. H., Gallick, G. E., Karantanos, T., Korentzelos, D., Azad, A. K., Kim, J., Corn, P. G., Aparicio, A. M., Logothetis, C. J., Troncoso, P., Heffernan, T., Toniatti, C., Lee, H. S., Lee, J. S., Zuo, X., Chang, W., Yin, J. & Thompson, T. C. 2017. Androgen receptor inhibitor-induced "BRCAness" and PARP inhibition are synthetically lethal for castration-resistant prostate cancer. *Sci Signal*, 10.
- Li, P., Nicosia, S. V. & Bai, W. 2001. Antagonism between PTEN/MMAC1/TEP-1 and androgen receptor in growth and apoptosis of prostatic cancer cells. *J Biol Chem*, 276, 20444-50.
- Li, Y., Alsagabi, M., Fan, D., Bova, G. S., Tewfik, A. H. & Dehm, S. M. 2011. Intragenic rearrangement and altered RNA splicing of the androgen receptor in a cell-based model of prostate cancer progression. *Cancer Res*, 71, 2108-17.
- Li, Y., Chan, S. C., Brand, L. J., Hwang, T. H., Silverstein, K. A. & Dehm, S. M. 2013. Androgen receptor splice variants mediate enzalutamide resistance in castration-resistant prostate cancer cell lines. *Cancer Res*, 73, 483-9.
- Li, Y., Hwang, T. H., Oseth, L. A., Hauge, A., Vessella, R. L., Schmechel, S. C., Hirsch, B., Beckman, K. B., Silverstein, K. A. & Dehm, S. M. 2012. AR intragenic deletions linked to androgen receptor splice variant expression and activity in models of prostate cancer progression. *Oncogene*, 31, 4759-67.
- Li, Y., Yang, R., Henzler, C. M., Ho, Y., Passow, C., Auch, B., Carreira, S., Nava Rodrigues, D., Bertan, C., Hwang, T. H., Quigley, D. A., Dang, H. X., Morrissey, C., Fraser, M., Plymate, S. R., Maher, C. A., Feng, F. Y., de Bono, J. S. & Dehm, S. M. 2020. Diverse AR Gene Rearrangements Mediate Resistance to Androgen Receptor Inhibitors in Metastatic Prostate Cancer. *Clin Cancer Res*, 26, 1965-1976.

- Libertini, S. J., Tepper, C. G., Rodriguez, V., Asmuth, D. M., Kung, H. J. & Mudryj, M. 2007. Evidence for calpain-mediated androgen receptor cleavage as a mechanism for androgen independence. *Cancer Res*, 67, 9001-5.
- Lin, H. K., Yeh, S., Kang, H. Y. & Chang, C. 2001. Akt suppresses androgen-induced apoptosis by phosphorylating and inhibiting androgen receptor. *Proc Natl Acad Sci U S A*, 98, 7200-5.
- Liu, C., Lou, W., Armstrong, C., Zhu, Y., Evans, C. P. & Gao, A. C. 2015. Niclosamide suppresses cell migration and invasion in enzalutamide resistant prostate cancer cells via Stat3-AR axis inhibition. *Prostate*, 75, 1341-53.
- Liu, L., Chen, X., Li, J., Wang, H., Buehl, C. J., Goff, N. J., Meek, K., Yang, W. & Gellert, M. 2022. Autophosphorylation transforms DNA-PK from protecting to processing DNA ends. *Mol Cell*, 82, 177-189 e4.
- Liu, L. L., Xie, N., Sun, S., Plymate, S., Mostaghel, E. & Dong, X. 2014a. Mechanisms of the androgen receptor splicing in prostate cancer cells. *Oncogene*, 33, 3140-50.
- Liu, N., Zhou, K. I., Parisien, M., Dai, Q., Diatchenko, L. & Pan, T. 2017. N6-methyladenosine alters RNA structure to regulate binding of a low-complexity protein. *Nucleic Acids Res*, 45, 6051-6063.
- Liu, Q., Wang, G., Li, Q., Jiang, W., Kim, J. S., Wang, R., Zhu, S., Wang, X., Yan, L., Yi, Y., Zhang, L., Meng, Q., Li, C., Zhao, D., Qiao, Y., Li, Y., Gursel, D. B., Chinnaiyan, A. M., Chen, K. & Cao, Q. 2019. Polycomb group proteins EZH2 and EED directly regulate androgen receptor in advanced prostate cancer. *Int J Cancer*, 145, 415-426.
- Liu, X. S., Wu, H., Ji, X., Stelzer, Y., Wu, X., Czauderna, S., Shu, J., Dadon, D., Young, R. A. & Jaenisch, R. 2016. Editing DNA Methylation in the Mammalian Genome. *Cell*, 167, 233-247 e17.
- Liu, Z., Merkurjev, D., Yang, F., Li, W., Oh, S., Friedman, M. J., Song, X., Zhang, F., Ma, Q., Ohgi, K. A., Krones, A. & Rosenfeld, M. G. 2014b. Enhancer activation requires trans-recruitment of a mega transcription factor complex. *Cell*, 159, 358-73.
- Lonergan, P. E. & Tindall, D. J. 2011. Androgen receptor signaling in prostate cancer development and progression. *J Carcinog*, 10, 20.

