IIP Studentship: Developing the next generation of predictive biomarkers in cancer

Thesis submitted in partial fulfillment of the requirement of the degree of

Doctor of Philosophy

Ruaridh Duncan

Translational and Clinical Research Institute

Faculty of Medical Sciences

Newcastle University

March 2025

Abstract

A key driver in prostate cancer (PC) development and progression is aberrantly functional DNA repair mechanisms. DNA double strand breaks can be repaired by either homologous repair (HR) or error prone non-homologous end-joining (NHEJ), often resulting in genomic instability and resultant cellular transformation. Within 4-24% of PC patients, the SPOP protein is mutated, causing a loss of function and dysregulation in the multiple cellular processes it governs, including androgen receptor signalling and DNA repair. With SPOP found to be a key driver of HR, SPOP-mutant cells will have increased reliance on NHEJ, which, through inhibition of the key NHEJ protein DNA-PKcs, would potentially result in cell senescence or death due to inability to repair DNA damage.

In vitro experiments conducted within this study indicate Cas9-mediated *SPOP* knock-out PC cells display elevated sensitivity to DNA-PKIs and are unable to repair ionising radiation-induced DNA damage providing evidence of a synthetic-lethality relationship. With current DNA-PKIs in early phase clinical trials and generally well tolerated, current data highlights a potential new sub-population in PC where patients with *SPOP* mutant tumours could show increased benefit of DNA-PKIs in combination with radiation therapy.

Successful detection and stratification of PC into cohorts requiring sequentially more aggressive treatment represents a major clinical challenge; with current methods, such as PSA testing, lacking in sensitivity and selectivity. DNA-based nanotechnology is a rapidly advancing field utilising simple Watson-Crick base pairing fundamentals to develop new biosensors which holds the potential for rapid detection of multiple cancer biomarkers from non-invasive liquid biopsy samples with increased selectivity and sensitivity than current detection platforms. Through novel DNA probe designs targeting mRNA, microRNA and point mutations, such as those present in SPOP, experiments have been conducted which show great promise and attain the level of, if not surpass, the selectivity and sensitivity seen in by current PC detection methods.

Acknowledgements

I would first like to thank my supervisor Dr Luke Gaughan who allowed me the opportunity to join his research group. The ongoing support, mentorship and guidance I received throughout my time in the lab was instrumental to the completion of the project. I am massively grateful for being offered the experience to work alongside him and the rest of his research group. I would also like to thank my industrial supervisor Dr Jerzy Kozyra for his positivity and enthusiasm throughout the project which made my time with Nanovery massively enjoyable. I also thank Prof Rakesh Heer and panel members Dr Stuart McCracken and Dr Lisa Russel for assisting and providing steer throughout the course of this project.

A special thanks to all of the lab members within the solid tumour target discovery group for providing guidance and support when needed. I would like to thank all the PhD students within Luke's group who helped me when I started up in the lab and particularly Laura Walker and Beth Adamson, who not only assisted with aspects of the project but for their continued friendship throughout our time in the lab. A further thanks to Roma and Alex, and the rest of the team, both past and present, at Nanovery for the sustained support and friendship throughout the time I spent in their labs.

Thank you to my funders, Nanovery and the ERDF for the funding required to undertake this research.

I would like to thank my family for the continued encouragement throughout the last 4 years. A special thanks to my parents for the immense support they have provided throughout my PhD. A final thanks to my friends and rugby club who provided a much-needed balance to my life whilst completing my PhD.

Contents

List of abbreviations

ADT	Androgen deprivation therapy
AF1	Activation function 1
AF2	Activation function 2
AR	Androgen receptor
AR-V	Androgen receptor-Variants
AREs	Androgen response elements
ATRi	ATR inhibition
BACK	BTB and C-terminal Kelch
BER	Base excision repair
BSA	Bovine serum albumin
BTB	Broad-complex, tram-track and bric-a-brac
ccRCC	Clear cell renal carcinoma
CE3	AR cryptic exon 3
CE3-24P	24 nucleotide CE3 based probe
CE3-24T	24 nucleotide CE3 based target
CE3-30P	30 nucleotide CE3 based probe
CE3-30T	30 nucleotide CE3 based target
CE3-40P	40 nucleotide CE3 based probe
CE3-40T	40 nucleotide CE3 based target
CE3-60T	60 nucleotide CE3 based target
CHD1	Chromodomain helicase DNA-binding protein 1
Chk1i	Chk1 inhibition
CRL	Cullin-RING E3 ligase
CRPC	Castrate resistant prostate cancer
CTC	Circulating tumour cells
ctDNA	Circulating tumour DNA

CUL	Cullin
DBD	DNA binding domain
ddPCR	Digital droplet PCR
DDR	DNA damage response
DEG	Differentially expressed genes
DHT	5a-dihydrotestoterone
DMEM	Dulbecco's modified medium
DNA-Pkcsi	DNA-PKcs inhibition
DSB	Double strand breaks
E3	AR exon 3
E3-30P	30 nucleotide E3 based probe
E3-30T	30 nucleotide E3 based target
EMT	Epithelial mesenchymal transition
ER	Endoplasmic reticulum
F	Fluorophore
FC	Fold change
FCS	Foetal calf serum
FSH	Follicle-stimulating hormone
GEO	Gene expression omnibus
GG-NER	Global genome nucleotide excision repair
GO	Gene ontology
GPS	Genomic protein score
GSEA	Gene-set enrichment analysis
Gy	Gray
HR	Homologous repair
IF	Immunofluorescence
INDEL	Insertion/deletion
IR	Ionising radiation

KEGG	Kyoto Encyclopaedia of Genes and Genome
LBD	Ligand binding domain
LBP	Ligand binding pocket
LH	Luteinising hormone
LOD	Limit of detection
MATH	Meprin and TRAF-C homology domain
mCRPC	Metastatic CRPC
miRNA	MicroRNA
MMR	Mismatch repair
MT	Mutant
MT-IL	Mutant long input
MT-IL+Bl	Mutant long input with annealed blockers
MT-IS	Mutant short input
NER	Nucleotide excision repair
NES	Normalised enrichment score
NHEJ	Nonhomologous end joining
NLS	Nuclear localisation signal
NTD	N-terminal domain
NUDoxo	NU7441 & Doxorubicin treatment
PARP	Poly (ADP-ribose) polymerase 1
PARPi	PARP inhibition
PBS	Phosphate buffered saline
PC	Prostate cancer
PCA	Principle component analysis
PCA3	prostate cancer antigen
PCR	polymerase chain reaction
PHI	Prostate health index
PI	Propidium iodide

POZ	Pox virus and zinc finger
pre-miRNA	Precursor RNAs
pri-miRNA	Primary miRNA transcripts
PSA	Prostate specific antigen
PTM	Post-translational modifications
Q	Quencher
RBX1/2	Ring-box protein
RISC	RNA-induced silencing complex
RNAseq	RNA-sequencing
RP	Radical prostatectomy
RT	Radiation therapy
RT-qPCR	Real-time qPCR
SDEG	Significantly differentially expressed genes
SDS-PAGE	SDS-poly acrylamide gel electrophoresis
SEM	Standard error means
SEP	Stochiometric entropy probe design
SF	Surviving fraction
sgScr	Non-specific sgRNA
SHBG	Serum sex hormone binding globulin
siScr	Non-specific siRNA
siSPOP	SPOP-targeting siRNA
siSPOP-UTR	SPOP 3'UTR targeting siRNA
SMPC	SPOP-mutant PC
SPOP	Speckle-type pox virus and zinc finger protein
SPOP-L	SPOP-like
SPOP-WT	Wild-type SPOP
SRB	Sulforhodamine B
SRS	Substrate recognition protein

SSB	Single strand breaks
SSBR	Single strand break repair
ssDNA	Single stranded DNA
Tau	Transactivation units
TC-NER	Transcription coupled nucleotide excision repair
TCA	Trichloroacetic acid
TEMED	N, N, N’, N’-Tetramethylethan-1,2-diamine
TMPRSS2	Transmembrane protease serine 2
TMSD	Toehold mediated strand displacement
TOPI	Topoisomerase I
TOPIi	Topoisomerase I inhibition
UGS	Urogenital sinus
UPS	Ubiquitin proteosome system
WT	Wild-type
WT-IL	Wild-type long input
WT-IL+Bl	Wild-type long input with annealed blockers
WT-IS	Wild-type short input

List of Figures

List of tables

Chapter 1 – Introduction

1.1 The Prostate and Prostate Cancer

1.1.1 The Prostate, development and signalling

The prostate is a small tubular-alveolar gland located at the base of the urinary bladder in males. The development of the prostate can be broken down into multiple stages, each of which are dependent on androgens (Cunha et al., 2018). Initially, the pre-bud urogenital sinus (UGS) forms; where testosterone production commences concurrently as the foetal testes begins development. This precedes the initial budding of solid prostatic epithelial buds from the UGS epithelium which is driven by the influence of testosterone and other androgens acting through the androgen receptor (AR) expressed in the prostatic stroma (Cunha et al., 2018, Aaron et al., 2016). The elongation and branching of these buds, and the final ductal canalisation of the solid epithelial cords, follows these stages. Once this is completed, epithelial cell differentiation and secretory cytodifferentiation takes place (Cunha et al., 2018).

Alongside androgens, the AR is also present during these developmental stages, where it plays a role in mediating the morphogenic effects seen in prostatic development (Majumder and Kumar, 1997, Singh et al., 2014). The AR is expressed initially after 10 weeks of development, following which point it is constantly expressed. The interaction with androgens, such as testosterone, and the more potent 5α-dihydrotestosterone (DHT), with the AR drives the process of sexual differentiation and results in the development of male phenotypic features (Adams et al., 2002, Cunha et al., 2018).

Testosterone and DHT are not only crucial to the formation and growth of the prostate gland, they are also key to its function after development. Following development, the hypothalamic-pituitary-gonadal axis regulates the biosynthesis of testosterone, with this axis comprising the hypothalamus, both the anterior and posterior pituitary glands and the testes [Figure 1.1]. Gonadotropin-releasing hormone is released by the hypothalamus in a pulsatile manner that controls the synthesis and secretion of luteinising hormone (LH) and follicle-stimulating hormone (FSH) from the pituitary gland (Kluth et al., 2014, Clavijo and Hsiao, 2018). LH then induces synthesis of testosterone from the testes, the systemic concentrations of which can feed-back and inhibit both GnRH and LH signalling. This systemic testosterone can reach the

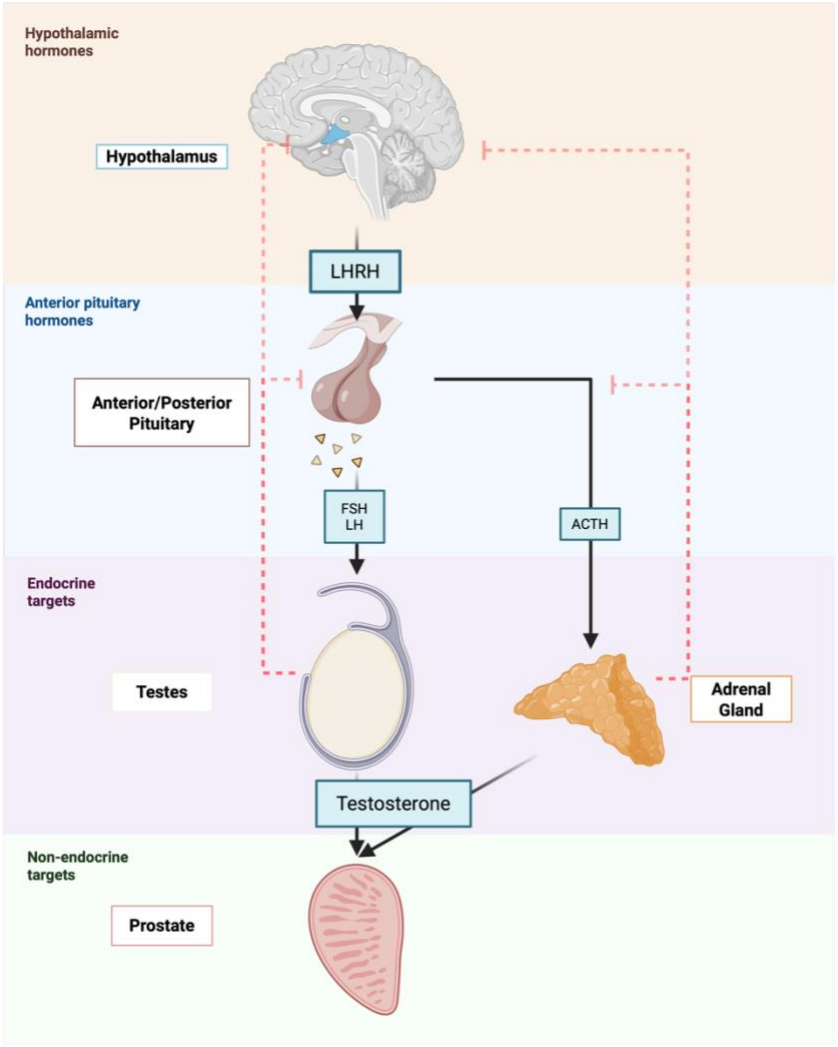


Figure 1.1 – Schematic of the hypothalamic-pituitary-gonadal axis. *Hormones released from tissues indicated by blue boxes with arrows displaying the target tissue where they have their effect in the axis. Red dotted flat head arrows indicate the presence of negative feedback mechanisms and the tissues these feedback mechanism target to bring about their effects.*

prostate and induce cell proliferative and survival effects through interaction with the AR in prostate cells (Kluth et al., 2014).

1.1.2 Prostate Cancer

Currently prostate cancer (PC) is one of the most common cancers worldwide and the most common cancer in men, with an estimated 1,600,000 cases and 366,00 deaths annually (Wang et al., 2018). In the UK alone, more than 47,500 men are diagnosed with PC and upwards of 11,500 patients dying of PC each year (PCUK, 2021). This makes PC the second most common form of cancer related death in the UK, with 5-year survival from the point of diagnosis varying dependent on tumour staging (PCUK, 2021).

There are many noted risk factors correlated with increased chance of an individual getting PC within their lifetime, as well as risk factors that imply a more aggressive PC phenotype. One such risk factor is a family history of PC (Kiciński et al., 2011). A male who has a member of their family diagnosed with PC at a young age (>40 years old) is more likely to develop PC themselves. Further, there is also an increased risk of developing PC if more than one family member has been diagnosed with PC previously (Johns and Houlston, 2003). With the wider PC screening methods, it was also found that a family history of PC also predisposes an individual for early PC onset (PC developing >40 years old) when compared to those without (Chen et al., 2008b). Another main risk factor in PC development is ethnicity, as confirmed in the PROCESS Cohort study which saw a significant two-fold increase PC risk in black males than that in white males (Ben-Shlomo et al., 2008). There are a number of potential causes of this increase, including higher circulating testosterone levels, higher AR protein levels and increased somatic and germline mutations of the *AR* gene, amongst others (Lowder et al., 2022). Other notable genetic factors seen within black men are alterations in both *CHD1* and *SPOP* genes, with nearly double the frequency of SPOP mutations being reported and increased prevalence of CHD1 deletions (Lowder et al., 2022). This difference in PC risk seen between ethnic populations has been shown not to be impacted by differences in cancer registration and healthcare differences globally (Grönberg, 2003). As well as those previously mentioned, other factors increasing PC risk have been described including genetic factors, age and lifestyle risk factors such as an individual's diet (Rawla, 2019). Many genetic factors have been described as risk factors in PC, with a 57% estimated genetic heritability of PC (Mucci et al., 2016, Bergengren et al., 2023). Of note, the most common genetic alterations that confer increased risk are found in *BRCA2* and *HOXB13* genes (Nyberg et al., 2019, Nyberg et al., 2022).

PC is typically characterised as abnormal prostate growth due to abnormally dividing cells within the prostate gland. The main cause of death from PC comes from when the tumour

metastasises and spreads to distal sites in the body (Schatten, 2018). PC can be graded based on tumour aggressiveness using Gleason scores that have been refined over the years, but were initially described by Donald Gleason in 1974 (Filella and Foj, 2018).

Early-stage PC is defined as when the tumour is localised to the prostate, whilst late-stage disease is when the tumour is locally advanced (remaining localised to the prostate) or the tumour has metastasised to distal sites in the body such as the pelvic lymph nodes (Kim and Bullock, 2018). If the tumour at the point of diagnosis is in the early stages, there is 100% 5-year survival of patients. However, if at diagnosis, PC is late stage there is only a 49% 5-year survival (PCUK, 2021). As PC progresses, the available options for treatment diminish, with the mainstay of treatment for late-stage PC being androgen deprivation therapy (ADT) which will eventually expedite progression to a more aggressive form of PC, termed castration resistant prostate cancer (CRPC), which is associated with decreased survival outcome for patients.

PC cells retain the androgen and AR dependencies seen in normal prostate cells (Gleason and Mellinger, 1974). A key finding in PC was described by Huggins and Hodges in 1941 that reported that PC was responsive to hormone treatment, and that castration resulted in tumour regression (Huggins and Hodges, 1972, Dehm and Tindall, 2011). This was the basis for the development of ADT; using agents to block AR signalling pathways through either pharmacological castration, surgical castration, or administration of anti-androgens (Gleason and Mellinger, 1974). This showed great initial results in inhibiting AR signalling and transcriptional activity, and most importantly tumour regression. However, after a median time of approximately 2 years, ADT treatment resistance can develop in patients via several resistance mechanisms that ultimately sees the resumption of AR signalling, resulting in more aggressive or metastatic CRPC (mCRPC), which are the main causes of PC morbidity (Wang et al., 2018, Gleason and Mellinger, 1974).