- Love, M. I., Huber, W. & Anders, S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*, 15, 550.
- Lu, S. X., De Neef, E., Thomas, J. D., Sabio, E., Rousseau, B., Gigoux, M., Knorr, D. A., Greenbaum, B., Elhanati, Y., Hogg, S. J., Chow, A., Ghosh, A., Xie, A., Zamarin, D., Cui, D., Erickson, C., Singer, M., Cho, H., Wang, E., Lu, B., Durham, B. H., Shah, H., Chowell, D., Gabel, A. M., Shen, Y., Liu, J., Jin, J., Rhodes, M. C., Taylor, R. E., Molina, H., Wolchok, J. D., Merghoub, T., Diaz, L. A., Jr., Abdel-Wahab, O. & Bradley, R. K. 2021. Pharmacologic modulation of RNA splicing enhances anti-tumor immunity. *Cell*, 184, 4032-4047 e31.
- Luo, H., Liu, Y., Li, Y., Zhang, C., Yu, B. & Shao, C. 2022. Androgen receptor splicing variant 7 (ARv7) promotes DNA damage response in prostate cancer cells. *FASEB J*, 36, e22495.
- Makridakis, N. M., di Salle, E. & Reichardt, J. K. 2000. Biochemical and pharmacogenetic dissection of human steroid 5 alpha-reductase type II. *Pharmacogenetics*, 10, 407-13.
- Makridakis, N. R., RK; Pike, MC; Chang, L; Stanczyk, FZ; Kolonel, LN; Shi, CY; Yu, MC; Henderson, BE; Reichardt, JK 1997. A prevalent missense substitution that modulates activity of prostatic steroid 5alpha-reductase. *Cancer Res*, 6, 1020-22.
- Martin, T. A., Goyal, A., Watkins, G. & Jiang, W. G. 2005. Expression of the transcription factors snail, slug, and twist and their clinical significance in human breast cancer. *Ann Surg Oncol*, 12, 488-96.
- Mateo, J., Carreira, S., Sandhu, S., Miranda, S., Mossop, H., Perez-Lopez, R., Nava Rodrigues, D., Robinson, D., Omlin, A., Tunariu, N., Boysen, G., Porta, N., Flohr, P., Gillman, A., Figueiredo, I., Paulding, C., Seed, G., Jain, S., Ralph, C., Protheroe, A., Hussain, S., Jones, R., Elliott, T., McGovern, U., Bianchini, D., Goodall, J., Zafeiriou, Z., Williamson, C. T., Ferraldeschi, R., Riisnaes, R., Ebbs, B., Fowler, G., Roda, D., Yuan, W., Wu, Y. M., Cao, X., Brough, R., Pemberton, H., A'Hern, R., Swain, A., Kunju, L. P., Eeles, R., Attard, G., Lord, C. J., Ashworth, A., Rubin, M. A., Knudsen, K. E., Feng, F. Y., Chinnaiyan, A. M., Hall, E. & de Bono, J. S. 2015. DNA-Repair Defects and Olaparib in Metastatic Prostate Cancer. *N Engl J Med*, 373, 1697-708.

- Mayeur, G. L., Kung, W. J., Martinez, A., Izumiya, C., Chen, D. J. & Kung, H. J. 2005. Ku is a novel transcriptional recycling coactivator of the androgen receptor in prostate cancer cells. *J Biol Chem*, 280, 10827-33.
- Medunjanin, S., Weinert, S., Schmeisser, A., Mayer, D. & Braun-Dullaeus, R. C. 2010. Interaction of the double-strand break repair kinase DNA-PK and estrogen receptor-alpha. *Mol Biol Cell*, 21, 1620-8.
- Menolfi, D. & Zha, S. 2020. ATM, ATR and DNA-PKcs kinases-the lessons from the mouse models: inhibition not equal deletion. *Cell Biosci*, 10, 8.
- Moilanen, A. M., Riikonen, R., Oksala, R., Ravanti, L., Aho, E., Wohlfahrt, G., Nykanen, P. S., Tormakangas, O. P., Palvimo, J. J. & Kallio, P. J. 2015. Discovery of ODM-201, a new-generation androgen receptor inhibitor targeting resistance mechanisms to androgen signaling-directed prostate cancer therapies. *Sci Rep*, 5, 12007.
- Mojica, F. J. M., Diez-Villasenor, C., Garcia-Martinez, J. & Almendros, C. 2009. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology (Reading)*, 155, 733-740.
- Mootha, V. K., Lindgren, C. M., Eriksson, K. F., Subramanian, A., Sihag, S., Lehar, J., Puigserver, P., Carlsson, E., Ridderstrale, M., Laurila, E., Houstis, N., Daly, M. J., Patterson, N., Mesirov, J. P., Golub, T. R., Tamayo, P., Spiegelman, B., Lander, E. S., Hirschhorn, J. N., Altshuler, D. & Groop, L. C. 2003. PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat Genet*, 34, 267-73.
- Morel, A. P., Hinkal, G. W., Thomas, C., Fauvet, F., Courtois-Cox, S., Wierinckx, A., Devouassoux-Shisheboran, M., Treilleux, I., Tissier, A., Gras, B., Pourchet, J., Puisieux, I., Browne, G. J., Spicer, D. B., Lachuer, J., Ansieau, S. & Puisieux, A. 2012. EMT inducers catalyze malignant transformation of mammary epithelial cells and drive tumorigenesis towards claudin-low tumors in transgenic mice. *PLoS Genet*, 8, e1002723.
- Morel, K. L., Sheahan, A. V., Burkhart, D. L., Baca, S. C., Boufaied, N., Liu, Y., Qiu, X., Canadas, I., Roehle, K., Heckler, M., Calagua, C., Ye, H., Pantelidou, C., Galbo, P., Panja, S.,

- Mitrofanova, A., Wilkinson, S., Whitlock, N. C., Trostel, S. Y., Hamid, A. A., Kibel, A. S., Barbie, D. A., Choudhury, A. D., Pomerantz, M. M., Sweeney, C. J., Long, H. W., Einstein, D. J., Shapiro, G. I., Dougan, S. K., Sowalsky, A. G., He, H. H., Freedman, M. L., Balk, S. P., Loda, M., Labbe, D. P., Olson, B. M. & Ellis, L. 2021. EZH2 inhibition activates a dsRNA-STING-interferon stress axis that potentiates response to PD-1 checkpoint blockade in prostate cancer. *Nat Cancer*, 2, 444-456.
- Munster, P., Mita, M., Mahipal, A., Nemunaitis, J., Massard, C., Mikkelsen, T., Cruz, C., Paz-Ares, L., Hidalgo, M., Rathkopf, D., Blumenschein, G., Jr., Smith, D. C., Eichhorst, B., Cloughesy, T., Filvaroff, E. H., Li, S., Raymon, H., de Haan, H., Hege, K. & Bendell, J. C. 2019. First-In-Human Phase I Study Of A Dual mTOR Kinase And DNA-PK Inhibitor (CC-115) In Advanced Malignancy. *Cancer Manag Res*, 11, 10463-10476.
- Myung, J. K., Banuelos, C. A., Fernandez, J. G., Mawji, N. R., Wang, J., Tien, A. H., Yang, Y. C., Tavakoli, I., Haile, S., Watt, K., McEwan, I. J., Plymate, S., Andersen, R. J. & Sadar, M. D. 2013. An androgen receptor N-terminal domain antagonist for treating prostate cancer. *J Clin Invest*, 123, 2948-60.
- Nakamura, K., Karmokar, A., Farrington, P. M., James, N. H., Ramos-Montoya, A., Bickerton, S. J., Hughes, G. D., Illidge, T. M., Cadogan, E. B., Davies, B. R., Dovedi, S. J. & Valge-Archer, V. 2021. Inhibition of DNA-PK with AZD7648 Sensitizes Tumor Cells to Radiotherapy and Induces Type I IFN-Dependent Durable Tumor Control. *Clin Cancer Res*, 27, 4353-4366.
- Oberhuber, M., Pecoraro, M., Ruzs, M., Oberhuber, G., Wieselberg, M., Haslinger, P., Gurnhofer, E., Schleder, M., Limberger, T., Lagger, S., Pencik, J., Kodajova, P., Hogler, S., Stockmaier, G., Grund-Groschke, S., Aberger, F., Bolis, M., Theurillat, J. P., Wiebringhaus, R., Weiss, T., Haitel, A., Brehme, M., Wadsak, W., Griss, J., Mohr, T., Hofer, A., Jager, A., Pollheimer, J., Egger, G., Koellensperger, G., Mann, M., Hantusch, B. & Kenner, L. 2020. STAT3-dependent analysis reveals PDK4 as independent predictor of recurrence in prostate cancer. *Mol Syst Biol*, 16, e9247.
- Orme, J. J., Pagliaro, L. C., Quevedo, J. F., Park, S. S. & Costello, B. A. 2022. Rational Second-Generation Antiandrogen Use in Prostate Cancer. *Oncologist*, 27, 110-124.

- Paix, A., Folkmann, A., Goldman, D. H., Kulaga, H., Grzelak, M. J., Rasoloson, D., Paidemarry, S., Green, R., Reed, R. R. & Seydoux, G. 2017. Precision genome editing using synthesis-dependent repair of Cas9-induced DNA breaks. *Proc Natl Acad Sci U S A*, 114, E10745-E10754.
- Pal, S. K., Moreira, D., Won, H., White, S. W., Dutttagupta, P., Lucia, M., Jones, J., Hsu, J. & Kortylewski, M. 2019. Reduced T-cell Numbers and Elevated Levels of Immunomodulatory Cytokines in Metastatic Prostate Cancer Patients De Novo Resistant to Abiraterone and/or Enzalutamide Therapy. *Int J Mol Sci*, 20.
- Paltoglou, S., Das, R., Townley, S. L., Hickey, T. E., Tarulli, G. A., Coutinho, I., Fernandes, R., Hanson, A. R., Denis, I., Carroll, J. S., Dehm, S. M., Raj, G. V., Plymate, S. R., Tilley, W. D. & Selth, L. A. 2017. Novel Androgen Receptor Coregulator GRHL2 Exerts Both Oncogenic and Antimetastatic Functions in Prostate Cancer. *Cancer Res*, 77, 3417-3430.
- Park, H. K., Lim, S. D. & Kwon, G. Y. 2019. mRNA expressions of androgen receptor and its variants in matched hormone-sensitive and castration-resistant prostate cancer. *Scand J Urol*, 53, 365-371.
- Parker, C. C., James, N. D., Brawley, C. D., Clarke, N. W., Hoyle, A. P., Ali, A., Ritchie, A. W. S., Attard, G., Chowdhury, S., Cross, W., Dearnaley, D. P., Gillessen, S., Gilson, C., Jones, R. J., Langley, R. E., Malik, Z. I., Mason, M. D., Matheson, D., Millman, R., Russell, J. M., Thalmann, G. N., Amos, C. L., Alonzi, R., Bahl, A., Birtle, A., Din, O., Douis, H., Eswar, C., Gale, J., Gannon, M. R., Jonnada, S., Khaksar, S., Lester, J. F., O'Sullivan, J. M., Parikh, O. A., Pedley, I. D., Pudney, D. M., Sheehan, D. J., Srihari, N. N., Tran, A. T. H., Parmar, M. K. B., Sydes, M. R. & Systemic Therapy for Advanced or Metastatic Prostate cancer: Evaluation of Drug Efficacy, i. 2018. Radiotherapy to the primary tumour for newly diagnosed, metastatic prostate cancer (STAMPEDE): a randomised controlled phase 3 trial. *Lancet*, 392, 2353-2366.
- Partin, A. W., Kattan, M. W., Subong, E. N., Walsh, P. C., Wojno, K. J., Oesterling, J. E., Scardino, P. T. & Pearson, J. D. 1997. Combination of prostate-specific antigen, clinical stage, and Gleason score to predict pathological stage of localized prostate cancer. A multi-institutional update. *JAMA*, 277, 1445-51.

- Paschalis, A., Sharp, A., Welti, J. C., Neeb, A., Raj, G. V., Luo, J., Plymate, S. R. & de Bono, J. S. 2018. Alternative splicing in prostate cancer. *Nat Rev Clin Oncol*, 15, 663-675.
- Paschalis, A., Welti, J., Neeb, A. J., Yuan, W., Figueiredo, I., Pereira, R., Ferreira, A., Riisnaes, R., Rodrigues, D. N., Jimenez-Vacas, J. M., Kim, S., Uo, T., Micco, P. D., Tumber, A., Islam, M. S., Moesser, M. A., Abboud, M., Kawamura, A., Gurel, B., Christova, R., Gil, V. S., Buroni, L., Crespo, M., Miranda, S., Lambros, M. B., Carreira, S., Tunariu, N., Alimonti, A., Al-Lazikani, B., Schofield, C. J., Plymate, S. R., Sharp, A., de Bono, J. S. & Team, S. U. C. P. C. F. I. P. C. D. 2021. JMJD6 Is a Druggable Oxygenase That Regulates AR-V7 Expression in Prostate Cancer. *Cancer Res*, 81, 1087-1100.
- Pineault, K. M., Novoa, A., Lozovska, A., Wellik, D. M. & Mallo, M. 2019. Two CRISPR/Cas9-mediated methods for targeting complex insertions, deletions, or replacements in mouse. *MethodsX*, 6, 2088-2100.
- Plummer, R., Jones, C., Middleton, M., Wilson, R., Evans, J., Olsen, A., Curtin, N., Boddy, A., McHugh, P., Newell, D., Harris, A., Johnson, P., Steinfeldt, H., Dewji, R., Wang, D., Robson, L. & Calvert, H. 2008. Phase I study of the poly(ADP-ribose) polymerase inhibitor, AG014699, in combination with temozolomide in patients with advanced solid tumors. *Clin Cancer Res*, 14, 7917-23.
- Polkinghorn, W. R., Parker, J. S., Lee, M. X., Kass, E. M., Spratt, D. E., Iaquina, P. J., Arora, V. K., Yen, W. F., Cai, L., Zheng, D., Carver, B. S., Chen, Y., Watson, P. A., Shah, N. P., Fujisawa, S., Goglia, A. G., Gopalan, A., Hieronymus, H., Wongvipat, J., Scardino, P. T., Zelefsky, M. J., Jasin, M., Chaudhuri, J., Powell, S. N. & Sawyers, C. L. 2013a. Androgen receptor signaling regulates DNA repair in prostate cancers. *Cancer Discov*, 3, 1245-53.
- Polkinghorn, W. R., Parker, J. S., Lee, M. X., Kass, E. M., Spratt, D. E., Iaquina, P. J., Arora, V. K., Yen, W. F., Cai, L., Zheng, D. Y., Carver, B. S., Chen, Y., Watson, P. A., Shah, N. P., Fujisawa, S., Goglia, A. G., Gopalan, A., Hieronymus, H., Wongvipat, J., Scardino, P. T., Zelefsky, M. J., Jasin, M., Chaudhuri, J., Powell, S. N. & Sawyers, C. L. 2013b. Androgen Receptor Signaling Regulates DNA Repair in Prostate Cancers. *Cancer Discovery*, 3, 1245-1253.