The reactivation of AR activity during treatment with ADT is a key event in almost all PC cases, with the progression from localised hormone-naïve tumours to the castrate resistant state generally attributed to aberrant AR signalling allowing treatment evasion (Dehm and Tindall, 2006, Lonergan and Tindall, 2011, Attard et al., 2016) . The restoration of androgen signalling in CRPC can be due to many mechanisms, including: overexpression or amplification of the *AR* gene; gain of function *AR* mutations, usually within the ligand binding domain (LBD); intracrine androgen production; overexpression of AR co-factors; ligand independent activation by cytokines or growth factors; and constitutively active splice variants of the AR, termed AR-Vs

(AR-Vs) (Lonergan and Tindall, 2011). Genetic alterations to the *AR* are found in 62.7% of mCRPC patients, with AR pathway associated aberrations seen in 71.3% of patients, further highlighting the aberrant activity in the AR and restoration of signalling being a key driver in PC development and progression (Robinson et al., 2015).

Like other cancers, PC has a massive number of genetic aberrations that can act either as cancer drivers or lead to increased cancer progression. In PC these aberrations are often present in PI3K, RAF, Wnt and AR signalling pathways, as well as in DNA repair and cell cycle control, ultimately leading to PC development and progression through enhanced cell survivability and proliferation (Robinson et al., 2015). Alterations can be either germline or somatic, with different somatic genetic alterations being implicated at different stages in disease progression (Wallis and Nam, 2015). One common mutation seen in PC is in *TP53* which is an early driver of disease progression, highlighted by alterations in this gene being found in 53.3% of mCRPC patients (Robinson et al., 2015, Abida et al., 2017). Alongside *TP53* mutations, another common oncogene *PTEN* is mutated in 40.7% of mCRPC patients, again leading to heightened disease progression (Robinson et al., 2015).

Of note in PC, DND damage repair pathways (discussed in Section 1.7) are frequently mutate in PC. DNA damage repair mutations, which are drivers of tumorigenesis, are seen in ~27% PC patients often seen across multiple stages of PC progression (Lang et al., 2019). Germline/somatic mutations in the homologous recombination (HR) driver *BRCA2* are often seen in early-stage disease and lead to a more aggressive disease, ultimately leading to metastasis. Germline *BRCA2* mutations have also been described in PC and increase the likelihood of tumour development ~6-fold (Kote-Jarai et al., 2011). Hereditary *HOXB13* mutations are also attributed with increased PC risk and are more common in familial early onset PC (Ewing et al., 2012). Other common DNA damage repair mutations include *ATM* mutations which are seen to be enriched in CRPC tumours (Abida et al., 2017). Aberrant activity in DNA damage repair can often lead to genetic rearrangements, which may contribute to the formation of ETS fusions, such as the common *TMPRSS2:ERG* fusion, which are seen in 56.7% of mCRPC patients. This event fuses the androgen regulated promoter of *TMPRSS2* to the oncogenic transcription factor ERG, leading to further disease progression (Glinsky et al., 2004, Lapointe et al., 2004). Alongside genetic alterations in tumour suppressors and oncogenes, there is now a growing body of evidence for massive changes in the microRNA

(miRNA) expression landscape in PC, with several miRNAs now identified as regulators of many key PC drivers, with their elevated or lost expression further impacting the already perturbed genetic landscape of PC tumours (discussed further in Chapter 8) (Wallis and Nam, 2015).

1.2 The Androgen Receptor

1.2.1 Androgen Receptor signalling

The AR is a member of the steroid hormone group of nuclear receptors and shares a similar general structure to other group members (Mangelsdorf et al., 1995). It is a ligand-dependent transcription factor that regulates many target genes relating to cell survival and proliferation [Figure 1.2]. Testosterone and DHT both bind to the AR and induce transcription of target genes, although DHT binds with a much greater affinity than testosterone. A majority of circulating testosterone is produced in Leydig cells (~90%) in the testes controlled by LH or the adrenal glands (~10%), with ~98% being bound to serum sex hormone binding globulin (SHBG) or albumin (Rosner et al., 1991, Harris et al., 2009, Kluth et al., 2014, Tan et al., 2015). DHT is also found in circulation being produced from the prostate, skin or testes (Harris et al., 2009). Only unbound androgens can enter prostate cells, where testosterone is converted to DHT by 5α-reductase and CYP450, where it can then bind to the C-terminal LBD of AR and displace Hsp90, alongside other intra-molecular chaperones and high molecular weight immunophilins, that are bound to the AR in the cytosol via the LBD (Rosner et al., 1991, Pratt and Toft, 1997, Baker, 2002, Srinivas-Shankar and Wu, 2006, Smith and Toft, 2008, Culig and Santer, 2014). This ligand binding induces a conformational change in the AR and allows receptor dimerisation, exposure of a nuclear localisation signal (NLS) and interaction with importin-α which permits translocation into the nucleus. Within the nucleus, active AR can bind to target gene promoter and enhancer regions via androgen response elements (AREs), where AR facilitates the recruitment of multiple proteins that are required for transcription. Androgen target genes include *prostate specific antigen* (*PSA*) and *transmembrane protease serine 2* (*TMPRSS2*), with AR mediated signalling pathways ultimately leading to growth and survival responses in prostate cells (Rosner et al., 1991, Shang et al., 2002, Dehm and Tindall, 2006).

1.2.2 AR signalling regulation

Regulation of the AR through post-translational modifications (PTM) is important for AR activity in normal prostate homeostasis, ultimately providing further AR signalling control within prostate cells. AR levels can be regulated indirectly through upstream pathways and histone acetylation/methylation, or through direct PTMs to the AR protein including

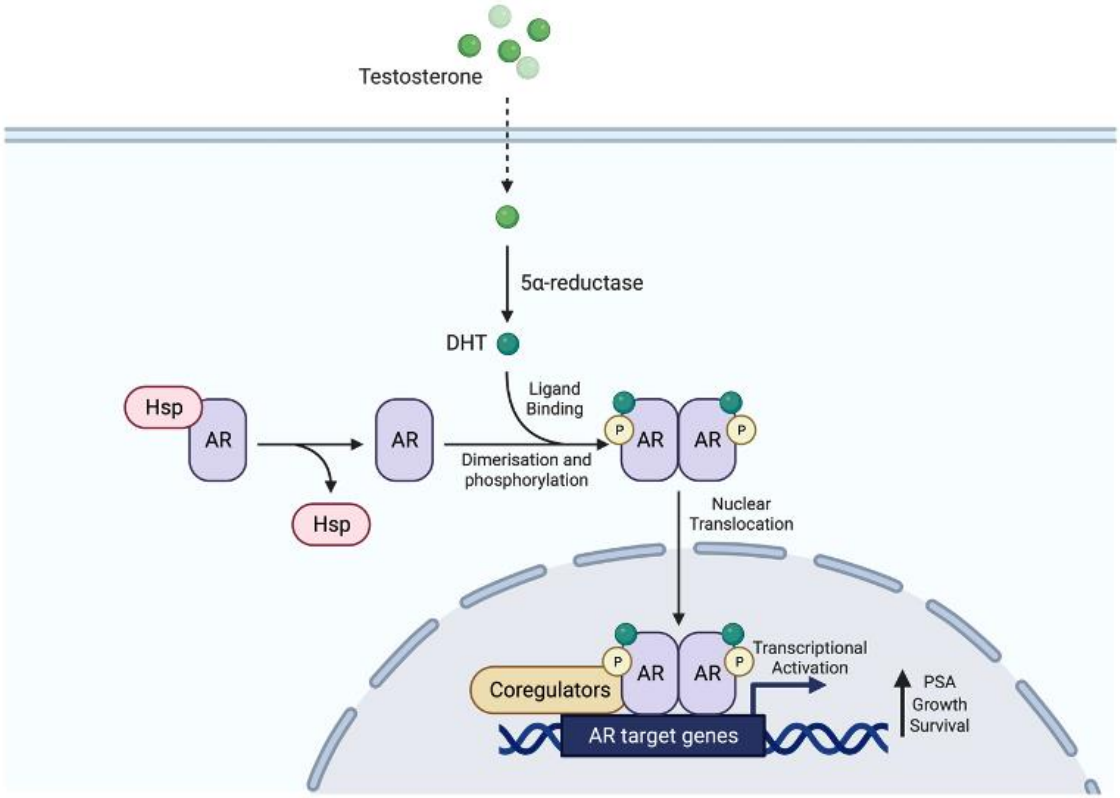


Figure 1.2 – Schematic diagram of the AR signalling axis within prostate cells. *Testosterone diffuses into epithelial prostate cells where it is metabolised by 5α-reductase to DHT in the cytoplasm. DHT binds to the androgen receptor (AR) in the cytoplasm causing the inhibitory foldosome complex, made up of HSPs and chaperone proteins, to dissociate. This dissociation allows AR dimerisation, phosphorylation and subsequent translocation into the nucleus where it binds androgen response elements (ARE's) in the enhancer and promoter regions of AR-regulated genes, where its transcription factor function facilitates the expression or repression of target genes.*

phosphorylation, acetylation, SUMOylation and ubiquitination (Coffey and Robson, 2012, Wen et al., 2020). Multiple sites on the AR can be subject to phosphorylation, with one of the best characterised being Serine 81 within the AR N-terminal domain (NTD) by AKT and CDK family members, which impacts AR activity through affecting its localisation ability and protein stability (Gordon et al., 2010). As a result of its effects on AR signalling this site is commonly phosphorylated in PC to stabilise AR levels and increase AR activity, resulting in activation of a pro-proliferative and survival phenotype (Coffey and Robson, 2012, Wen et al., 2020). Other phosphorylation sites on the AR include Tyr 287 and Tyr 363 which are both targets of the kinase ACK1, leading to increased AR target gene expression and ultimate promotion of androgen independent PC progression (Mahajan et al., 2007). Furthermore, ACK1 has been found to modulate AR activity through its kinase activity by phosphorylation of histone H4-Tyr88 upstream of the AR transcriptional start site. This in turn is read by the WDR5/MLL2 complex, leading to H3K4-trimethylation resulting in promotion of AR transcription (Mahajan et al., 2017).

The AR is also the target for ubiquitination, with multiple E3 ligases being implicated in the addition of single or multiple ubiquitin proteins, ultimately leading to ubiquitin-mediated proteasomal degradation. MDM2, responsible for p53 ubiquitination and degradation, also polyubiquitinates AR, with the interaction being enhanced by Akt-mediated phosphorylation of both AR and MDM2 (Lin et al., 2002). This polyubiquitination leads to AR proteosome-mediated degradation, further regulating AR activity. Another protein responsible for poly-ubiquitinating AR is RNF6 which, unlike MDM2-mediated ubiquitination, promotes AR signalling. Addition of K6/K27 ubiquitin chains to AR by RNF6 act as scaffold proteins allowing the recruitment of AR co-factors and promoting expression of AR target genes (Xu et al., 2009). Another E3 ligase that directly targets AR for ubiquitination is Speckle-type pox virus and zinc finger (POZ) protein (SPOP), due to the AR harbouring the specific SPOP-recognition motif within its hinge region. Indirect polyubiquitination by SPOP, like MDM2-mediated ubiquitination, leads to subsequent proteosome degradation of AR and reduced AR activity (An et al., 2014). As a result, due to their effects on AR activity, these E3 ligases are commonly found to be aberrantly functional in PC to resume AR signalling through alterations in the proposed MDM2 pathways, upregulation of RNF6 and point mutations in *SPOP* leading to lack of substrate binding.

1.2.3 AR structure

The *AR* gene is located on the X-chromosome with a 2757 nucleotide coding region that spans across 8 exons, encoding a 919 amino acid, 110 kDa protein (Brown et al., 1989). The AR protein itself is made up of three functional domains, the NTD, a DNA binding domain (DBD) and C-terminal LBD. The DBD and LBD are also joined by a hinge region that is important for the function of the AR protein as a whole (Gelmann, 2002). Also of note are two activation function domains that are key for AR activity. Activation function 1 (AF1) is found within the NTD and is constitutively active (McEwan, 2004), whereas the ligand dependent activation function 2 (AF2) is found in the LBD (He et al., 1999).

The AR NTD is entirely encoded by exon 1 of the *AR* gene and spans amino acids 1-555, although this length can vary depending on the size of polymorphic (CAG)n and (GGN)n repeats (Ding et al., 2004, Ferro et al., 2002). Within the NTD is the AF1 region, spanning from residues 142-485, and acts as the main transcriptional effector region of the NTD. AF1 is a

much more potent transcriptional activator than its counterpart AF2, with this activity being directed by two transactivation units (Tau) that are essential for full AR activity, Tau-1 and Tau-5, spanning residues 100-370 and 360-485, respectively (Callewaert et al., 2006). These domains harbour specific motifs that allow and mediate the inter-domain interactions of the NTD and the LBD upon ligand binding, termed the N/C-interaction, which stabilises the AR dimers formed and slows ligand dissociation (Doesburg et al., 1997, Langley et al., 1998, He et al., 1999).

The AR DBD spans residues 555-623 and is made up of a core of two zinc finger domains that target AREs. The AREs themselves are made up of two hexameric half-sites separated by a 6 base pair spacer (Shaffer et al., 2004). The individual zinc fingers are encoded by exon 2 and exon 3; with the N-terminal zinc finger harbouring the α-helix known as the 'recognition helix' which is responsible for the recognition and interaction with the nucleotides in the major groove of AREs (Beato et al., 1995). The second zinc finger is the dimerisation interface for interaction with another AR monomer, as well as making non-specific DNA backbone contacts to support DNA binding (Shaffer et al., 2004).

The DBD is followed by the hinge region spanning residues 623-655, with the junction of these domains harbouring the NLS that is responsible for nuclear import of the AR (Jenster et al., 1993, Zhou et al., 1994). Androgen binding induces a conformational switch in the receptor which exposes the NLS, which facilitates the binding of importin-α and allows subsequent nuclear import (Ni et al., 2013). The hinge region is also the target for methylation by Set9 (Gaughan et al., 2011), ubiquitination by MDM2 and SPOP (Lin et al., 2002, An et al., 2014) and acetylation by p300 and TIP60 (Fu et al., 2000, Gaughan et al., 2002) and plays further roles in co-activator recruitment, DNA binding and the N/C interaction (Clinckemalie et al., 2012).

The LBD spans residues 665-919 and is made up of a three-layered anti-parallel 11 α-helical sandwich fold, encoded by exons 4-8. The N-terminal ends of α-helices 3, 5 and 11 make up the ligand binding pocket (LBP), and helix 12 is the core of the AF2 region located in the LBD, which closes upon androgen binding (Matias et al., 2000). Although the AR LBD contains 11 helices, the nomenclature of the structural fold has been maintained with other steroid hormone receptors which contain 12 helices; as helix 2 is absent in the AR LBD. Upon androgen binding, a conformational change in the LBD takes place which exposes the NLS in the hinge region, allowing the interaction with importin-α and the formation of a hydrophobic pocket

within the LBD which facilitates the dimerization of the androgen receptor (Zhou et al., 1994, Cutress et al., 2008, Culig and Santer, 2014). Following nuclear entry, the AR dimerises and binds AREs resulting in the recruitment of the basal transcriptional activity that is required for the tightly controlled transcription of target genes.

1.3 PC treatment

There are currently multiple treatment options available to PC patients, and with PC being a heterogenous disease, the treatment decision varies from patient to patient, and is dependent on age, tumour grading and staging (Wang et al., 2018). For most localised cancers, the Gleason score and patient risk is used to guide treatment regimen. Treatment options in PC can range from active surveillance to combinations of surgery, chemotherapy, radiation and/or ADT, dependent on PC stage at presentation [Figure 1.3] (Kim and Bullock, 2018).

1.3.1 Localised PC

Localised PC is when the tumour is confined to the prostate. Treatment options available when presenting at this stage include radical prostatectomy (RP), radiation therapy (RT) and active surveillance (Mottet et al., 2017, Tselis et al., 2017). Currently RP serves as the gold standard for the surgical management of early PC however this may not be the best treatment option in some later stages (Kim and Bullock, 2018). RP can be used in all PC stages, and in most cases offers increased survival benefit. However, in patients with advanced metastatic disease the potential benefit of RP may be outweighed by the negative effects on quality-of-life associated with RP. In early disease, active surveillance may be a better suited treatment option. Active surveillance is when patients are not initially treated and are instead monitored closely with plans in place should the disease progress. Expression of prostate specific AR target gene *KLK3* which encodes PSA can be used as a biomarker to monitor the tumour stage along with Gleason score assessed via prostate biopsy. Although basal PSA levels can vary, in the UK PSA levels of 3 ng/ml or higher in the blood would promote further tests to be conducted and enrol a patient onto active surveillance. Usually, PSA levels are used in conjunction with Gleason grade (≤ 6) and biopsy results to inform the decision to enrol patients on active surveillance or continue with other treatment options such as RP or RT. With the advent of improved imaging options available, such as 3T multiparametric resonance imaging, there can be increased reductions in morbidity due to delaying RP (Kim and Bullock, 2018). Also of note, PC can be described as locally advanced, in which the disease is in early stages of metastasis and spreads to the lymph nodes before distant organs (Wang et al., 2018). Here, standard of care treatment includes RT in combination with ADT which shows increased survival benefits (Widmark et al., 2009, Warde et al., 2011).

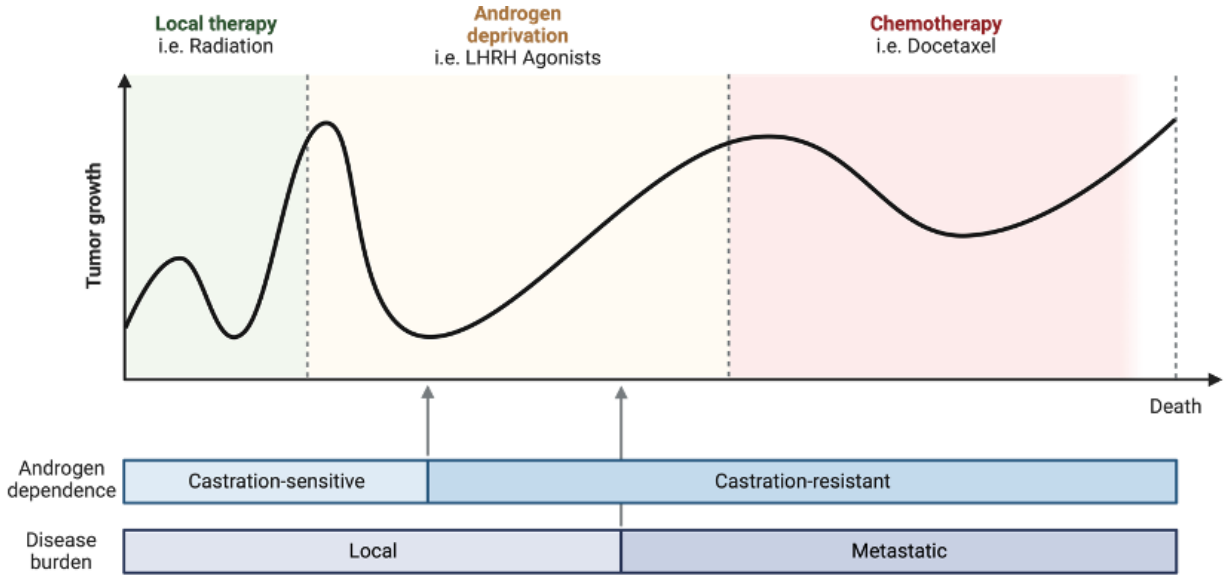


Figure 1.3- Schematic representation of typical PC progression with available treatment options at each disease stage. *Treatment options available for PC very depending on disease stage at the time of presentation at the clinical, usually assessed through monitoring of multiple parameters including Gleason score from tumour biopsy and circulating PSA levels. PSA levels are often monitored at clinical presentation through active surveillance, where treatment interventions are in place should PSA levels increase. Early stage localised PC sees beneficial results from surgery and radiotherapy and continued active surveillance. If PSA levels increase further patients will be receive androgen-deprivation therapy which shows great initial responses although will ultimately relapse. At this point patients can receive second-generation anti-androgens such as enzalutamide, although eventually patient tumours will transition to a castrate resistant state and metastasise, resulting in patients left with chemotherapeutics such as docetaxel as palliative treatment options.*

1.3.2 Advanced PC

Advanced PC is when the tumour has spread to distal sites in the body and would no longer benefit from treatment options targeting the prostate such as RT or RP. Following RT or RP for treatment of localised PC, 20% of patients will eventually show increased PSA levels and progress to a more advanced disease (Cornford et al., 2017, Messner et al., 2020). In this case, and if a patient presents with advanced PC, patients will undergo ADT. ADT can involve chemical castration using drugs that target the hypothalamic-pituitary-gonadal axis, or surgical castration with the ultimate aim to reduce the levels of circulating androgens in the body, due to their key role in PC progression [Figure 1.4] (Wang et al., 2018). Anti-androgens are commonly used in ADT and over the years many different anti-androgens have been developed that target and inhibit the AR signalling pathway. Initial anti-androgens were steroidal, which decreased testosterone binding to AR. However, these compounds showed an unfavourable therapeutic index and were subsequently replaced by non-steroidal anti-androgens which were more specific to the AR and, as a result, much safer. First generation anti-androgens

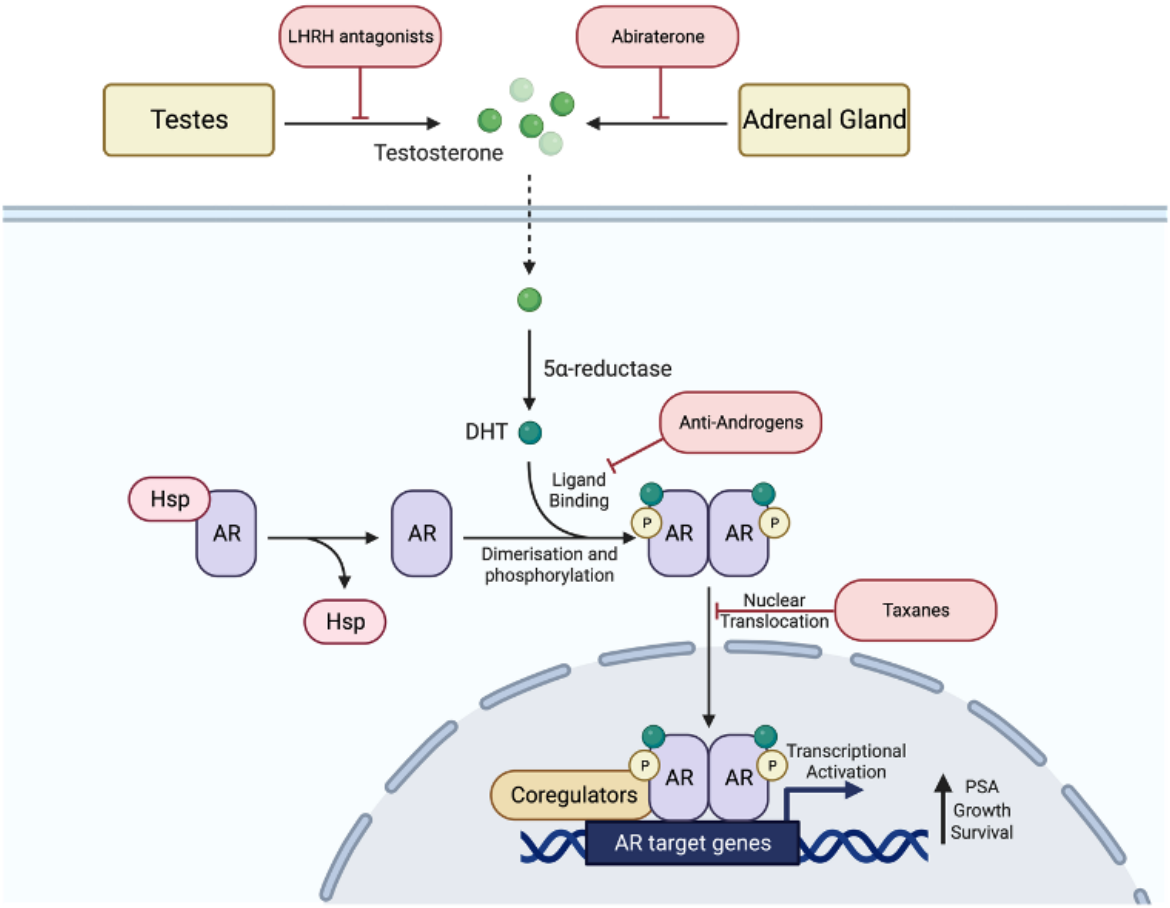


Figure 1.4 - Schematic indicating current therapeutic options for PC treatment*. Androgen receptor signalling axis with therapeutic options highlighted in red boxes with flat head arrows indicating the regions of the signalling axis at which they exert their effects.*

included flutamide and nilutamide which, after time, would develop agonist activity due to *AR* mutations converting these agents from antagonists to agonists. The second-generation anti-androgens, such as bicalutamide, that followed these showed the same efficacy but enabled once-daily dosing due to increased half-lives (Crawford et al., 2018).

1.3.3 CRPC

In most cases, ADT is initially well received with patients showing signs of tumour regression. However, in almost all cases, the recurrence of the cancer is seen leading to the more aggressive or metastatic CRPC. As stated previously, this can be due to several mechanisms ultimately resulting in the resumption of AR signalling which is a major driver of cancer progression (Crawford et al., 2018). This prompted the development of third-generation anti-androgens, such as enzalutamide. Enzalutamide inhibits AR translocation into the nucleus and binds the AR with a 5-8 fold greater binding affinity when compared to second-generation anti-androgens. Alongside third generation anti-androgens, drugs such as abiraterone can also be

used to treat CRPC. Abiraterone inhibits CYP17A1, which depletes adrenal and intratumoral androgens (O'Donnell et al., 2004, Attard et al., 2008) and both enzalutamide and abiraterone have been seen to lead to an increase in patient survival in both localised and mCRPC patients (Crawford et al., 2018). However, following treatment with these agents, secondary resistance can develop, meaning there is an increased demand for future therapeutics that overcome these resistance mechanisms, or better methods of staging and grading patients PC so as to not start unnecessary treatments at later stages of the disease. Recently new potential treatment options have been investigated and seen use in PC in the form of selective androgen modulators and degraders. Modulators act through similar functions as steroidal androgens, acting as AR agonists and resulting in the activation of canonical AR signalling and ultimately supress growth of PC cells (Nyquist et al., 2021). Degraders facilitate the blocking of AR signalling through the degradation of the AR by the ubiquitin proteome through compromised AR stability through binding to the AF-1 region and show great promise from positive response to treatment in enzalutamide-resistant PC cells (Hwang et al., 2019, Ponnusamy et al., 2019, Wu et al., 2023). Ultimately, drugs that target this resistance mechanisms are in development, however methods of detecting resistance that allow delivery of these agents earlier is likely to improve the outcome for patients.

1.4 DNA Damage Repair

DNA repair mechanisms are vital for maintaining DNA integrity and hence normal cell viability. As such, dysfunctional DNA repair processes result in genomic instability which is one of the classically described hallmarks of cancer (Hanahan, 2022). The central dogma for DNA repair goes as follows: detection of DNA damage via the DNA damage response (DDR); halting of the cell cycle; and activation of the repair machinery to repair the damaged DNA. If the repair is successful, the cell can continue as normal. However, if the damage is unable to be repaired successfully, either senescence or cell death pathways are triggered (Mateo et al., 2017). DNA damage can be caused through endogenous sources, such as replication errors and oxidative DNA damage, or exogenous sources in the form of ionising or UV radiation, chemical sources, or environmental stresses [Figure 1.5] (Chatterjee and Walker, 2017). These are highly common and often unavoidable, resulting in different forms of DNA lesions that lead to DNA single strand breaks (SSBs) or double strand beaks (DSBs) which must be repaired for normal cell activity. Thus, much work has been devoted to increasing knowledge of the mechanisms that facilitate DNA damage repair.

The DDR is responsible for the detection of DNA lesions and directing the appropriate repair response. Sensor proteins within cells detect specific DNA lesions which then facilitate the recruitment of DDR factors in a highly regulated process, directing repair pathway choices specific to the lesion detected (Harper and Elledge, 2007). This is coupled with chromatin remodelling, allowing free access of the sensor and repair factors to the site of the DNA lesion. With DDR acting as the initiation step for DNA damage repair, mutations within the range of proteins involved in this step can result in the promotion of tumorigenesis (Ciccia and Elledge, 2010).

1.5 Repair of DNA lesions and SSBs

1.5.1 Nucleotide excision repair (NER)

NER is used in the repair of bulky DNA lesions or DNA crosslinking often caused by UV radiation exposure. Here, and in Base Excision Repair (BER), prior to repair pathway activation there is initial chromatin remodelling which allows access of the NER machinery to the DNA lesion (Schärer, 2013). Following chromatin remodelling one of the two NER pathways can take place; global genome-NER (GG-NER) or transcription coupled-NER (TC-NER) [Figure 1.6].

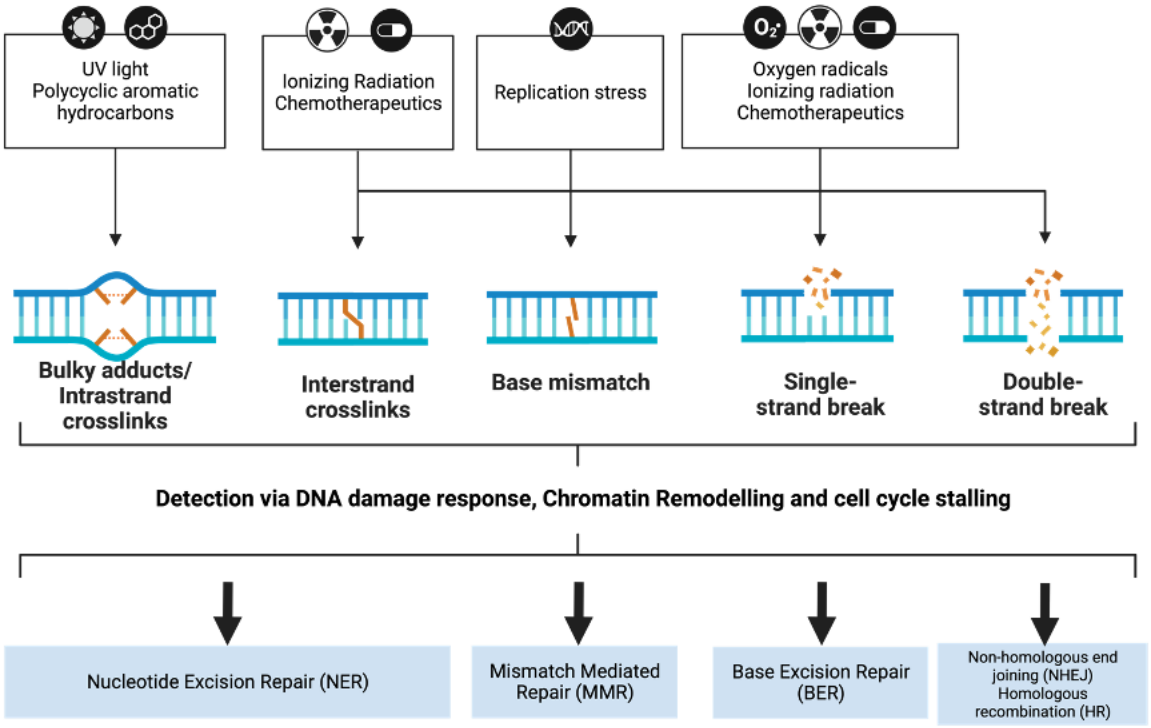


Figure 1.5 – Schematic displaying sources of DNA damage and the pathways which they are repaired. *Exogenous sources of DNA damage listed with the effects on DNA below. DNA aberrations are detected by the DNA damage response typically leading to the remodelling of chromatin and stalling of the cell cycle, after which each specific DNA aberration is repaired by the indicated repair mechanism.*

In GG-NER, regions of single stranded DNA (ssDNA), caused by disrupted base pairing due to bulky DNA lesions, are detected by the XPC/RAD23B complex and CETN2 (Masutani et al., 1994, Nishi et al., 2005). DNA lesions bound by XPC are then the substrate for the 10-protein subunit complex TFIIH which opens the region of DNA to allow cleavage and gap filling in a coordinated manner that does not further trigger the DDR. Cleavage is performed by XPF-ERCC1 at the 5' end and XPG at the 3' end of the ssDNA containing the lesion (Fagbemi et al., 2011). From here replication proteins are recruited and facilitate gap-filling and ligation. TC-NER is often initiated by the recruitment of CSA and CSB to lesion stalled RNA polymerase II. CSA and CSB recruit multiple downstream proteins to ultimately remove RNA polymerase II and allow access to the lesion site. Once available TFIIH is then recruited to the lesion site and downstream repair takes place as in GG-NER (Spivak, 2015).