- Poulidakos, P. I., Persaud, Y., Janakiraman, M., Kong, X., Ng, C., Moriceau, G., Shi, H., Atefi, M., Titz, B., Gabay, M. T., Salton, M., Dahlman, K. B., Tadi, M., Wargo, J. A., Flaherty, K. T., Kelley, M. C., Misteli, T., Chapman, P. B., Sosman, J. A., Graeber, T. G., Ribas, A., Lo, R. S., Rosen, N. & Solit, D. B. 2011. RAF inhibitor resistance is mediated by dimerization of aberrantly spliced BRAF(V600E). *Nature*, 480, 387-90.
- Prevo, R., Fokas, E., Reaper, P. M., Charlton, P. A., Pollard, J. R., McKenna, W. G., Muschel, R. J. & Brunner, T. B. 2012. The novel ATR inhibitor VE-821 increases sensitivity of pancreatic cancer cells to radiation and chemotherapy. *Cancer Biol Ther*, 13, 1072-81.
- Qi, L. S., Larson, M. H., Gilbert, L. A., Doudna, J. A., Weissman, J. S., Arkin, A. P. & Lim, W. A. 2013. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell*, 152, 1173-83.
- Radaeva, M., Ban, F., Zhang, F., LeBlanc, E., Lallous, N., Rennie, P. S., Gleave, M. E. & Cherkasov, A. 2021. Development of Novel Inhibitors Targeting the D-Box of the DNA Binding Domain of Androgen Receptor. *Int J Mol Sci*, 22.
- Rasool, R. U., Natesan, R., Deng, Q., Aras, S., Lal, P., Sander Efron, S., Mitchell-Velasquez, E., Posimo, J. M., Carskadon, S., Baca, S. C., Pomerantz, M. M., Siddiqui, J., Schwartz, L. E., Lee, D. J., Palanisamy, N., Narla, G., Den, R. B., Freedman, M. L., Brady, D. C. & Asangani, I. A. 2019. CDK7 Inhibition Suppresses Castration-Resistant Prostate Cancer through MED1 Inactivation. *Cancer Discov*, 9, 1538-1555.
- Rehman, Y. & Rosenberg, J. E. 2012. Abiraterone acetate: oral androgen biosynthesis inhibitor for treatment of castration-resistant prostate cancer. *Drug Des Devel Ther*, 6, 13-8.
- Reid, A. H., Attard, G., Danila, D. C., Oommen, N. B., Olmos, D., Fong, P. C., Molife, L. R., Hunt, J., Messiou, C., Parker, C., Dearnaley, D., Swennenhuis, J. F., Terstappen, L. W., Lee, G., Kheoh, T., Molina, A., Ryan, C. J., Small, E., Scher, H. I. & de Bono, J. S. 2010. Significant and sustained antitumor activity in post-docetaxel, castration-resistant prostate cancer with the CYP17 inhibitor abiraterone acetate. *J Clin Oncol*, 28, 1489-95.
- Rhee, H. W., Zou, P., Udeshi, N. D., Martell, J. D., Mootha, V. K., Carr, S. A. & Ting, A. Y. 2013. Proteomic mapping of mitochondria in living cells via spatially restricted enzymatic tagging. *Science*, 339, 1328-1331.

- Richters, A., Doyle, S. K., Freeman, D. B., Lee, C., Leifer, B. S., Jagannathan, S., Kabinger, F., Koren, J. V., Struntz, N. B., Urgiles, J., Stagg, R. A., Curtin, B. H., Chatterjee, D., Mathea, S., Mikochik, P. J., Hopkins, T. D., Gao, H., Branch, J. R., Xin, H., Westover, L., Bignan, G. C., Rupnow, B. A., Karlin, K. L., Olson, C. M., Westbrook, T. F., Vacca, J., Wilfong, C. M., Trotter, B. W., Saffran, D. C., Bischofberger, N., Knapp, S., Russo, J. W., Hickson, I., Bischoff, J. R., Gottardis, M. M., Balk, S. P., Lin, C. Y., Pop, M. S. & Koehler, A. N. 2021. Modulating Androgen Receptor-Driven Transcription in Prostate Cancer with Selective CDK9 Inhibitors. *Cell Chem Biol*, 28, 134-147 e14.
- Rivera-Calzada, A., Maman, J. D., Spagnolo, L., Pearl, L. H. & Llorca, O. 2005. Three-dimensional structure and regulation of the DNA-dependent protein kinase catalytic subunit (DNA-PKcs). *Structure*, 13, 243-55.
- Robinson, D., Van Allen, E. M., Wu, Y. M., Schultz, N., Lonigro, R. J., Mosquera, J. M., Montgomery, B., Taplin, M. E., Pritchard, C. C., Attard, G., Beltran, H., Abida, W., Bradley, R. K., Vinson, J., Cao, X., Vats, P., Kunju, L. P., Hussain, M., Feng, F. Y., Tomlins, S. A., Cooney, K. A., Smith, D. C., Brennan, C., Siddiqui, J., Mehra, R., Chen, Y., Rathkopf, D. E., Morris, M. J., Solomon, S. B., Durack, J. C., Reuter, V. E., Gopalan, A., Gao, J., Loda, M., Lis, R. T., Bowden, M., Balk, S. P., Gaviola, G., Sougnez, C., Gupta, M., Yu, E. Y., Mostaghel, E. A., Cheng, H. H., Mulcahy, H., True, L. D., Plymate, S. R., Dvinge, H., Ferraldeschi, R., Flohr, P., Miranda, S., Zafeiriou, Z., Tunariu, N., Mateo, J., Perez-Lopez, R., Demichelis, F., Robinson, B. D., Schiffman, M., Nanus, D. M., Tagawa, S. T., Sigaras, A., Eng, K. W., Elemento, O., Sboner, A., Heath, E. I., Scher, H. I., Pienta, K. J., Kantoff, P., de Bono, J. S., Rubin, M. A., Nelson, P. S., Garraway, L. A., Sawyers, C. L. & Chinnaiyan, A. M. 2015. Integrative clinical genomics of advanced prostate cancer. *Cell*, 161, 1215-1228.
- Rogakou, E. P., Boon, C., Redon, C. & Bonner, W. M. 1999. Megabase chromatin domains involved in DNA double-strand breaks in vivo. *J Cell Biol*, 146, 905-16.
- Rosner, W., Hryb, D. J., Khan, M. S., Nakhla, A. M. & Romas, N. A. 1991. Sex hormone-binding globulin: anatomy and physiology of a new regulatory system. *J Steroid Biochem Mol Biol*, 40, 813-20.

- Rothkamm, K., Kruger, I., Thompson, L. H. & Lobrich, M. 2003. Pathways of DNA double-strand break repair during the mammalian cell cycle. *Mol Cell Biol*, 23, 5706-15.
- Roux, K. J., Kim, D. I., Raida, M. & Burke, B. 2012. A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. *J Cell Biol*, 196, 801-10.
- Ruis, B. L., Fattah, K. R. & Hendrickson, E. A. 2008. The catalytic subunit of DNA-dependent protein kinase regulates proliferation, telomere length, and genomic stability in human somatic cells. *Mol Cell Biol*, 28, 6182-95.
- Schellhammer, P., Sharifi, R., Block, N., Soloway, M., Venner, P., Patterson, A. L., Sarosdy, M., Vogelzang, N., Jones, J. & Kolvenbag, G. 1995. A controlled trial of bicalutamide versus flutamide, each in combination with luteinizing hormone-releasing hormone analogue therapy, in patients with advanced prostate cancer. Casodex Combination Study Group. *Urology*, 45, 745-52.
- Scher, H. I., Fizazi, K., Saad, F., Taplin, M. E., Sternberg, C. N., Miller, K., de Wit, R., Mulders, P., Chi, K. N., Shore, N. D., Armstrong, A. J., Flaig, T. W., Flechon, A., Mainwaring, P., Fleming, M., Hainsworth, J. D., Hirmand, M., Selby, B., Seely, L., de Bono, J. S. & Investigators, A. 2012. Increased survival with enzalutamide in prostate cancer after chemotherapy. *N Engl J Med*, 367, 1187-97.
- Scher, H. I., Graf, R. P., Schreiber, N. A., Jayaram, A., Winkquist, E., McLaughlin, B., Lu, D., Fleisher, M., Orr, S., Lowes, L., Anderson, A., Wang, Y., Dittamore, R., Allan, A. L., Attard, G. & Heller, G. 2018. Assessment of the Validity of Nuclear-Localized Androgen Receptor Splice Variant 7 in Circulating Tumor Cells as a Predictive Biomarker for Castration-Resistant Prostate Cancer. *JAMA Oncol*, 4, 1179-1186.
- Schiewer, M. J., Goodwin, J. F., Han, S., Brenner, J. C., Augello, M. A., Dean, J. L., Liu, F., Planck, J. L., Ravindranathan, P., Chinnaiyan, A. M., McCue, P., Gomella, L. G., Raj, G. V., Dicker, A. P., Brody, J. R., Pascal, J. M., Centenera, M. M., Butler, L. M., Tilley, W. D., Feng, F. Y. & Knudsen, K. E. 2012. Dual roles of PARP-1 promote cancer growth and progression. *Cancer Discov*, 2, 1134-49.

- Schmidt-Hansen, M., Hoskin, P., Kirkbride, P., Hasler, E. & Bromham, N. 2014. Hormone and radiotherapy versus hormone or radiotherapy alone for non-metastatic prostate cancer: a systematic review with meta-analyses. *Clin Oncol (R Coll Radiol)*, 26, e21-46.
- Shaffer, P. L., Jivan, A., Dollins, D. E., Claessens, F. & Gewirth, D. T. 2004. Structural basis of androgen receptor binding to selective androgen response elements. *Proc Natl Acad Sci U S A*, 101, 4758-63.
- Sharp, A., Coleman, I., Yuan, W., Sprenger, C., Dolling, D., Rodrigues, D. N., Russo, J. W., Figueiredo, I., Bertan, C., Seed, G., Riisnaes, R., Uo, T., Neeb, A., Welti, J., Morrissey, C., Carreira, S., Luo, J., Nelson, P. S., Balk, S. P., True, L. D., de Bono, J. S. & Plymate, S. R. 2019. Androgen receptor splice variant-7 expression emerges with castration resistance in prostate cancer. *J Clin Invest*, 129, 192-208.
- Shen, M. M. & Abate-Shen, C. 2010. Molecular genetics of prostate cancer: new prospects for old challenges. *Genes Dev*, 24, 1967-2000.
- Sheth, N., Roca, X., Hastings, M. L., Roeder, T., Krainer, A. R. & Sachidanandam, R. 2006. Comprehensive splice-site analysis using comparative genomics. *Nucleic Acids Res*, 34, 3955-67.
- Shibata, K., Kajiyama, H., Ino, K., Terauchi, M., Yamamoto, E., Nawa, A., Nomura, S. & Kikkawa, F. 2008. Twist expression in patients with cervical cancer is associated with poor disease outcome. *Ann Oncol*, 19, 81-5.
- Shin, C., Feng, Y. & Manley, J. L. 2004. Dephosphorylated SRp38 acts as a splicing repressor in response to heat shock. *Nature*, 427, 553-8.
- Shoag, J. & Barbieri, C. E. 2016. Clinical variability and molecular heterogeneity in prostate cancer. *Asian J Androl*, 18, 543-8.
- Siegel, R. L., Miller, K. D. & Jemal, A. 2016. Cancer statistics, 2016. *CA Cancer J Clin*, 66, 7-30.
- Smith, M. R., Hussain, M., Saad, F., Fizazi, K., Sternberg, C. N., Crawford, E. D., Kopyltsov, E., Park, C. H., Alekseev, B., Montesa-Pino, A., Ye, D., Parnis, F., Cruz, F., Tammela, T. L. J., Suzuki, H., Utriainen, T., Fu, C., Uemura, M., Mendez-Vidal, M. J., Maughan, B. L., Joensuu, H., Thiele, S., Li, R., Kuss, I., Tombal, B. & Investigators, A. T. 2022.