1.5.2 BER

BER is predominantly active in G1 of the cell cycle, where SSBs or single-base damage DNA lesions are detected by DNA glycosylases [Figure 1.6]. These can either be mono-functional or

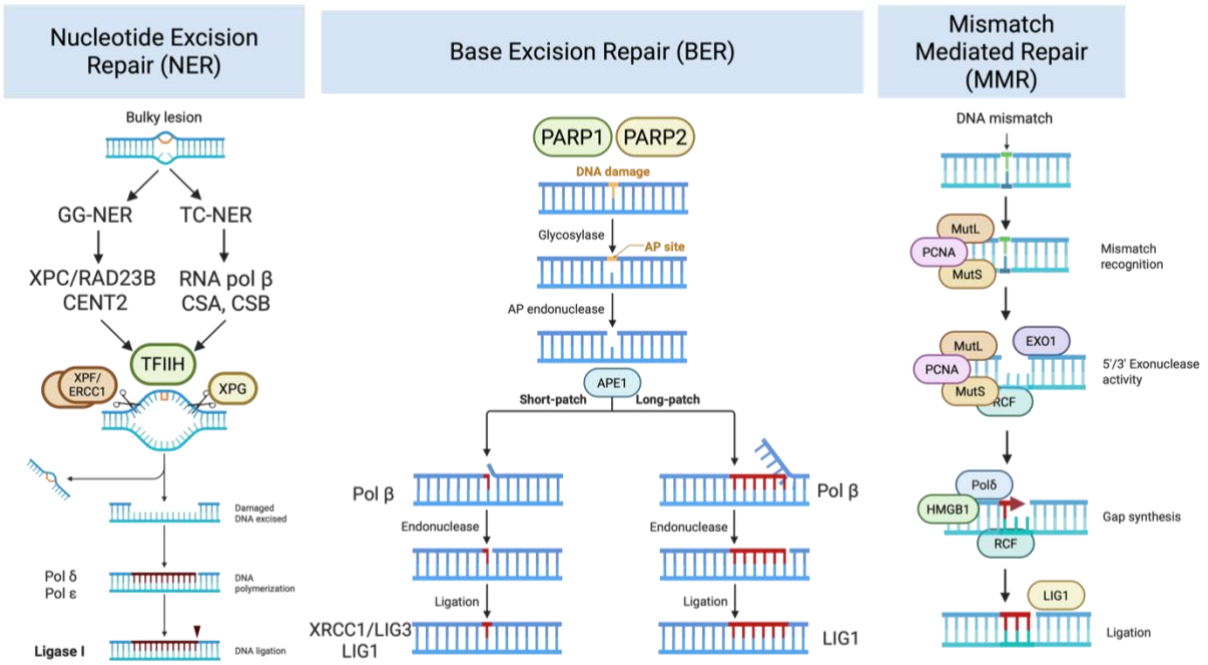


Figure 1.6 – Schematic indicating the main single strand DNA damage repair mechanisms. *Overviews of the three predominant mechanisms for resolving single stranded DNA damage, namely Nucleotide Excision Repair, Base Excision Repair and Mismatch Mediated Repair. Key proteins within each respective process are highlighted alongside general descriptions of the processes taking place at each step of the repair process.*

bi-functional, with each directing the short-patch or long-patch repair sub-pathways within BER (Dianov and Hübscher, 2013, Krokan and Bjørås, 2013). In short-patch repair, an abasic site is created by monofunctional DNA glycosylases which serve as a target site for APE1. APE1 then cleaves the DNA 5' to the lesion site and leaves a repair gap recognised by DNApolβ which fills the newly created single nucleotide gap. Following gap filling, LIG1 or the XRCC1/LIG3 complex ligates the DNA back together, resulting in complete repair of the DNA lesion (Almeida and Sobol, 2007). In long-patch repair bi-functional DNA glycosylases recognise the lesion site and allow targeting by APE1, however in this case utilising its 3' phosphodiesterase activity. This results in a repair gap again targeted by Polβ which displaces the strand containing the lesion and creating an overhang which is removed by flap endonucleases, with the resultant repaired strand being ligated by LIG1 (Akbari et al., 2009, Svilar et al., 2011).

1.5.3 Mismatch repair (MMR)

MMR is a post-replicative DNA repair mechanism that targets base mismatches caused during DNA replication, increasing its fidelity 100-fold (Kunkel, 2009). Key to the MMR pathway is the MSH family of proteins [Figure 1.6]. Specifically, the MutS (MSH2/MSH6) and MutL

(MLH1/PMS2) heterodimers are responsible for recognising the mismatch site, assisted by PCNA. From here the endonuclease activity of the MutS/MutL heterodimer directs EXO1 to the mismatch site which nicks 5' and 3' of the mismatch in a PCNA/RCF dependent manner. The gap created is then filled by Polδ, RFC and HMGB1 mediated strand synthesis with subsequent ligation by LIG1 (Kunkel, 2009).

1.5.4 SSB repair (SSBR)

Through failures in BER, NER or MMR in repairing DNA lesions or from error in DNA topoisomerase I (TOPI) activity, oxidative damage or abasic sites, DNA SSBs will be created. If left unrepaired these SSBs will result in collapse of DNA replication forks, leading to DSB formation, stalled transcription and PARP1 activation (Hegde et al., 2008). There are three SSBR pathways available to cells, long-patch, short-patch or TOPI-SSB, with the decision being dependent on the source of the SSB. In each the SSB is often detected by either PARP1/PARP2 which leads to the PARylation of itself and other substrate proteins facilitating the recruitment of downstream factors (Caldecott, 2014, Groelly et al., 2023). Long-patch SSBR is based on PARP1-mediated detection of SSBs, which will undergo cyclical poly(ADP) ribosylation of DNA at the break site. This allows 3' end processing by APE1, PNKP and APTX before 5' processing by FEN1, PARP1 and PCNA creating a ssDNA gap. Polβ and Polδ/ε then fill this ssDNA gap before final XRCC1 and PCNA dependent LIG1 ligation (Lan et al., 2004, Mortusewicz et al., 2006). Short-patch SSBR works to repair DNA lesions generated via BER, where following APE1 recognition of the DNA lesion is followed by the same end-processing steps as long-patch SSBR, with subsequent gap filling performed by Polβ, and LIG3 mediated ligation. The final SSBR pathway is TOPI-SSBR which shares the PARP1 dependent lesion recognition of long-patch SSBR but differs in the use of TDP-1 in end processing as removal of TOPI bound to the 3' end of DNA is necessary (Caldecott, 2014).

1.6 DNA DSB repair

If cells are unable to repair SSBs they will become DSBs which are much harder for cells to repair, hence pose a much higher risk to genome integrity and are toxic to cells (McKinnon and Caldecott, 2007). As such, methods of DBS repair are essential to cell homeostasis. There are two main mechanisms used in the repair of DSBs; non-homologous end joining (NHEJ) and homologous recombination (HR) [Figure 1.7]. DSBs themselves can be either two-ended,

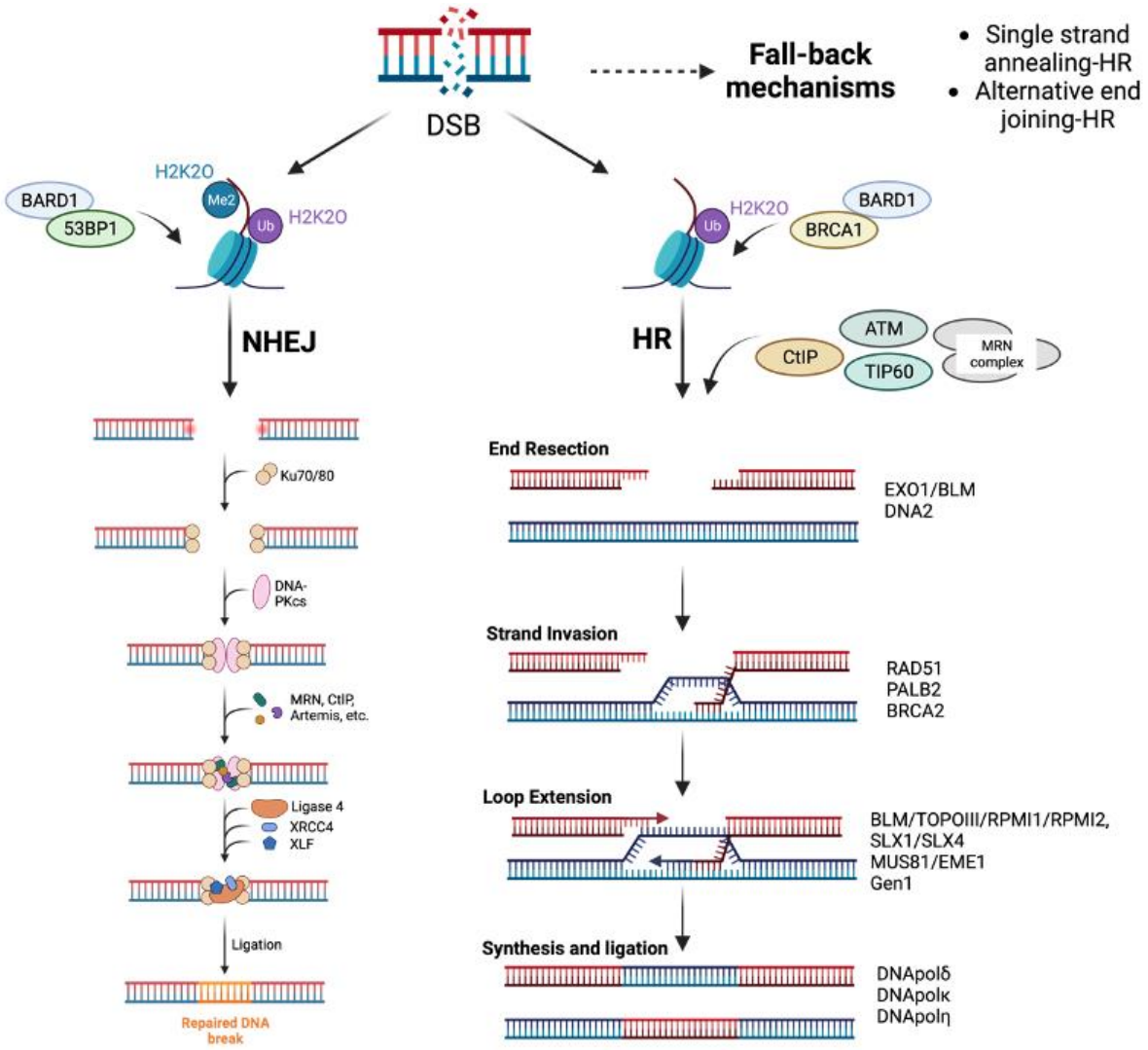


Figure 1.7 – Schematic of the main double strand DNA break repair mechanisms. *Overviews of both non-homologous end joining (NHEJ) and homologous recombination (HR), with associated histone marks used to direct double strand repair pathway choice. Key proteins at play in each step in the repair process highlighted. Alternate DNA damage repair mechanisms also highlighted as fall-back mechanisms available to cells if both NHEJ and HR are unavailable.*

where NHEJ will be utilised for the majority of repair, or one-ended caused by the collapse of replication forks in S phase. If NHEJ is utilised, chromosomal rearrangements can be formed by re-ligating the DNA ends of different chromosomes. As a result, NHEJ repair is repressed by HR proteins at replication forks, further highlighting the link between DNA damage repair and cell cycle signalling (Groelly et al., 2023). ATM/ATR are key kinases in the cellular response to DSBs as they facilitate the phosphorylation of multiple downstream substrates, including Chk1/Chk2, which once phosphorylated will down regulate CDK activity to halt cell cycle progression until the DNA damage is repaired (Blackford and Jackson, 2017). ATM and Chk2 work within G1 to further halt cell cycle progression through the phosphorylation and stabilisation of p53, ultimately promoting p21 expression which works to further inhibit CDKs.

Within the S/G2 phases of the cell cycle ATR and Chk1 act in checkpoint activation though the phosphorylation and inactivation of the CDC25 family of proteins, ultimately preventing early mitosis before DNA is repaired and replication is completed (Kastan and Bartek, 2004).

The initiation steps of both pathways upon DSB detection are similar and involve remodelling of chromatin around the DSB which results in a cascade of events including: activation of ATM/ATR; H2AX phosphorylation by ATM; MDC1 recruitment; ATM-mediated phosphorylation; and chromatin PARylation. From here either 53BP1 and BRCA1 can be recruited to further direct repair down either the NHEJ or HR pathways (Groelly et al., 2023). Repair of DSBs predominantly takes place during interphase through ATM/ATR activation as previously described. However, if DSBs are detected during mitosis neither 53BP1 nor BRCA1 are recruited. Instead CK2 phosphorylates MDC1, allowing interaction with TOPB1 and recruitment of CIP3A and TOPBP1, which form structures tethering DNA ends which allows proper chromosome segregation and cell cycle progression until the DSB can be repaired next interphase (Leimbacher et al., 2019).

1.6.1 DSB repair pathway choice

NHEJ based repair has the potential to cause chromosomal rearrangements if deployed in S phase and is highly toxic to cells. As a result, modifications to chromatin around DSB sites are used to direct specific repair pathways, with promotion of NHEJ during G1 and HR during S/G2 [Figure 1.7] (Beucher et al., 2009, Karanam et al., 2012). At the chromatin level, DSBs are recognised by BARD1 which is bound by 53BP1 and BRCA1, antagonising and promoting resection respectively. BARD1 and 53BP1 read H2AK15 ubiquitination, which is induced by DNA damage and is deposited by RNF168, and H2K20 methylation status (Botuyan et al., 2006, Fradet-Turcotte et al., 2013). 53BP1 protects DSBs from end resection, stopping the formation of ssDNA and resultant RAD51 nucleoprotein filaments essential for HR, promoting NHEJ over HR.

A similar chain of events takes place when directing DSB repair towards HR. In this case BRCA1/BARD1 complex recognises H2AK15 ubiquitination, but instead of methylated H4K20 the complex recognises unmethylated H4K20 (Nakamura et al., 2019, Becker et al., 2021). This is due to unmethylated H4K20 being enriched around the replication forks, and is only methylated on mature chromatin post-replication, meaning BRCA1/BARD1 binding and

directing HR during S phase where NHEJ-mediated repair could result in catastrophic outcomes. In the case of HR promotion 53BP1 will still bind to distal chromatin sites, preventing repression of any unnecessary long-range end resection during HR (Groelly et al., 2023).

1.6.2 NHEJ

NHEJ utilises a non-homologous template to direct DSB repair, which is why it is often thought to be an error prone process. In reality, NHEJ is used in for a majority of DSB repair, is highly efficient and rarely results in negative effects on the cell (Blackford and Jackson, 2017). Over-reliance on NHEJ through compromised HR, promoting its use at stalled replication forks results, in the commonly described error-prone repair.

In NHEJ, DSBs are recognised by 53BP1 which is then responsible for recruitment of other proteins essential for NHEJ, whilst also activating checkpoint proteins to stall the cell cycle to allow repair to take place and spatially tethering the two DNA ends together (Panier and Boulton, 2014). From here the KU70/KU80 heterodimer binds DNA ends, preventing end resection and functioning as a scaffold for the recruitment of downstream NHEJ proteins, such as DNA-PKcs, XRCC4, APLF, XLF and LIG4 (Chatterjee and Walker, 2017). Critical is the DNA dependent activation of DNA-PKcs, which not only displaces Ku but also auto-phosphorylates itself and surrounding proteins (Hammel et al., 2010, Dobbs et al., 2010). Following this XRCC4 binds DNA ends in place of the displaced Ku heterodimer and tethers ends, acting as an additional scaffold for further recruitment (Hammel et al., 2011, Andres et al., 2012). Once XRCC4 is in place a complex of NHEJ proteins including Artemis, APLF, PNKP, WRN, Aprataxin and Ku proteins begin end processing, with the subsequent gap filling performed by Polμ or Polλ, in a template dependent or independent manner, respectively, followed by LIG4 based ligation (Chatterjee and Walker, 2017). Most of the factors recruited in the latter stages of NHEJ, such as Artemis, are not essential for NHEJ mediated repair, although help promote repair accuracy (Groelly et al., 2023).

1.6.3 HR

HR utilises strand invasion for DNA template-based repair, resulting in highly accurate repair of DNA damage in S/G2 phase of the cell cycle. The MRN complex, made up of MRE11/RAD50/NBS1, initiates HR at the site of DSBs which then recruits TIP60 and ATM (Stracker and Petrini, 2011, Sun et al., 2005). TIP60 activates ATM, which in turn phosphorylates H2AX, providing a binding sire for MDC1. MDC1 is then phosphorylated by ATM, allowing it to act as a scaffold for further recruitment of the E3 ubiquitin ligases RNF8 and RNF168, resulting in the ubiquitination of H2AX. H2AX ubiquitination provides a site for binding of BRCA1, preventing the binding of 53BP1, and facilitating ubiquitination of CtIP and subsequent recruitment of RPA and RAD51 (Chatterjee and Walker, 2017). End resection of the 3' overhang is performed by the MRN complex and CtIP followed by long-range resection by the EXO1/BLM complex and DNA2 (Chen et al., 2008a). End resection only takes place in S/G2 phase of the cell cycle due to CDK-dependent phosphorylation of CtIP initiating the process (Huertas and Jackson, 2009). This is especially important in diploid cells as it prevents the use of the homologous chromosome as a template rather than the sister chromatid. The 3' overhang created from end resection is initially bound by RPA, which is subsequently replaced by RAD51 to create nucleoprotein filaments. Using these filaments PALB2 and BRCA2 initiate strand invasion into the nearby DNA duplex forming a D-loop, with RAD54 and RAD54B playing key roles in this invasion through removal of RAD51 and template-based DNA synthesis by DNApol$\delta/\kappa/\nu$. Ultimately, a Holliday Junction is formed and processed by the complex of BLM/TOPOIII/RPMI1/RPMI2, SLX1/SLX4 complex, MUS81/EME1 complex and Gen1 endonuclease (Chatterjee and Walker, 2017).

There are other forms of HR available to cells, such as single strand annealing-HR and alternative end joining-HR, which rely on RAD52/PolQ activity respectively. These are often used as back-up pathways when the NHEJ and HR pathways are compromised as they are both highly mutagenic processes. However, the results are much more favourable than the effects of unrepaired DSBs progressing into mitosis, as upon chromosome segregation they can cause chromothripsis, formation of micronuclei and breakage-fusion-bridge cycles (Rubin et al., 2016, Groelly et al., 2023).

1.7 DNA repair in PC and DNA repair as a therapeutic target

1.7.1 DNA repair defects in PC

Analysis of PC clinical samples has revealed an increased prevalence of mutations, either germline or somatic, within genes playing key roles in DNA repair (Mateo et al., 2017). The presence of mutations in DNA repair genes increases from 5-10% in localised PC to 20% in advanced mCRPC, making these repair pathways a target for research and a potential avenue for developing new therapeutic options for PC (Grypari et al., 2023).

From a whole-exome and transcriptome sequencing of 150 matched normal versus mCRPC biopsies performed by the SU2C-PCF International Dream Team, essential DNA repair genes were seen to be inactivated in around 23% of cases (Robinson et al., 2015). Alongside this, a TCGA study revealed that around 19% of primary prostate tumours harboured alterations in key DNA repair pathway genes playing key roles in HR such as *RAD51*, *ATM*, *BRCA1* and *BRCA2* (TCGA, 2015, Gulhan et al., 2019). Further, defects in DDR pathways have been reported to increase when comparing primary tumour samples to metastatic samples, with an increase in DDR mutations from 10% to 27% (Lozano et al., 2021). Through germline analysis of 221 patients, Abida et al., 2017 reported 27% of patients demonstrating either germline or somatic alterations in key HR genes *BRCA1*, *BRCA2*, *ATM* and *CHEK2*, which was mirrored in work presented by Pritchard et al., 2017 which conducted similar germline analysis reporting DDR defects in 11.8% of patients (Abida et al., 2017, Pritchard et al., 2016).

These data together depict a new substantial sub-population of PC patients with DNA repair defects. Germline mutations in *BRCA2* are seen in PC and are associated with worse prognosis. This poor prognosis due to mutation in *BRCA2*, either via germline or somatic means, is through defective HR repair of DBSs, leading to an increase in genomic instability. Increased genomic instability can lead to increased copy number alterations within cells, for example *MYC* amplification which leads to increased disease progression and aggressiveness and promotes CRPC (Qiu et al., 2022, Grypari et al., 2023).

Cross-regulation between DNA damage repair and AR signalling could be relevant for PC progression (Goodwin et al., 2013). DSBs can be generated as a direct consequence of AR-mediated transcription that can lead to rearrangements across the genome (Lin et al., 2009). The AR can bind to chromosomal sites and create a proximity to distant chromatin loci that can result in translocations like the *TMPRSS2-ERG* translocation seen in many PC clinical

samples (Lin et al., 2009, Mani et al., 2009). Within these tumours, mutations are seen in the *SPOP* gene (discussed further in Section 1.9), which has roles in stabilising proteins such as the AR and its transcriptional regulators. The elevated level of DNA aberrations observed as a result of oncogenic *SPOP* mutations are likely due to compromised HR; and as a result, tumours rely upon the error prone NHEJ as the predominant repair mechanism (Boysen et al., 2015).

Recently the DNA damaging effects of new methods of PC treatment known as bi-polar androgen treatment (BAT) have been investigated. BAT brings about supraphysiological androgen (SPA) concentrations and subsequent increased AR signalling through intermittent dosing with testosterone, causing oscillation between supraphysiological levels and castrate levels of circulating testosterone (Chatterjee et al., 2019). Ultimately this has been shown to increase DNA damage within PC cells and bring about growth arrest and senescence (Chatterjee et al., 2019, Markowski et al., 2022). Interestingly, SPA has been shown to cause a significant increase in the downregulation of HR-associated genes over that of seen in response to ADT. Further, SPA in AR overexpressing PC cells have a lack of ability to complete NHEJ DNA repair due to inability to displace and subsequent retention of DNA-PKcs on chromatin (Chatterjee et al., 2019). As a result, patients with HR-associated genetic aberrations (specifically *BRCA2* or *TP53* inactivating mutations) show favourable responses to BAT, due to compromised NHEJ and HR repair processes (Chatterjee et al., 2019, Markowski et al., 2022). Also of note, SPA has been shown to overcome and ultimately repress AR-V-mediated activation of HR DNA repair pathway (Chatterjee et al., 2019). This has led to investigation into the combinations of BAT and PARP inhibition to introduce both increased DNA DSBs alongside compromised DDR capacity of cells (Chatterjee et al., 2019). With promising initial results, ongoing clinical trials are in place to investigate BAT in mCRPC patients with HR or *TP53* mutations (NCT03522064)(Markowski et al., 2022). Alongside this, a phase II clinical trial investigating combinations of Olaparib and BAT have been conducted, displaying high response rates and extended progression-free survival of patients in response to the combination, regardless of HR gene mutational status (Schweizer et al., 2023).

1.7.2 Therapeutically targeting DNA damage repair defects in PC.

The identification of this large patient subgroup with compromised DNA repair mechanisms, due to highly prevalent mutations in DNA repair genes, presents an ideal sub-population of

patients where remaining intact DNA repair mechanisms can be therapeutically targeted, termed synthetic lethality, providing novel therapeutic sensitivities in advanced PC patients.

One of the most well-known examples is exploiting *BRCA1/2* mutations; a defect that results in compromised HR-mediated DNA repair. *BRCA1/2* mutations confer genome instability due to an accumulation of DNA lesions caused by failure to repair collapsed replication forks by HR and instead relying on the error-prone NHEJ pathway (Thompson et al., 2002). This accumulation of genetic aberrations as a result of *BRCA1/2* mutations leads to enhanced tumorigenesis. The reliance of on error-prone DSB repair mechanisms can be exploited by further increasing the amount of unrepaired SSBs being translated into DSB through inhibition of key SSB repair proteins PARP1/2. Through PARP inhibition (PARPi) there will be a lack in detection, and as a result repair, of SSBs leading to increased DBS burden when these SSBs are encountered by replication forks, leaving one-ended DSBs that must be repaired by HR. However, with compromised HR in the background of *BRCA1/2* mutations, NHEJ has to be used in repair of DSBs resulting in highly toxic gene rearrangements. Ultimately this toxicity leads to cell death, presenting a synthetic lethality in PC cells harbouring *BRCA1/2* mutations in response to treatment with PARP inhibitors (Bryant et al., 2005, Farmer et al., 2005).

Other potential druggable targets within DNA damage repair pathways include the repertoire of kinases which function in the sensing and repair stages of the DDR. For example, ATM is one of the main kinases at play in DNA repair, with mutations or loss of *ATM* activity promoting NHEJ as the main DSB repair pathway. Loss of ATM activity also confers further sensitivity to PARPi or TOPI inhibition (TOPIi) as both promote the generation of one-ended DSBs and with loss of ATM activity, much like in *BRCA* mutant backgrounds, there is promotion of NHEJ and resultant toxic gene fusions and cell death (Bryant and Helleday, 2006). Interestingly mutations within key NHEJ proteins, such as XRCC4 or LIG4, confer resistance to PARPi and TOP1i due to cells being much more capable of repairing the one-ended DSBs caused (Groelly et al., 2023).

With ATR playing key roles in S-phase to induce HR-mediated repair of DSBs concurrent with DNA replication, its importance is highlighted in cancer further due to the increased incidence of stalled replication forks. As such, there is an increased reliance on ATR signalling in cancer making it a prime candidate for therapeutic intervention. ATR inhibitors (ATRi), which attenuate the catalytic activity of the kinase, results in failure of cells to undergo cell cycle arrest and an increase in replication fork stalling. This leads to DSBs when cells progress into mitosis, resulting in chromosomal breaks and cell cytotoxicity. As such, cancers with increased

replication stress (for example due p53 deficiencies and/or cyclin E overexpression) or *ATM* mutations are exquisitely sensitive to ATRi. ATRis also act synergistically with PARP inhibition regardless of *BRCA* mutation status (Toledo et al., 2011, Groelly et al., 2023). Using modalities to cause DNA damage, such as ionising radiation (IR), in combination with ATRi can enhance the effects of ATRi alone, and has been demonstrated in triple negative breast cancer (Tu et al., 2018). Further, ATRi potentiates radiation induced-interferon signalling, providing an avenue of combination therapy with immunotherapy. So far, combination of ATRi has shown synergistic anti-cancer effects when combined with immune checkpoint blockade via PDL1 targeted therapies in mouse models and has been taken further to ongoing clinical trials (Dillon et al., 2019, Feng et al., 2020, Tang et al., 2021). Another avenue for eliciting similar responses to ATRi is inhibition of its downstream target Chk1. As such, Chk1 inhibition (Chk1i) shows enhanced anti-tumour effects in cancers with increased replication stress, whilst also showing synergistic anti-tumour effects in combination with DNA damaging agents, PARPi and anti-PDL1 based therapeutics in non-PC backgrounds (Booth et al., 2013, Sen et al., 2019).

ATM, ATR and Chk1 targeting inhibitors all target HR DNA damage repair and hence promote error prone HR back-up pathways or use of NHEJ in cell cycle stages to bring about toxic effects.

Therapeutically targeting DNA-PKcs, the catalytic subunit of DNA-PK, provides a method of targeting and reducing NHEJ mediated DNA repair capacity in cells. As a result, cancer cells lacking DNA-PKcs activity are sensitive to DNA damaging agents, with similar effects seen through DNA-PKcs inhibition (DNA-PKcsi); with DNA-PKcsi treatment further potentiating the anti-cancer effects of radiotherapy (Fok et al., 2019). Further, dual loss of ATM and DNA-PKcs activity is lethal to cells as they rely wholly on the highly toxic fall-back DSB repair mechanisms due to loss of both NHEJ and HR mediated DNA repair (Riabinska et al., 2013). This has potential to be recapitulated in response to combinations of ATMi and DNA-PKcsi as a dual treatment to cause a synthetic lethality in cells. As such DNA-PKcsi has been a candidate for multiple clinical trials, either as single agent treatments or in combination with other DNA damage-inducing therapeutics (Groelly et al., 2023).

Although much work has been done in developing inhibitors targeting key kinases in DNA damage repair pathways, there is ongoing work to identify new DNA repair associated targets for therapeutic intervention. Current clinically approved applications of DDR targeting agents in PC and currently active clinical trials for DDR inhibitors are summarised in Table 1.1 and Table .2 respectively (Li et al., 2023b). New targets could lie within upstream regulators of DNA

repair pathways, with some commonly mutated in PC, such as *SPOP,* which have seen to further potentiate the effects of current DNA repair targeting therapeutics.

Table 1.1 – Currently approved DDR targeting agents in PC alongside their respective biomarkers. *Adapted from Li et al., 2023.*

Drug	**Indication**	**Biomarkers**
Olaparib	Pre-treated metastatic castration-resistant prostate cancer; monotherapy	HR repair gene muatant
Rucaparib	Androgen receptor-directed therapy and a taxane-based chemotherapy pretreated metastatic castration-resistant prostate cancer; monotherapy	*BRCA* mutant

Table 1.2 - Currently active clinical trials for DDR inhibitors in advance solid tumours and PC.
Adapted from Li et al., 2023.

Pathway	Target	Drug	Combination	Phase	Cancer type	Clinical Trial Identifier
SSBR	PARP1-selective	NMS-03305293	Monotherapy	1	Advanced Solid Tumour	NCT04182516
		AZD5305	Monotherapy; Paclitaxel; Carboplatin; Trastuzumab Deruxtecan; Datopotamab Deruxtecan	1/2	Advanced Solid Tumour	NCT04644068
			Ceralasertib	1/2	Advanced Solid Tumour	NCT02264678
			Datopotamab deruxtecan or Durvalumab + Datopotamab deruxtecan	1/2	Advanced Solid Tumour	NCT05489211
			Enzalutamide; Abiraterone acetate; Darolutamide	1/2	Metastatic prostate cancer	NCT05367440
		AZD9574	Monotherapy; Temozolomide	1/2	Advanced Solid Tumour, breast cancer, glioma	NCT05417594
DSBR	Polθ	ART4215	Monotherapy; Talazoparib; Niraparib	1/2	Advanced Solid Tumour	NCT04991480
	RAD51	CYT0851	Monotherapy; Gemcitabine; Capecitabine; Rituximab and Bendamustine	1/2	Advanced Solid Tumour, B-Cell Malignancies	NCT03997968

1.8 SPOP in DNA repair and its uses as a potential PC biomarker

1.8.1 SPOP structure

SPOP is a 42 kDa 374 amino acid protein comprised of three main domains that was first identified in 1997 [Figure 1.8] (Nagai et al., 1997). The N-terminal meprin and TRAF-C homology domain (MATH) domain within SPOP spans residues 28-166 and is made up of multiple antiparallel β-sheets. The resultant cleft formed by these sheets contains SPOP binding motifs with which substrates can interact (Maekawa and Higashiyama, 2020). Also present in SPOP is the broad-complex, tram-track and bric-a-brac (BTB) domain, spanning residues 186-300 which facilitates the interaction with the E3 ligase scaffold protein Cul3 (Zhuang et al., 2009). The final domain present in SPOP is the BTB and C-terminal Kelch (BACK) domain which spans residues 300-359 and acts as a dimerisation domain, which indicates the possibility for self-association (Bouchard et al., 2018, Cuneo and Mittag, 2019).

1.8.2 SPOP function

One key process essential in maintaining cell homeostasis is protein degradation. Aberrant activity in this process can lead to accumulation of proteins, which can potentiate disease development and is commonly seen in cancer [Figure 1.9]. 80% of all intracellular protein degradation is mediated by the ubiquitin proteosome system (UPS) which is responsible for maintaining key processes such as cell proliferation and apoptosis (Chen and Chen, 2016). The UPS works through a regimented multi-step process facilitated by E1 proteins which activate ubiquitin, E2 proteins responsible for ubiquitin conjugation and E3 ligase proteins which are essential for substrate recognition and specific degradation (Bedford et al., 2011). Of the E3 ligases, the Cullin-RING E3 ligase (CRL) family is the largest and are made up of a complex of proteins, generally containing a Cullin (CUL) protein acting as a scaffold, a Ring-box protein (RBX1/2) which facilitates E2 recruitment, a substrate recognition protein (SRS) and an adaptor protein responsible for linking the SRS to the CUL scaffold (Buetow and Huang, 2016). Most of the CRL family are made up of these 4 subunits, however the CRL3 complex is made up of 3 in which CUL3 acts as a scaffold for RBX1 and a BTB protein which acts as both the adaptor and SRS protein. Commonly, SPOP acts within this CUL3 complex as the site of substrate recognition (Song et al., 2020).

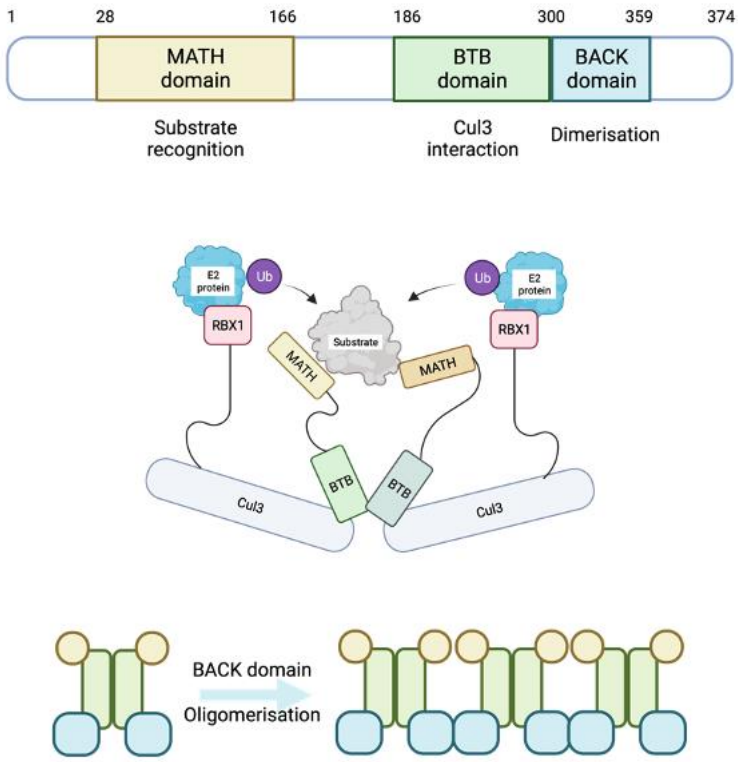


Figure 1.8 – Schematic representation of the SPOP protein domain structure, how it functions within the Cul3 ubiquitin complex and its dimerisation and oligomerisation capabilities. Top *- SPOP amino acid structure with annotated meprin and TRAF-C homology domain (MATH) domain between 28-166aa, broad-complex, tram-track and bric-a-brac (BTB) domain between 186-300aa and BTB and C-terminal Kelch (BACK) domain between 300-359aa.* ***Middle*** *– SPOP within the Cul3 ubiquitin ligase complex, binding to the Cul3 scaffold and dimerising with other SPOP proteins through BTB domain interactions. SPOP functions as the substrate binding interface of the Cul3 ubiquitin complex, binding substrate proteins through MATH domain interactions with SPOP-binding motifs present on substrate proteins. Ultimately allows E2 proteins to deposit ubiquitin onto substrate proteins.* ***Bottom*** *– SPOP can form oligomers comprised of multiple SPOP proteins through BACK domain interactions.*

SPOP recognises a specific 5 amino acid long φ-π-S-S/T-S/T consensus sequence, where φ is a nonpolar and π is a polar amino acid. These regions on substrate proteins are often found in disordered regions allowing accessibility for SPOP binding (Song et al., 2020). Using the BTB domain, SPOP is able to dimerise, whilst through the self-association BACK domain, SPOP is able to form polymers which enhance its ubiquitination ability as it is able to interact with multiple SPOP recognition motifs on substrate proteins. (Zhuang et al., 2009, Cuneo and Mittag, 2019). This ability to form dimers is the reason for SPOP mutations showing dominant negative effects, as seen in PC patients as they do not display loss of heterozygosity, meaning that both wild-type and mutant *SPOP* is present. In this case patients still lose SPOP function and accumulation of SPOP substrates, due to the presence of mixed polymers and heterodimers, resulting in reduced activity and disease progression (Cuneo and Mittag, 2019, Zhang et al., 2023). SPOP dimerization also provides a mechanism of regulation of its ubiquitination activity, through interaction with SPOP-like (SPOP-L) protein. SPOP-L shares high

sequence homology with SPOP, including the ability to form heterodimers with SPOP. These heterodimers are drastically less active than SPOP homodimers, thus providing regulation of SPOP activity through expression of SPOP-L (Clark and Burleson, 2020). Further, SPOP has the ability to self-ubiquitinate and direct itself for degradation, indicating a potential negative feedback loop allowing for greater regulation of SPOP activity (Bunce et al., 2008).