- Darolutamide and Survival in Metastatic, Hormone-Sensitive Prostate Cancer. *N Engl J Med*, 386, 1132-1142.
- Smith, M. R., Saad, F., Chowdhury, S., Oudard, S., Hadaschik, B. A., Graff, J. N., Olmos, D., Mainwaring, P. N., Lee, J. Y., Uemura, H., Lopez-Gitlitz, A., Trudel, G. C., Espina, B. M., Shu, Y., Park, Y. C., Rackoff, W. R., Yu, M. K., Small, E. J. & Investigators, S. 2018. Apalutamide Treatment and Metastasis-free Survival in Prostate Cancer. *N Engl J Med*, 378, 1408-1418.
- Sramkoski, R. M., Pretlow, T. G., 2nd, Giaconia, J. M., Pretlow, T. P., Schwartz, S., Sy, M. S., Marengo, S. R., Rhim, J. S., Zhang, D. & Jacobberger, J. W. 1999. A new human prostate carcinoma cell line, 22Rv1. *In Vitro Cell Dev Biol Anim*, 35, 403-9.
- Steketee, K., Berrevoets, C. A., Dubbink, H. J., Doesburg, P., Hersmus, R., Brinkmann, A. O. & Trapman, J. 2002. Amino acids 3-13 and amino acids in and flanking the 23FxxLF27 motif modulate the interaction between the N-terminal and ligand-binding domain of the androgen receptor. *Eur J Biochem*, 269, 5780-91.
- Stockley, J., Markert, E., Zhou, Y., Robson, C. N., Elliott, D. J., Lindberg, J., Leung, H. Y. & Rajan, P. 2015. The RNA-binding protein Sam68 regulates expression and transcription function of the androgen receptor splice variant AR-V7. *Sci Rep*, 5, 13426.
- Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., Paulovich, A., Pomeroy, S. L., Golub, T. R., Lander, E. S. & Mesirov, J. P. 2005. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*, 102, 15545-50.
- Sun, S., Sprenger, C. C., Vessella, R. L., Haugk, K., Soriano, K., Mostaghel, E. A., Page, S. T., Coleman, I. M., Nguyen, H. M., Sun, H., Nelson, P. S. & Plymate, S. R. 2010. Castration resistance in human prostate cancer is conferred by a frequently occurring androgen receptor splice variant. *J Clin Invest*, 120, 2715-30.
- Sun, Y., Daemen, A., Hatzivassiliou, G., Arnott, D., Wilson, C., Zhuang, G., Gao, M., Liu, P., Boudreau, A., Johnson, L. & Settleman, J. 2014. Metabolic and transcriptional profiling reveals pyruvate dehydrogenase kinase 4 as a mediator of epithelial-mesenchymal transition and drug resistance in tumor cells. *Cancer Metab*, 2, 20.

- Tacke, R. & Manley, J. L. 1999. Determinants of SR protein specificity. *Curr Opin Cell Biol*, 11, 358-62.
- Takayama, K. I., Suzuki, T., Fujimura, T., Yamada, Y., Takahashi, S., Homma, Y., Suzuki, Y. & Inoue, S. 2017. Dysregulation of spliceosome gene expression in advanced prostate cancer by RNA-binding protein PSF. *Proc Natl Acad Sci U S A*, 114, 10461-10466.
- Tang, S., Sethunath, V., Metaferia, N. Y., Nogueira, M. F., Gallant, D. S., Garner, E. R., Lairson, L. A., Penney, C. M., Li, J., Gelbard, M. K., Alaiwi, S. A., Seo, J. H., Hwang, J. H., Strathdee, C. A., Baca, S. C., AbuHammad, S., Zhang, X., Doench, J. G., Hahn, W. C., Takeda, D. Y., Freedman, M. L., Choi, P. S. & Viswanathan, S. R. 2022. A genome-scale CRISPR screen reveals PRMT1 as a critical regulator of androgen receptor signaling in prostate cancer. *Cell Rep*, 38, 110417.
- Taplin, M. E., Bubley, G. J., Ko, Y. J., Small, E. J., Upton, M., Rajeshkumar, B. & Balk, S. P. 1999. Selection for androgen receptor mutations in prostate cancers treated with androgen antagonist. *Cancer Res*, 59, 2511-5.
- Taplin, M. E., Bubley, G. J., Shuster, T. D., Frantz, M. E., Spooner, A. E., Ogata, G. K., Keer, H. N. & Balk, S. P. 1995. Mutation of the androgen-receptor gene in metastatic androgen-independent prostate cancer. *N Engl J Med*, 332, 1393-8.
- Taylor, B. S., Schultz, N., Hieronymus, H., Gopalan, A., Xiao, Y., Carver, B. S., Arora, V. K., Kaushik, P., Cerami, E., Reva, B., Antipin, Y., Mitsiades, N., Landers, T., Dolgalev, I., Major, J. E., Wilson, M., Socci, N. D., Lash, A. E., Heguy, A., Eastham, J. A., Scher, H. I., Reuter, V. E., Scardino, P. T., Sander, C., Sawyers, C. L. & Gerald, W. L. 2010. Integrative genomic profiling of human prostate cancer. *Cancer Cell*, 18, 11-22.
- Tepper, C. G., Boucher, D. L., Ryan, P. E., Ma, A. H., Xia, L., Lee, L. F., Pretlow, T. G. & Kung, H. J. 2002. Characterization of a novel androgen receptor mutation in a relapsed CWR22 prostate cancer xenograft and cell line. *Cancer Res*, 62, 6606-14.
- Tilbrook, A. J. & Clarke, I. J. 2001. Negative feedback regulation of the secretion and actions of gonadotropin-releasing hormone in males. *Biol Reprod*, 64, 735-42.
- Tomlins, S. A., Rhodes, D. R., Perner, S., Dhanasekaran, S. M., Mehra, R., Sun, X. W., Varambally, S., Cao, X., Tchinda, J., Kuefer, R., Lee, C., Montie, J. E., Shah, R. B., Pienta,

- K. J., Rubin, M. A. & Chinnaiyan, A. M. 2005. Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. *Science*, 310, 644-8.
- Tonotsuka, N., Hosoi, Y., Miyazaki, S., Miyata, G., Sugawara, K., Mori, T., Ouchi, N., Satomi, S., Matsumoto, Y., Nakagawa, K., Miyagawa, K. & Ono, T. 2006. Heterogeneous expression of DNA-dependent protein kinase in esophageal cancer and normal epithelium. *Int J Mol Med*, 18, 441-7.
- Tran, C., Ouk, S., Clegg, N. J., Chen, Y., Watson, P. A., Arora, V., Wongvipat, J., Smith-Jones, P. M., Yoo, D., Kwon, A., Wasielewska, T., Welsbie, D., Chen, C. D., Higano, C. S., Beer, T. M., Hung, D. T., Scher, H. I., Jung, M. E. & Sawyers, C. L. 2009. Development of a second-generation antiandrogen for treatment of advanced prostate cancer. *Science*, 324, 787-90.
- Trials.gov, C. 2018. *Safety and Anti-Tumor Study of Oral EPI-506 for Patients With Metastatic Castration-Resistant Prostate Cancer* [Online]. Available: <https://clinicaltrials.gov/ct2/show/NCT02606123> [Accessed 1/5 2019].
- Trincado, J. L., Entizne, J. C., Hysenaj, G., Singh, B., Skalic, M., Elliott, D. J. & Eyras, E. 2018. SUPPA2: fast, accurate, and uncertainty-aware differential splicing analysis across multiple conditions. *Genome Biol*, 19, 40.
- van Bussel, M. T. J., Awada, A., de Jonge, M. J. A., Mau-Sorensen, M., Nielsen, D., Schoffski, P., Verheul, H. M. W., Sarholz, B., Berghoff, K., El Bawab, S., Kuipers, M., Damstrup, L., Diaz-Padilla, I. & Schellens, J. H. M. 2021. A first-in-man phase 1 study of the DNA-dependent protein kinase inhibitor peposertib (formerly M3814) in patients with advanced solid tumours. *Br J Cancer*, 124, 728-735.
- van der Kwast, T. H., Schalken, J., Ruizeveld de Winter, J. A., van Vroonhoven, C. C., Mulder, E., Boersma, W. & Trapman, J. 1991. Androgen receptors in endocrine-therapy-resistant human prostate cancer. *Int J Cancer*, 48, 189-93.
- Varambally, S., Dhanasekaran, S. M., Zhou, M., Barrette, T. R., Kumar-Sinha, C., Sanda, M. G., Ghosh, D., Pienta, K. J., Sewalt, R. G., Otte, A. P., Rubin, M. A. & Chinnaiyan, A. M. 2002. The polycomb group protein EZH2 is involved in progression of prostate cancer. *Nature*, 419, 624-9.