SPOP has been found to play key roles in many different cellular processes including development, apoptosis, cellular senescence and, importantly, DNA damage repair. In development, high levels of SPOP expression has been found to be essential for embryo implantation and endometrial decidualisation, whilst also targeting GLI2 and GLI3 which are downstream effectors of sonic hedgehog signalling. SPOP-mediated degradation of DAXX has been linked to apoptosis signalling, as this reverses repression of p53 and ETS1 transcription, thus inducing apoptosis. SENP7 degradation by SPOP has also been implicated in promoting cellular senesce (Zhu et al., 2015).

SPOP is essential for the mRNA expression of many DNA repair related genes (e.g., *ATR*, *BRCA2*, *ChK1*, *RAD51*), which was supported by proteomics data (Hjorth-Jensen et al., 2018). This study also found wild-type SPOP (SPOP-WT) interacted with proteins involved in many other cellular processes such as splicing (Hjorth-Jensen et al., 2018). SPOP was initially implicated in DNA damage repair signalling *in vitro* by Zhang *et al.,* and was later confirmed to be functional in this pathway *in vivo* (Bernasocchi et al., 2021). Further, SPOP foci are seen to accumulate in an ATM dependent manner around sites of DNA damage with SPOP depletion being linked with decreased RAD51 foci formation and reduced Chk1 activation (Zhang et al., 2014a). Further, in response to DNA damage, ATM has been found to phosphorylate S119 of SPOP resulting in SPOP-mediated degradation of cell division-related proteins and non-degrative ubiquitination of 53BP1. As such, this promote HR for DSB repair over NHEJ (Xiao et al., 2021). Ultimately, SPOP has been found to be a key regulator in HR of DNA damage, and mutations within the SPOP MATH domain resulted in cells displaying a '*BRCA*-like' phenotype (Boysen et al., 2015).

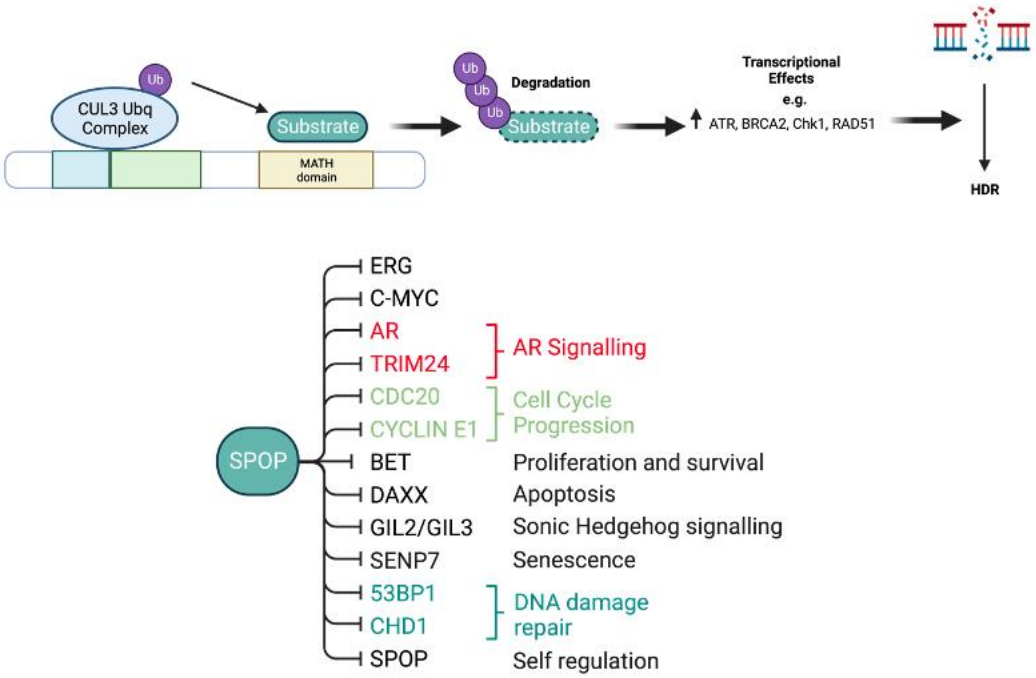


Figure 1.9 – Schematic representing how SPOP regulates multiple faucets of cell homeostasis through substrate degradation. Top *– An example of SPOP bringing about cellular effects, namely SPOP mediated substrate degradation leading to increases in transcription of key homologous recombination (HR) genes causing the promotion of HR for DNA repair.* ***Bottom*** *– List of SPOP substrate proteins and the cellular processes they play roles in.*

1.8.3 SPOP Regulation

SPOP activity itself can be regulated by PTMs that result in either stabilisation of the SPOP protein or through inhibiting its activity. SPOP can be phosphorylated by LIMK2 and AURKA which functionally decreases both SPOP stability and activity. Both kinases are known SPOP substrates, indicating the existence of a potential negative feedback loop to regulate SPOP activity in cells (Nikhil et al., 2020, Nikhil et al., 2021). CDK4/6 also phosphorylate SPOP, stabilising it and further potentiating its activity. Use of CDK4/6 inhibitors has shown reduced CDK-mediated SPOP phosphorylation which promotes SPOP ubiquitination and degradation by APC/C via CDH1 activation (Zhang et al., 2018a). In response to ATRi, downregulation of ATR-Chk1 signalling results in CDK1 activation and subsequent SPOP phosphorylation further stabilising its protein levels (Tang et al., 2021). This hints at a potential compensatory mechanism to stabilise cell's DNA repair capabilities.

In terms of regulation of SPOP, through transcriptional or translational changes, not much is known in a PC setting, although there have been large amounts of work done in relation to other cancer backgrounds. Notably, the *SPOP* gene displays an increased transcriptional response to hypoxic stress in clear cell renal carcinoma. Conversely, hypermethylation of CpG sites in the SPOP promoter causes transcriptional repression leading to metastasis and disease

progression in colorectal cancer. Furthermore, several miRNAs have been implemented in silencing SPOP expression in breast cancer, such as miR-145 (Song et al., 2020). Alongside these direct effects on SPOP levels, there is growing evidence of upstream regulators of SPOP which have been implicated in PC progression. One such example is G3BP1 which acts as an upstream negative inhibitor of SPOP-mediated protein degradation through direct competitive binding to SPOP. In PC, this results in increased AR signalling which in turn can promote G3BP1 expression in a positive feed-forward loop to maintain AR signalling and drive tumour progression (Mukhopadhyay et al., 2021).

1.9 SPOP and PC

SPOP mutant expression can result in tumour formation, due to the changes in substrate protein turnover [Figure 1.10]. Particularly, missense mutations within the SPOP substrate binding MATH domain have been seen to occur in PC (Maekawa and Higashiyama, 2020). These MATH domain mutations often interfere with the interaction of SPOP and its substrate proteins, ultimately resulting in the accumulation of substrate proteins such as the AR, which can have oncogenic effects (Maekawa and Higashiyama, 2020). Examples of these PC associated mutations include the Y87C and the F133V mutation which assert dominant negative effects resulting in more frequent gene rearrangements in PC when compared to PC samples with unmutated SPOP (Theurillat et al., 2014, Maekawa and Higashiyama, 2020).

Mutations in *SPOP* are the most common non-synonymous mutations in PC and is seen mutated in 4-28.6% and downregulated in 25.2-93.5% of PC patient samples (study dependent) (Clark and Burleson, 2020). With both these resulting in repressed SPOP activity, it is evident that SPOP plays key roles as a tumour suppressor. Further, *SPOP* mutations are highly correlated to increased rearrangements and significantly higher copy number alteration burden, thus are highly correlated with negative prognosis and increased tumour progression (Boysen et al., 2015). This hints at a rising new sub-class of PC where *SPOP* mutations happen early in PC and result in genomic instability that enhances PC progression through deregulation of the wide range of SPOP substrates.

With SPOP being implicated in DNA damage repair, it is suggested that its function in these pathways is impaired when PC-associated SPOP mutants are expressed (Mateo et al., 2017). The F133V mutation displays a similar transcriptional response to *BRCA1* inactivation, which is consistent with SPOPs role in DNA repair (Maekawa and Higashiyama, 2020). SPOP-WT promotes DNA repair via HR; with mutant SPOP-expressing cells, as observed in PC samples, lacking HR and promoting the error prone NHEJ (Boysen et al., 2015). This offers an explanation for the increased gene rearrangements seen in SPOP mutant PCs, with NHEJ resulting in increased genomic instability and increased probability of rearrangement.

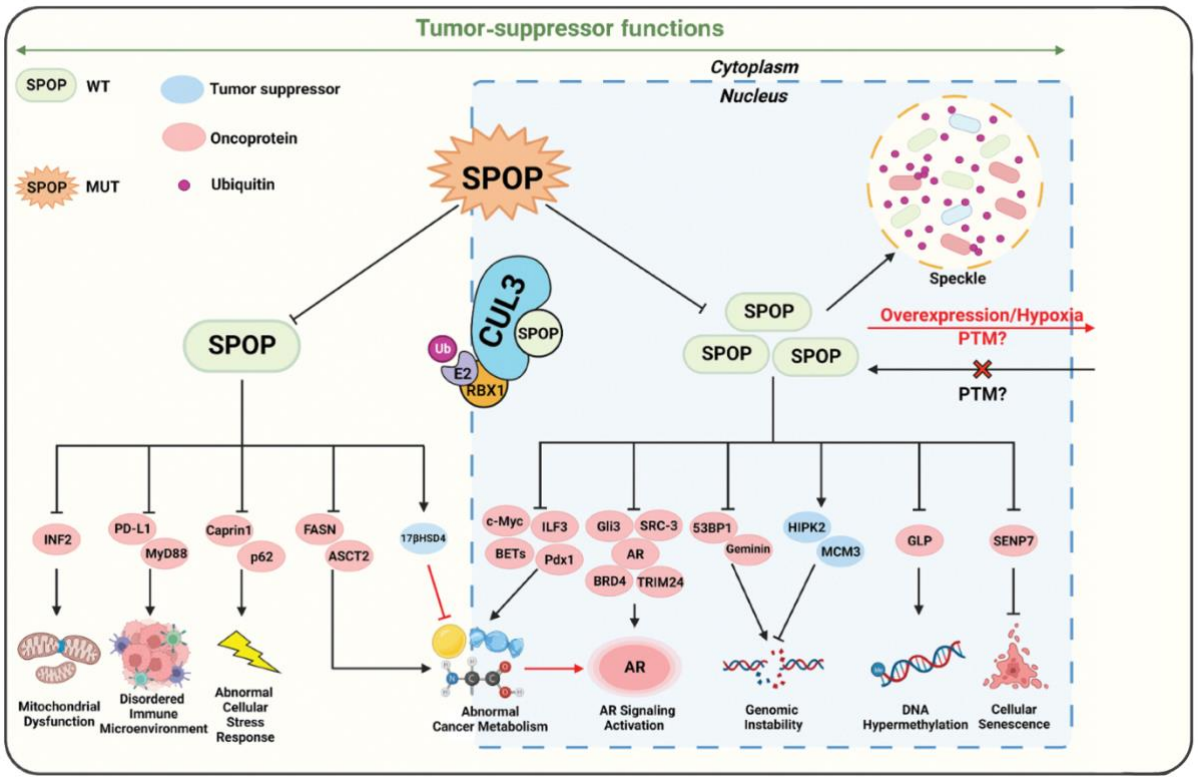


Figure 1.10- SPOP substrate proteins and the effects of cancer associated SPOP mutations on the pathways they regulate. *Adapted from Zhang et al., 2023*

SPOP mediated degradation has been heavily implicated in regulation of AR signalling, through direct degradation of AR or through degradation of AR coregulators, such as SRC-3, TRIM24, BRD4, Gil3 and p300 (Zhang et al., 2023). SRC-3 was one of the first identified SPOP substrates and functions as a co-activator of AR. In the case of SPOP-mutant PC (SMPC) there is stabilisation of SRC-3 levels resulting in increased transcriptional activity of the AR, resulting in AR driven cancer progression and tumorigenesis. SRC-3 stabilisation also promotes cancer progression independent of AR activity through enhanced PI3K/mTOR signalling. Typically, the AR and AKT pathways exert negative feedback on each other to reduce tumorigenesis. However, in SMPC this negative feedback is decoupled and thus sees a two-fold increase in oncogenic signalling (Geng et al., 2013, Blattner et al., 2017). Furthermore, AR itself is an SPOP substrate due to an optimum SPOP binding motif being present in the AR hinge region. As a result, SMPC sees AR accumulation resulting in increased AR signalling. Also of note, AR-Vs associated with PC progression and treatment resistant lack the AR hinge region, which harbours an SPOP binding motif, making AR-Vs unable to be targeted for ubiquitination by SPOP and subsequent degradation (An et al., 2014). AR signalling can be enhanced in the

presence of SPOP mutations through the accumulation of another substrate, TRIM24. SMPC conditions will result in TRIM24 accumulation and promotion of PC cell proliferation in low androgen conditions through enhanced AR signalling, giving an example of a mechanism of PC reoccurrence following ADT. TRIM28 is an upstream regulator of TRIM24 and binds directly to TRIM24 and increases its chromatin occupancy (Clark and Burleson, 2020). Upon TRIM24/TRIM28 binding, the SPOP binding motif on TRIM24 is blocked, resulting in competitive binding of both SPOP and TRIM28 playing roles in regulation of TRIM24 activity, reliant on TRIM28 levels. Due to aberrant regulation of AR signalling in SMPC, patient data has shown hypersensitivity and increased overall survival in response to ADT in early disease (Abida et al., 2017, Grbesa et al., 2021), enzalutamide and abiraterone treatments in late stage mCRPC (Bernasocchi et al., 2021, Swami et al., 2022).

Chromodomain helicase DNA-binding protein 1 (CDH1) has also been implicated in PC. CHD1 is a protein essential for DNA repair via HR and has also been implicated in regulating AR occupancy (Zhu et al., 2021). *CDH1* deletions are commonly seen in PC, which is of interest as SPOP and CDH1 have both been implicated as regulators of the DNA damage response (Boysen et al., 2015, Maekawa and Higashiyama, 2020). *SPOP* mutations are also highly associated with *CHD1* deletion, with defects in both genes resulting in a synergistically higher level of DNA damage in prostate epithelial cells, indicating that these proteins work in tandem to maintain genome integrity (Zhu et al., 2021).

Another PC driver ERG is a SPOP substrate, specifically through the SPOP binding motif present in the N-terminal domain of ERG (An et al., 2015, Gan et al., 2015). Genomic rearrangements, including the *TMPRSS2-ERG* translocation, are a key mechanism that drives PC tumorigenesis through increased expression of the *ERG* oncogene (Boysen et al., 2015). These fusions result in the truncation of the ERG N-terminal domain and this loss of the SPOP binding motif resulting in the inability for SPOP-mediated degradation of ERG. In PC there is mutual exclusivity seen between *TMPRSS2-ERG* fusions and SPOP-mutations (Bernasocchi et al., 2021). This presents two separate mechanisms of ERG overexpression, either through SMPC resulting in ERG accumulation or through the *TMPRSS2-ERG* fusion event.

Other SPOP substrates that accumulate in SMPC to promote tumorigenesis are cMYC and BET proteins. cMYC is commonly altered in PC in which increased activity enhances cell proliferation and survival. In SMPC there is heightened cMYC signalling and thus enhanced tumorigenesis. This was confirmed through gene set enrichment analysis (GSEA) of both SPOP-

MT- and cMYC-overexpressing PC cell line transcriptomics, which demonstrated considerable transcriptomic overlap suggesting both cMYC- and mutant SPOP-overexpression result in the effects on similar processes (Geng et al., 2017). Elevated levels of BET proteins BRD2, BRD3 and BRD4 play roles in cMYC independant activation of the AKT-mTOR signalling pathway through *RAC1* upregulation, thus in SMPC there is enhanced AKT-mTOR signalling resulting cell proliferation and survival (Zhang et al., 2017).

DNA hypermethylation caused by accumulation of SPOP substrates is now described as a major feature in SMPC (Robinson et al., 2015). The histone methylase GLP functions in the recruitment of DNA methyltransferases to chromatin. GLP is an SPOP substrate, hence in SMPC an elevation in GLP levels is observed. This results in heightened recruitment of DNA methyltransferases to chromatin, with a resultant increase in silencing of key tumour suppressor genes, such as *FOXO3*, *GATA5*, and *NDRG1* (Zhang et al., 2021). SPOP has also been implicated in regulating the immune microenvironment through MyD88, RAS/MAPK signalling alterations and importantly SPOP-mediated PD-L1 ubiquitination and degradation. SPOP ubiquitinates and degrades PD-L1 to facilitate suppression of immune evasion, thus SMPC allows immune escape through accumulation of PD-L1 (Zhang et al., 2018a).

As previously mentioned, SPOP plays roles in multiple other cellular processes such as apoptosis and cell cycle regulation. Its role in apoptosis has been linked to its substrate DDIT3 which plays roles in endoplasmic reticulum (ER) stress mediated apoptosis (Zhang et al., 2014b). In SMPC, DDIT3 accumulates and results in compromised ER stress-induced apoptosis which promotes cell survival and heightened tumorigenesis. Cell cycle perturbations in SMPC can be attributed to the SPOP substrate CDC20, the accumulation of which in SMPC promotes cell cycle progression through interaction with the anaphase promotion complex (Wu et al., 2017). Alongside this SPOP targets cyclin E1 for degradation, thus in SMPC its interaction with CDK2 is lost causing further effects on the cell cycle (Ju et al., 2019).

As well as the processes mentioned above, SMPC causes aberrant activity in the cell stress response, cancer metabolism and cellular senescence. With a high number of PC patients presenting with *SPOP* mutations resulting in loss of functional protein and aberrant activity in multiple cellular processes, the emergence of a new subset of PC cancer patients with increased opportunities for treatment is clear. With such a range of impacts seen in response to *SPOP* mutations there are many potential therapeutic interventions which hold the

potential to enhance patient response to current PC treatments whilst opening the door to applying new treatments not previously implicated in a PC setting.

1.10 PC biomarkers, their detection, and DNA Nanotechnology

1.10.1 Cancer Biomarker Detection

In most cancer cases, early diagnosis is strongly correlated with increased survival. Thus, identifying and detecting specific cancer biomarkers that may be present in the early stages of the disease would be of great benefit. Biomarker detection offers reduced costs and increased sensitivity when compared to traditional methods of cancer screening, whilst allowing the tracking of tumour origin and cancer progression. The use in patient stratification is not only important for individual treatment decision making but also for enhancing clinical trial design (Adamaki and Zoumpourlis, 2021). However, often these biomarkers are only present in extremely low numbers and are present in complex extracellular environments, making detection much more complicated (Dong et al., 2021). Further inter-patient and intra-tumour heterogeneity pose massive challenges to developing biomarker tests looking into specific targets. Development of biomarker detection assays to enable detection of a panel of PC related biomarkers has increased their effectiveness. This highlights the necessity for development of technology that is highly adaptable to the detection of any biomarker as opposed to specific detection of a select few (Adamaki and Zoumpourlis, 2021). Currently next-generation sequencing is the only method to get a precise snapshot of the genetic landscape of a tumour but the associated costs of this, though decreasing, is still very high. With one of the key molecular phenotypes in PC being a high level of intratumor heterogeneity, standard next generation sequencing techniques may lose sensitivity in biomarker detection. Thus, recently there has been a shift towards single cell RNA sequencing in patient biomarker detection which aims to overcome this issue whilst offering other a host of other benefits over traditional bulk next generation sequencing (Feng et al., 2024).

Ideal biomarker detection platforms boast highly reproducible results with clear readouts allowing for great ease of use, whilst providing high specificity (correctly identifying negative samples) and high sensitivity (correctly identifying positive samples) at a low cost. Current available methods of biomarker detection hold some of these qualities, yet none provide all, with each having their own key advantages and drawbacks [Figure 1.11] (Kohaar et al., 2019).

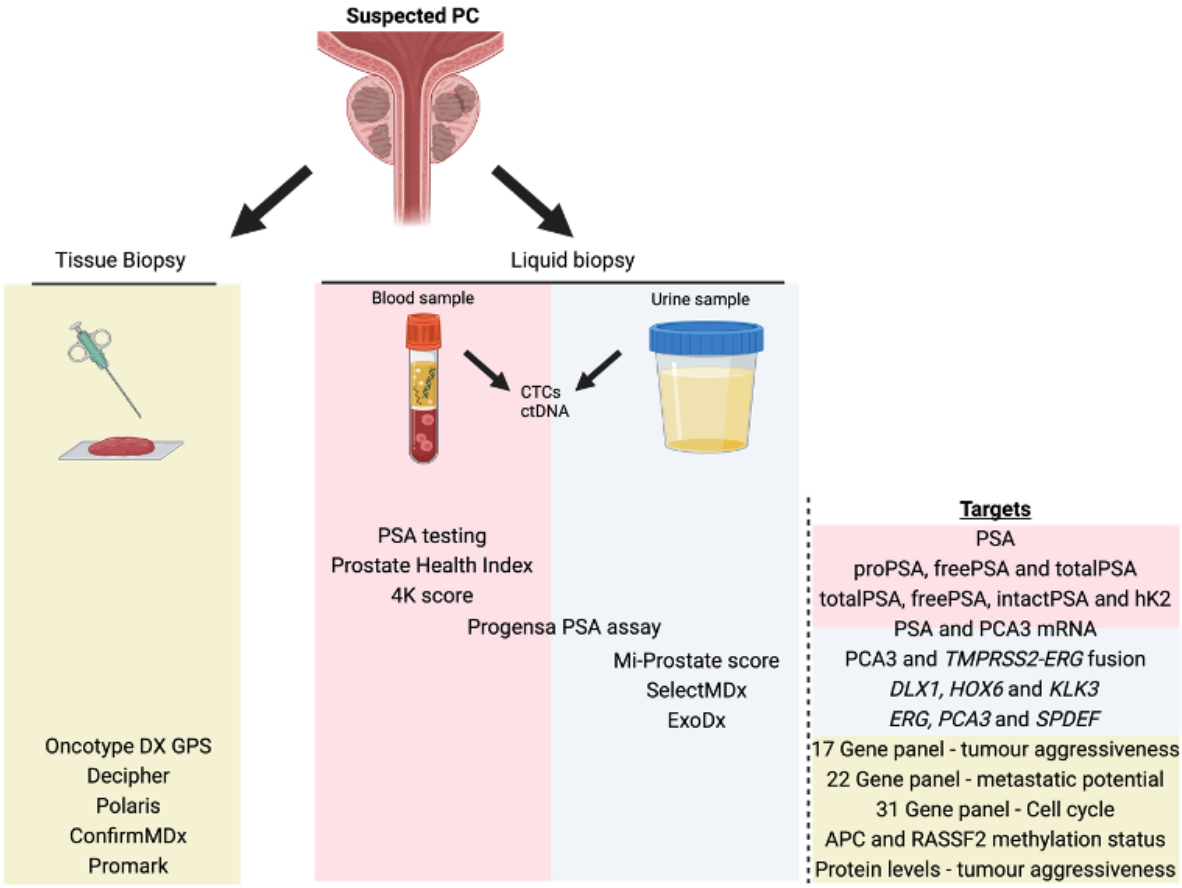


Figure 1.11 – Schematic highlighting the range of clinical biomarker tests available for both tissue and liquid PC biopsies with their associated target biomarkers. *Liquid biopsies in prostate cancer consist of either blood or urine samples and are generally preferred due to their non-invasive nature. Blood based biomarker tests are highlighted in red and mainly target circulating PSA levels. Urine samples, highlighted in blue, generally look for increased PC associated gene expression levels from circulating tumour cells and DNA. Tissue biopsy biomarker tests are highlighted in green and tend to cover larger gene panels compared to liquid biopsy tests.*

1.10.2 PSA testing and blood-based biomarker detection

PSA is encoded by the *KLK3* gene and sees increased expression in PC compared to normal tissues, making it an ideal potential biomarker. Use of PSA testing as a PC biomarker was approved by the FDA in 1986 for monitoring of patients with existing PC. Later in 1994, PSA testing in combination with digital rectal examination was approved for diagnostic screening of PC in clinic (Catalona et al., 1991). Following approval, PSA testing has seen extensive use and had a massive positive impact on PC mortality and has been iterated and improved since (Etzioni et al., 2008). Typically, a set PSA threshold is used, however other methods such as rising PSA over a set period of time, termed PSA velocity, can provide much greater diagnostic accuracy. However, despite its great positive impact, PSA has been found to be a very imprecise biomarker, with increased PSA levels not always being related to the presence of cancer as it can spike in response to non-cancer related triggers. (Thompson et al., 2004). From met-

analysis of PSA diagnostic accuracy in the detection of prostate in patients presenting with suspected symptoms performed by Merriel et al., 2022, PSA testing reports a high sensitivity of 0.93 for PSA however the selectivity was reported to be much lower at 0.20, indicating that PSA testing lacks the ability for PC detection in symptomatic patients(Merriel et al., 2022). The knock-on effect of this is a large false positive rate and increased number of prostate biopsies which result in downstream patient complications such as urinary incontinence and erectile dysfunction (Loeb et al., 2013).

Due to the inaccuracies with PSA testing alone, PSA levels have been incorporated into multi-target tests focusing on other prognostic biomarkers to enhance specific and selectivity for PC detection. The Prostate Health Index (PHI) gained FDA approval in 2012 and combines proPSA, unbound or freePSA and totalPSA, the ratios of which direct diagnostic PC detection. This was seen to significantly increase specificity of PC detection in men with increased general PSA levels through standard PSA testing (Catalona et al., 2011, Wang et al., 2014). This has been taken further in the form of 4K score testing, which looks into levels of 4 kallikrein proteins. 4K scores look into expression levels of totalPSA, freePSA, intactPSA and hK2 (a prostate-specific protein target) to provide further increased accuracy in PC detection than PHI or PSA testing alone. 4K scores not only detect presence or absence of PC but can also predict the PC risk alongside the aggressiveness of the disease. Although 4K scores have yet to gain FDA approval, it has the potential to reduce the risk of unnecessary biopsies seen by current PSA-based biomarker detection platforms (Vickers et al., 2011, Carlsson et al., 2013, Voigt et al., 2014).

1.10.3 Urine-based biomarkers

Currently one of the most extensively used biomarkers form urine is prostate cancer antigen (PCA3), discovered to be elevated in transcriptome data comparing normal to cancerous prostate tissue, and later confirmed to be elevated in all PC samples (Kohaar et al., 2019). As such, it was integrated into biomarker detection assays such as the FDA approved Progensa PSA assay (Groskopf et al., 2006). The assay, used in men over 50 with negative biopsy results, compares ratios of PCA3 mRNA and PSA mRNA from urine samples, using a set cut-off value to predict presence or absence of PC. The test sees beneficial results compared to PSA testing alone, but displays variable sensitivity and specificity depending on the cut-off used. Multiple studies have been performed showing that instead of using urine mRNA levels for both PCA3

and PSA, urine mRNA levels of PCA3 and serum levels of PSA should be used instead for more accurate PC diagnosis whilst also being able to accurately predict disease aggressiveness (Kohaar et al., 2019).

As seen in the Progensa PSA assay, multiplexing biomarker detection can allow for much greater specificity and selectivity in PC detection. Due to this, multiplexed assays such as the Mi-Prostate score, SelectMDx and ExoDx Prostate assay have been developed in urinary PC biomarker detection, focusing on the detection of other commonly upregulated mRNAs in PC. The FDA approved Mi-Prostate score utilises PCA3 and *TMPRSS2-ERG* fusion levels as biomarkers for PC identification. This shows increased predictive accuracy than PSA testing alone, although PSA levels can be factored into the assay to further increase accuracy (Rice et al., 2010). The SelectMDx assay is a multiplexed RTqPCR assay which looks at expression levels of *DLX1* and *HOXC6* from urine samples from patients being considered for biopsy. *KLK3* expression is used as a reference, with the 3 gene panel approach showing improved detection of PC when compared to PSA and PCA3 testing (Van Neste et al., 2016a). Finally, the ExoDx Prostate assay is used in the detection of genetic material within exosomes in urine. Detection of *ERG* and *PCA3* levels, with *SPDEF* as a control, has been shown to perform better in PC detection than PSA testing in high grade vs benign PC (Donovan et al., 2015, McKiernan et al., 2016).

1.10.4 Tissue-based biomarker detection

Patient tissue is the most ideal source of material for biomarker detection and as a result multiple tissue-based assays have been developed. Currently available tissue-based biomarker tests focus on expression of different PC biomarkers in biopsy positive patents. These include the Oncotype DX Genomic Protein Score (GPS), Decipher and Prolaris assays looking at mRNA expression levels, the ConfirmMDx assay looking into methylation status of proteins, and Promark assay looking directly at protein levels.

Oncotype DX GPS focuses on mRNA expression of a panel of 17 genes linked with growth and survival which has been validated in studies focussing on stratifying patient tumour aggressiveness (Cullen et al., 2015, Van Den Eeden et al., 2018). The Decipher assay similarly looks at mRNA expression of a 22 gene panel developed from studies into differential expression datasets from cancer vs control samples, focusing specifically on predicting disease

aggressiveness and metastatic potential (Karnes et al., 2013, Klein et al., 2016). The Prolaris assay instead detects the expression changes in 31 cell cycle progression associated genes (HQO, 2017). The commercially available ConfirmMDx assay focuses on DNA methylation through multiplexed detection of APC and RASSF2 methylation status alongside GSTP1 in the detection of PC (Van Neste et al., 2016b). The Promark assay again aims to provide prognostic information through predicting disease aggressiveness, although focuses on protein levels directly (Shipitsin et al., 2014, Blume-Jensen et al., 2015).

While these tests show promising results, they all rely on biopsy samples being taken which have been shown to impact patient quality of life and thus are ideally avoided for use in prognostic PC detection (Loeb et al., 2013). As a result, use of liquid biopsies from blood or urine over tissue biopsies are preferred sources of patient material in biomarker testing, and new biomarker detection platforms are being developed focussing on other sources of material from these liquid biopsies that better represent a patient's tumour.

1.10.5 Circulating tumour cells, circulating tumour DNA and extracellular vesicles

Tumour biomarkers can be detected in circulating tumour cells (CTC) and circulating tumour DNA (ctDNA) within the blood, and at the protein and mRNA levels within living cells (Dong et al., 2021, Srivastava et al., 2013, Hong et al., 2016). Detection of CTCs from blood samples presents the optimum method of detecting tumour cells in a non-invasive way. Unfortunately, however, they are present at very low levels within the blood and only readily detectable as the disease progresses. Couple this with the amount of non-cancerous cells present within the blood means isolating and detecting the presence of biomarkers from CTCs is a very challenging process. A suitable alternative would be ctDNA, which is the genomic material released from dying or dead tumour cells into the blood. This DNA will carry some of the generic material present in the primary tumour cell, including mutations and genome alterations, which if detected, could allow for identification and characterisation of primary tumour cells (Dong et al., 2021). A limitation with the detection here is again the abundance of ctDNA, as it is only present at very low concentrations, with it being estimated that the amounts of ctDNA within total circulating DNA is around 1-0.01% (Dong et al., 2021). Further detection of a single mutation within the ctDNA poses an even more difficult challenge as there will often be an excess of WT DNA present (Tang et al., 2020). Alongside these, due to ctDNA

being released in response to cancer cell death, it is difficult to ascertain if circulating ctDNA has come from increased tumour burden or from increases in cell death in response to successful anti-cancer treatments.

Current options for detection of mutations within ctDNA include PCR and nucleic acid sequencing (Tang et al., 2020). These techniques hold a high error rate for mutation discrimination and have lengthy operating procedures whilst requiring expensive infrastructure to be in place. Abnormal RNA expression has been associated with many cancers and as a result offers another alternative target for early cancer biomarker detection (Srivastava et al., 2013). Much like CTCs and ctDNA, a problem with using RNA as a target for biomarker detection are its limited abundance, degradation by nucleases present in the plasma, and varying concentration (Dong et al., 2021).

Another source of potential biomarkers within liquid biopsy samples are those proteins and RNA contained within extracellular vesicles. An example of this can be seen in the promising early data that has been generated looking into utilising lipid biomarkers for the identification of poor-prognosis in mCRPC patients (Scheinberg et al., 2024). Here PCPro (a combination of multiple lipid biomarkers) levels were assessed utilising a bespoke mass spectrometry-based assay, with PCPro positive mCRPC patients displaying significantly shorter overall survival that those who were PCPro negative. These results highlight extracellular vesicles as a viable source of biomarkers however there are some inherent issues with them as a biomarker source, namely their heterogeneity due their release by various mechanisms and cell types, and the dilution of cancer-derived extracellular vesicles with non-cancer derived extracellular vesicles, ultimately decreasing the abundance of cancer associated biomarkers (Ramirez-Garrastacho et al., 2022).