- Velot, L., Lessard, F., Berube-Simard, F. A., Tav, C., Neveu, B., Teyssier, V., Boudaoud, I., Dionne, U., Lavoie, N., Bilodeau, S., Pouliot, F. & Bisson, N. 2021. Proximity-dependent Mapping of the Androgen Receptor Identifies Kruppel-like Factor 4 as a Functional Partner. *Mol Cell Proteomics*, 20, 100064.
- Venables, J. P., Klinck, R., Koh, C., Gervais-Bird, J., Bramard, A., Inkel, L., Durand, M., Couture, S., Froehlich, U., Lapointe, E., Lucier, J. F., Thibault, P., Rancourt, C., Tremblay, K., Prinos, P., Chabot, B. & Elela, S. A. 2009. Cancer-associated regulation of alternative splicing. *Nat Struct Mol Biol*, 16, 670-6.
- Verrijdt, G., Schoenmakers, E., Alen, P., Haelens, A., Peeters, B., Rombauts, W. & Claessens, F. 1999. Androgen specificity of a response unit upstream of the human secretory component gene is mediated by differential receptor binding to an essential androgen response element. *Mol Endocrinol*, 13, 1558-70.
- Walker, K. J., Nicholson, R. I., Turkes, A. O., Turkes, A., Griffiths, K., Robinson, M., Crispin, Z. & Dris, S. 1983. Therapeutic potential of the LHRH agonist, ICI 118630, in the treatment of advanced prostatic carcinoma. *Lancet*, 2, 413-5.
- Wang, S., Yang, S., Nan, C., Wang, Y., He, Y. & Mu, H. 2018. Expression of Androgen Receptor Variant 7 (AR-V7) in Circulated Tumor Cells and Correlation with Drug Resistance of Prostate Cancer Cells. *Med Sci Monit*, 24, 7051-7056.
- Warburg, O. 1956. On the origin of cancer cells. *Science*, 123, 309-14.
- Warde, P., Mason, M., Ding, K., Kirkbride, P., Brundage, M., Cowan, R., Gospodarowicz, M., Sanders, K., Kostashuk, E., Swanson, G., Barber, J., Hiltz, A., Parmar, M. K., Sathya, J., Anderson, J., Hayter, C., Hetherington, J., Sydes, M. R., Parulekar, W. & investigators, N. C. P. M. U. P. 2011. Combined androgen deprivation therapy and radiation therapy for locally advanced prostate cancer: a randomised, phase 3 trial. *Lancet*, 378, 2104-11.
- Ware, K. E., Garcia-Blanco, M. A., Armstrong, A. J. & Dehm, S. M. 2014. Biologic and clinical significance of androgen receptor variants in castration resistant prostate cancer. *Endocr Relat Cancer*, 21, T87-T103.

- Warnmark, A., Treuter, E., Wright, A. P. & Gustafsson, J. A. 2003. Activation functions 1 and 2 of nuclear receptors: molecular strategies for transcriptional activation. *Mol Endocrinol*, 17, 1901-9.
- Watson, P. A., Chen, Y. F., Balbas, M. D., Wongvipat, J., Socci, N. D., Viale, A., Kim, K. & Sawyers, C. L. 2010. Constitutively active androgen receptor splice variants expressed in castration-resistant prostate cancer require full-length androgen receptor. *Proc Natl Acad Sci U S A*, 107, 16759-65.
- Weigel, N. L., Carter, T. H., Schrader, W. T. & Omalley, B. W. 1992. Chicken Progesterone-Receptor Is Phosphorylated by a DNA-Dependent Protein-Kinase during Invitro Transcription Assays. *Molecular Endocrinology*, 6, 8-14.
- Welti, J., Rodrigues, D. N., Sharp, A., Sun, S., Lorente, D., Riisnaes, R., Figueiredo, I., Zafeiriou, Z., Rescigno, P., de Bono, J. S. & Plymate, S. R. 2016. Analytical Validation and Clinical Qualification of a New Immunohistochemical Assay for Androgen Receptor Splice Variant-7 Protein Expression in Metastatic Castration-resistant Prostate Cancer. *Eur Urol*, 70, 599-608.
- Welti, J., Sharp, A., Brooks, N., Yuan, W., McNair, C., Chand, S. N., Pal, A., Figueiredo, I., Riisnaes, R., Gurel, B., Rekowski, J., Bogdan, D., West, W., Young, B., Raja, M., Prosser, A., Lane, J., Thomson, S., Worthington, J., Onions, S., Shannon, J., Paoletta, S., Brown, R., Smyth, D., Harbottle, G. W., Gil, V. S., Miranda, S., Crespo, M., Ferreira, A., Pereira, R., Tunariu, N., Carreira, S., Neeb, A. J., Ning, J., Swain, A., Taddei, D., Team, S. C. P. I. P. C. D., Schiewer, M. J., Knudsen, K. E., Pegg, N. & de Bono, J. S. 2021. Targeting the p300/CBP Axis in Lethal Prostate Cancer. *Cancer Discov*, 11, 1118-1137.
- Westra, E. R., Semenova, E., Datsenko, K. A., Jackson, R. N., Wiedenheft, B., Severinov, K. & Brouns, S. J. 2013. Type I-E CRISPR-cas systems discriminate target from non-target DNA through base pairing-independent PAM recognition. *PLoS Genet*, 9, e1003742.
- Willmore, E., Elliott, S. L., Mainou-Fowler, T., Summerfield, G. P., Jackson, G. H., O'Neill, F., Lowe, C., Carter, A., Harris, R., Pettitt, A. R., Cano-Soumillac, C., Griffin, R. J., Cowell, I. G., Austin, C. A. & Durkacz, B. W. 2008. DNA-dependent protein kinase is a therapeutic

- target and an indicator of poor prognosis in B-cell chronic lymphocytic leukemia. *Clin Cancer Res*, 14, 3984-92.
- Willoughby, C. E., Jiang, Y., Thomas, H. D., Willmore, E., Kyle, S., Wittner, A., Phillips, N., Zhao, Y., Tudhope, S. J., Prendergast, L., Junge, G., Lourenco, L. M., Finlay, M. R. V., Turner, P., Munck, J. M., Griffin, R. J., Rennison, T., Pickles, J., Cano, C., Newell, D. R., Reeves, H. L., Ryan, A. J. & Wedge, S. R. 2020. Selective DNA-PKcs inhibition extends the therapeutic index of localized radiotherapy and chemotherapy. *J Clin Invest*, 130, 258-271.
- Woolbright, B. L., Choudhary, D., Mikhalyuk, A., Trammel, C., Shanmugam, S., Abbott, E., Pilbeam, C. C. & Taylor, J. A., 3rd 2018. The Role of Pyruvate Dehydrogenase Kinase-4 (PDK4) in Bladder Cancer and Chemoresistance. *Mol Cancer Ther*, 17, 2004-2012.
- Wu, J. & Manley, J. L. 1989. Mammalian pre-mRNA branch site selection by U2 snRNP involves base pairing. *Genes Dev*, 3, 1553-61.
- Xing, J., Wu, X., Vaporciyan, A. A., Spitz, M. R. & Gu, J. 2008. Prognostic significance of ataxia-telangiectasia mutated, DNA-dependent protein kinase catalytic subunit, and Ku heterodimeric regulatory complex 86-kD subunit expression in patients with nonsmall cell lung cancer. *Cancer*, 112, 2756-64.
- Xiong, H., Nie, X., Zou, Y., Gong, C., Li, Y., Wu, H., Qiu, H., Yang, L., Zhuang, L., Zhang, P., Zhang, J., Wang, Y. & Xiong, H. 2017. Twist1 Enhances Hypoxia Induced Radioresistance in Cervical Cancer Cells by Promoting Nuclear EGFR Localization. *J Cancer*, 8, 345-353.
- Yang, C., Wang, Q., Liu, X., Cheng, X., Jiang, X., Zhang, Y., Feng, Z. & Zhou, P. 2016. NU7441 Enhances the Radiosensitivity of Liver Cancer Cells. *Cell Physiol Biochem*, 38, 1897-905.
- Yang, J., Mani, S. A., Donaher, J. L., Ramaswamy, S., Itzykson, R. A., Come, C., Savagner, P., Gitelman, I., Richardson, A. & Weinberg, R. A. 2004. Twist, a master regulator of morphogenesis, plays an essential role in tumor metastasis. *Cell*, 117, 927-39.
- Yin, Y., Li, R., Xu, K., Ding, S., Li, J., Baek, G., Ramanand, S. G., Ding, S., Liu, Z., Gao, Y., Kanchwala, M. S., Li, X., Hutchinson, R., Liu, X., Woldu, S. L., Xing, C., Desai, N. B., Feng, F. Y., Burma, S., de Bono, J. S., Dehm, S. M., Mani, R. S., Chen, B. P. C. & Raj, G. V. 2017.

- Androgen Receptor Variants Mediate DNA Repair after Prostate Cancer Irradiation. *Cancer Res*, 77, 4745-4754.
- Yu, Q., Zhang, K., Wang, X., Liu, X. & Zhang, Z. 2010. Expression of transcription factors snail, slug, and twist in human bladder carcinoma. *J Exp Clin Cancer Res*, 29, 119.
- Yu, Y., Guo, Y., Tian, Q., Lan, Y., Yeh, H., Zhang, M., Tasan, I., Jain, S. & Zhao, H. 2020. An efficient gene knock-in strategy using 5'-modified double-stranded DNA donors with short homology arms. *Nat Chem Biol*, 16, 387-390.
- Yu, Z., Chen, S., Sowalsky, A. G., Voznesensky, O. S., Mostaghel, E. A., Nelson, P. S., Cai, C. & Balk, S. P. 2014. Rapid induction of androgen receptor splice variants by androgen deprivation in prostate cancer. *Clin Cancer Res*, 20, 1590-600.
- Zenke, F. T., Zimmermann, A., Sirrenberg, C., Dahmen, H., Kirkin, V., Pehl, U., Grombacher, T., Wilm, C., Fuchss, T., Amendt, C., Vassilev, L. T. & Blaukat, A. 2020. Pharmacologic Inhibitor of DNA-PK, M3814, Potentiates Radiotherapy and Regresses Human Tumors in Mouse Models. *Mol Cancer Ther*, 19, 1091-1101.
- Zhang, Z., Connolly, P. J., Lim, H. K., Pande, V., Meerpoel, L., Teleha, C., Branch, J. R., Ondrus, J., Hickson, I., Bush, T., Luistro, L., Packman, K., Bischoff, J. R., Ibrahim, S., Parrett, C., Chong, Y., Gottardis, M. M. & Bignani, G. 2021. Discovery of JNJ-63576253: A Clinical Stage Androgen Receptor Antagonist for F877L Mutant and Wild-Type Castration-Resistant Prostate Cancer (mCRPC). *J Med Chem*, 64, 909-924.
- Zhao, Y., Thomas, H. D., Batey, M. A., Cowell, I. G., Richardson, C. J., Griffin, R. J., Calvert, A. H., Newell, D. R., Smith, G. C. & Curtin, N. J. 2006. Preclinical evaluation of a potent novel DNA-dependent protein kinase inhibitor NU7441. *Cancer Res*, 66, 5354-62.
- Zhou, K. I., Shi, H., Lyu, R., Wylder, A. C., Matuszek, Z., Pan, J. N., He, C., Parisien, M. & Pan, T. 2019. Regulation of Co-transcriptional Pre-mRNA Splicing by m(6)A through the Low-Complexity Protein hnRNPG. *Mol Cell*, 76, 70-81 e9.
- Zhou, Z. X., Sar, M., Simental, J. A., Lane, M. V. & Wilson, E. M. 1994. A ligand-dependent bipartite nuclear targeting signal in the human androgen receptor. Requirement for the DNA-binding domain and modulation by NH<sub>2</sub>-terminal and carboxyl-terminal sequences. *J Biol Chem*, 269, 13115-23.