In ctDNA and RNA detection, DNA nanotechnology presents an alternative form of detection that can overcome the challenges presented. DNA nanomachines offer increased sensitivity, selectivity, versatility, and biocompatibility than the current enzyme-associated methods used for biomarker detection (Dong et al., 2021). For these reasons, extensive work has been done to develop mechanisms for these DNA nanomachines to be used in the detection of cancer biomarkers. The fundamental basis for these nanomachines is the ability to change upon interaction with a specific target and release a signal. Some DNA nanomachine-based methods of detection can be enzyme associated (including the use of PCR) and offer high selectivity. However, the use of enzyme-free alternatives offers many advantages over the enzyme-

associated methods (Liu et al., 2019, Dong et al., 2021). Most of these enzyme-free methods centre around toehold mediated strand displacement (TMSD) which results in decreased background signals, increased catalytic efficiency and much more stable reaction systems.

1.11 DNA Nanotechnology

One of the first presented uses of DNA nanotechnology was by Seeman *et al*., which created programmable DNA-DNA interactions that were based on fundamental Watson-Crick base pairing (Srinivas et al., 2017, Liu et al., 2019). This was first confirmed in a study by Kallenback *et al*., which was based on utilising 4 arm junctions, resembling Holliday junctions observed in DNA replication, and has since been expanded to create complex 3D structures and complex folded DNA origami to demonstrate the high level of programmability of hybrid fabricated DNA nanostructures able to house proteins, nanoparticles and carbon nanotubes (Kallenbach et al 1983)(Chao et al 2016)(Shen et al 2021). This was dubbed 'Structural DNA Nanotechnology' and offers nanoscale controllability and high predictability of nanomachines designed in this way. Ultimately these DNA nanomachines could be used to design molecular reaction pathways and chemical reaction networks that would allow detection of specific reaction pathways (Srinivas et al., 2017). Later work in this field moved towards 'Dynamic DNA nanotechnology' which involves the dynamic movement of these nanostructures, stimulated by nucleic acid transitions between different nanostructures (Liu et al., 2019). In this process, overhangs present on DNA molecules are able to act as triggers that initiate branch migration reactions, facilitating the displacement of DNA strands. As a result, DNA nanotechnology has the potential for highly selective and specific detection of target DNA oligonucleotides, that allows signal amplification upon analyte detection, allowing much lower limits of detection than current biomarker detection platforms (Shen et al 2021). Thus, applying DNA nanotechnology systems in the detection of PC associated biomarkers holds the potential to out-perform the current repertoire of biomarker detection platforms.

DNA holds many attractive qualities that suit it for use as building blocks for molecular nanostructures capable of biomarker detection. Highly predictable secondary structure, high biocompatibility, easy programmability, and self-assembly properties are some of the features that allow the design of functional nanostructures that can be easily engineered for use in a wide range of applications. Furthermore, DNA can be combined with nanostructures to broaden its possible applications. Examples of such applications can be divided into: (i) organic nanomachines, which uses only bioavailable DNA material in creating DNA hydrogels or DNA signal cascades; and (ii) inorganic nanomachines which involves conjugation of DNA nanoparticles in combination with gold or magnetic nanomaterials (Chen et al 2018). What is shared by all these nanostructures is their exploitation of DNAs programmable nature and

their critical use of overhang displacement in a reaction called toehold mediated strand displacement.

1.11.1 Toehold Mediated Strand Displacement

One of the main methods DNA nanomachines use that allows for detection of targets is TMSD and are used ubiquitously across many dynamic DNA nanotechnology applications (Dong et al., 2021, Liu et al., 2019, Tang et al., 2020). TMSD was first introduced by Yurke *et al*., and involves a duplex structure comprised of two ssDNA oligonucleotides with free energy powering the reaction [Figure 1.12]. (Liu et al., 2019, Yurke et al., 2000). One strand, the 'complement' is complementary to the target sequence and contains the toehold, a sticky end which facilitates target binding. The other strand, the 'protector', resembles the target sequence but is shorter than the complement which makes up the toehold (Tang et al., 2020). The TMSD reaction progresses through a branch migration process. As the target binds the toehold and displaces the 'protector' in a process that resembles a random walk. However, as more base pairs are formed, the displacement reaction becomes more energetically favourable, ultimately resulting in the displacement of the 'protector' strand from the original duplex and a new stable duplex being formed between the target and the 'complement' strand that is unable to dissociate (Tang et al., 2020).

When there are short duplex sequences being used, the kinetics of the reaction depend on both the length and sequence of the toehold (Liu et al., 2019, Tang et al., 2020). Using the TMSD mechanism, discrimination between wild type sequences and mutant sequences is possible by including the point mutation in the 'complement' strand shortly after the toehold, which will have an impact on reaction kinetics (Tang et al., 2020). This allows for the construction of assays that can use this change in kinetics to indicate whether or not samples contain a point mutation of interest. TMSD has seen many applications in DNA nanotechnology and has allowed the design of many different reactions such as Remote Toehold reactions and Toehold exchange (Tang et al., 2020). First introduced by Zhang *et al.*, toehold exchange works through a similar method of action as TMSD (Zhang et al., 2007). The difference between the two lies in the reverse toehold present in the duplex substrate. The reverse toehold allows the decoupling of the kinetics and thermodynamics of the reaction, allowing the reaction to progress regardless of the thermodynamics. As a result, this is one of the most powerful and

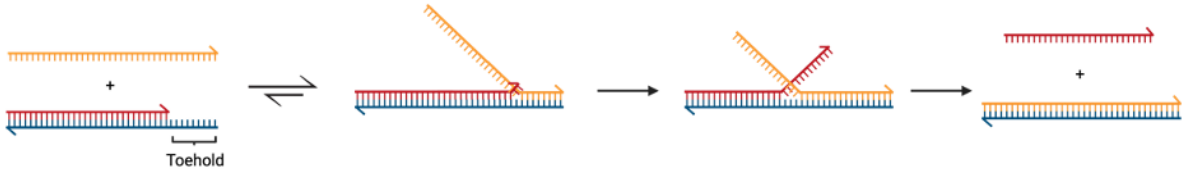


Figure 1.12 – Schematic representation of TMSD reactions. *Probe consisting of a blocker (red) and complementary strand to the target (blue) with resultant overhang or 'toehold' due to shorter blocker strand. Upon addition of target strand (yellow) stepwise displacement of the blocker strand will take place, resulting in a fully complementary duplex formed.*

widely used mechanisms when designing probes and assays, however the final product is in equilibrium with the preceding intermediate formed, meaning there may potentially be signal fluctuations (Tang et al., 2020).

Chapter 2 – Aims

With PC now being one of the most common cancers across the world, and the second most common form of cancer-related death in the UK, there is massive scope for increasing our capability of better treating the disease. The key point of mortality in the disease process is progression from early-stage PC to late-stage advanced PC, where an over 50% drop in 5-year survival is observed. PC progression also comes with fewer therapeutic options, with available treatments ultimately leading to progression to the more aggressive androgen-independent form of the disease, termed CRPC.

One of the mechanisms at play in PC development and progression is chromosomal rearrangements, potentially due to dysfunctional DNA repair mechanisms. Mutations within the *SPOP* gene, which is the most commonly occurring mutation in early PC, results in a functionally inactive protein leading to compromised HR and promotion of error-prone NHEJ-mediated DNA repair. This reliance of NHEJ in *SPOP* mutant patients contributes to an increase in chromosomal rearrangements and, ultimately, tumour progression. Critically, the dependency of *SPOP* mutant cells on the NHEJ pathway may offer a vulnerability that can be exploited therapeutically. As such, targeting DNA-PKcs, a key protein in NHEJ, has the capacity to render cells unable to efficiently repair DNA damage, presenting a potential synthetic lethality in these cells in response to exogenous DNA damage.

Mutations within *SPOP* are found in ~11% of PC patients, presenting *SPOP* mutations as a biomarker for susceptibility for DDR targeting therapeutic agents. Therefore, methods of characterising and categorising tumour subtype are of importance to enable more effective and patient personalised treatment. However, current methods of biomarker detection in PC are limited by their specificity, sensitivity, and are inflexible in target selection. DNA nanotechnology-based probes, developed by Nanovery, offer a more streamlined assay alternative that is amenable to detection of any target RNA/DNA, with the potential for much greater specificity and sensitivity over other commercially available assays. However, there still remains considerable validation and optimisation of these DNA nanotechnology-based probes before this technology can be applied in a clinical setting.

This project looks to further validate SPOP as a new PC biomarker of synthetic lethality with inhibitors of the NHEJ pathway. Alongside this, work with Nanovery looks to develop the future

of PC biomarker detection using DNA nanotechnology-based probes targeting not only *SPOP* mutations but a range of other clinically relevant PC biomarkers. This work will be carried out through two main objectives:

i. Investigate the potential synthetic-lethality in response to DNA damage in cells defective for SPOP activity upon DNA-PKcs inhibition and develop new models to assist in further exploring this relationship.

ii. Further develop Nanovery DNA-based nanotechnology probes to detect clinically-relevant PC associated biomarkers, find new applications of the technology and compare against current commercially available methods of biomarker detection.

Chapter 3 – Materials and Methods

3.1 Mammalian cell culture

3.1.1 Cell Culture

CWR22Rv1 (ATCC® CRL-2505™), PC3 (ATCC® CRL-1435™), VCaP (CRL-2876™), LNCaP (CRL-1740™) and HEK293FT (ATCC® CRL-3216™) cell lines were purchased from ATCC (Virginia, US). HEK293FT, PC3, CWR22Rv1 and the derivative CWR22Rv1-AR-EK cell line were cultured in RPMI-1640 Media (R5886, Sigma Aldrich) (Table 3.1). The VCaP cell line was cultured in Dulbecco's Modified Medium (DMEM) (D5030, Sigma Aldrich). These media were supplemented with 10% (v/v) foetal calf serum (FCS)(Sigma), 2 nM L-glutamate (Sigma) and 1% penicillin/streptomycin to make up what will hereby be referred to as full media. All cell lines were cultured at 37°C in 5% CO_2 in a humidified incubator (MCO-20AIC, Sanyo).

Table 3.1 - *Cell lines used in mammalian cell culture and their origin.*

Cell line	FL-AR/AR-V expression in full media	Origin
CWR22Rv1	+++/++	Xenograft tumour from primary tumour.
CWR22Rv1-AR-EK	-/++	CWR22Rv1 CRISPR-derivative (Kounatidou et al., 2019).
PC3	-/-	Bone metastasis, neuroendocrine lineage.
VCaP	+++/+	Xenograft tumour from metastasis.
LNCaP	+++/-	Lymph node metastasis.
HEK293T	-/-	Embryonic kidney.

3.1.2 Cell sub-culturing

BioMat class II microbiological safety cabinets were used to carry out all cell culture. After reaching approximately 70% confluency, cells were passaged by removing media from the culture flask, washed twice with warmed phosphate buffered saline (PBS) (Invitrogen) then incubated with 1 x Trypsin-EDTA (Merck) to detach cells. Following this, trypsin was neutralized using full media and cells were transferred to a sterile universal tube and centrifuged at 400 x

g for 5 minutes after which the supernatant was removed, and cell pellets were resuspended and seeded into new flasks. Mycoplasma tests were carried out in-house every 2 months.

3.1.3 Cell storage

Cells were centrifuged at 400 x g for 5 minutes, the supernatant removed and resuspended in freezing media (RPMI media supplemented with 10% DMSO (Sigma) and 10% FCS). 1 mL of this cell suspension was aliquoted into cryovials (Thermofisher Scientific) and subsequently frozen at -80 ^{O}C. For long-term storage, once frozen as above at -80 ^{O}C, cells were transferred to liquid nitrogen. Cells were thawed by centrifuging at 400 x g for 5 minutes followed by removal of the freezing media supernatant, then seeded into a culture flask and passaged twice before experimental use.

3.1.4 Compounds

Compounds commonly used in mammalian cell culture were used as described in Table 3.2.

Table 3.2 – *Compounds used in mammalian cell culture with stock and final treatment concentrations.*

Compound	Origin and use	Supplier	Stock concentration	Treatment concentration	Storage
Doxycycline	Synthetic broad-spectrum antibiotic derived from oxytetracycline	Merck	1 mg/mL in DMSO	1 µg/ mL	-20 ^{o}C
Puromycin dihydrochloride	Antibiotic protein synthesis inhibitor from *Streptomyces alboniger*	Sigma	1 mg/mL in DMSO	1 µg / mL	-20 ^{o}C

3.2 Plasmid DNA, siRNA and sgRNA transfections

3.2.1 siRNA and sgRNA Transfection

siRNA was transfected into cells using Lipofectamine® RNAiMAX Transfection Reagent (Invitrogen) at a ratio of 1:3 siRNA (μg) : RNAiMAX (μl). This mixture was suspended in the necessary basal media (non-supplemented media) and left to incubate for 30 minutes before being administered dropwise to pre-seeded cells (forward transfection) to a final concentration of 25 nM in cells. Cells were then incubated for 48 hours before further application. siRNA or sgRNA used are described in Table 3.3 and 3.4 respectively.

Table 3.3 - *siRNA targets and sequences*

Gene target	siRNA sequence/target
Scr (non-specific)	UUCUCCGAACGUGUCACGU[dT][dT]
SPOP (ON-TARGETplus siRNA SMARTpool)	GAGAGUCAACGGGCAUAUA
	GGUAAAGGUUCCUGAGUGC
	CAACUAUCAUGCUUCGGAU
	AAAUGGUGUUUGCGAGUAA

Table 3.4 - *sgRNA targets and sequences*

Gene target	sgRNA sequence
Scr (non-specific)	*AACCCCTGATTGTATCCGCA*
SPOP-KO-1 (Exon 3)	*TAGTGAAATTCTCCTACATG*
SPOP-KO-2 (Exon 3)	*GTTATTGATGGTCCACATGT*
SPOP-KO-3 (Exon 3)	*CAATAACTTTAGCTTTTGCC*

3.2.2 Plasmid DNA transfection

Plasmid DNA was transfected into cell lines using the LT-1 transfection reagent (Mirus Bio). A 1:3 plasmid (μg) : LT-1 (μl) ratio was made up in RPMI-1640 media (Thermo Scientific) in an Eppendorf tube and incubated for 30 minutes at room temperature to equilibrate. Following incubation, the transfection mixture was added dropwise to pre-seeded cells.

3.3 SDS-poly acrylamide gel electrophoresis (SDS-PAGE) and western blotting

3.3.1 SDS-PAGE

Reagents used for cell lysate generation and western blotting can be found in Table 3.5. Cell lysates were harvested by direct addition of 120 µL SDS-sample buffer to cells. Lysates were then boiled at 100^0C for approximately 10 minutes before being loaded onto gels for SDS-PAGE.

10% PAGE gels (see Table 3.6 for ingredients) were made using the mini-PROTEAN® tetra casting system (Bio-Rad). The pre-boiled protein samples were loaded onto gels alongside a Spectra Multicolor Broad Range Protein Ladder (Thermo Science) and separated by electrophoresis using the mini-PROTEAN® Tetra Vertical Electrophoresis Cell (Bio-Rad) at a constant voltage of 130 V in running buffer (Table 3.5).

3.3.2 Western Blotting

After separation by gel electrophoresis, proteins were transferred to a Hybond ECL Nitrocellulose membrane (GE Healthcare) using the Mini Trans-Blot Electrophoretic Transfer Cell system (Bio-Rad) at a constant voltage of 30V overnight in transfer buffer (Table 3.5). Following transfer, membranes were blocked using 5% (w/v) milk powder (Marvel) dissolved in TBS (Table 3.5) for around 1 hour at room temperature. Membranes were incubated with the specific primary antibody (Table 3.7) diluted in 4 mL of 1% (w/v) milk/TBST overnight at 4 ^{0}C. Membranes were then washed twice with TBST for around 5 minutes per wash with gentle rocking at room temperature before being added to the appropriate secondary antibody for 1 hour at room temperature. Following this, membranes were washed again twice in TBST before being developed using Clarity™ ECL Western substrate (Bio-Rad) and imaged using ChemiDoc™ Imaging systems (Bio-Rad) to detect signal.

Table 3.5 – *SDS-PAGE and Western Blotting Reagents*

Solution	Composition

SDS sample buffer	*125 mM Tris-HCl, pH6.8* *5% SDS* *10% glycerol* *10% β-mercaptoethanol* *0.01% bromophenol blue*
Buffer 'A' (2X)	*750 mM Tris-HCl* *0.2% SDS* *pH 8.8*
Buffer 'B' (2X)	*250 mM Tris-HCl* *0.2% SDS* *pH 6.8*
Running buffer	*25 mM Tris* *190 mM glycine* *0.1% SDS*
Transfer buffer	*25 mM Tris-HCl* *150 mM glycine* *10% methanol* *pH 8.3*
TBS	*500 mM NaCl* *200 mM Tris-HCl* *pH 7.5*
TBST	*500mM NaCl* *200 mM Tris-HCl* *0.001% Tween-20* *pH 7.5*

Table 3.6 – *Reagents required for x2 10% polyacrylamide gels*

Reagent	**10% Gel**	
	Running gel	**Stacking gel**
2X Buffer A	*5 mL*	-
2X Buffer B	-	*5 mL*
Acrylamide (30%)	*3.33 mL*	*1.67 mL*
diH_2O	*1.67 mL*	*3.33 mL*
Ammonium persulphate	*100 µL*	*100 µL*

N, N, N', N'-Tetramethylethan-1,2-diamine (TEMED)	*20 µL*	*20 µL*

Table 3.7 – *Antibodies used for Western Blot and immunofluorescence.*

Antibody	**Species**	**Supplier**	**Applications**	**Catalogue number**	**Dilution in 1% milk/TBST (WB) BSA/PBS (IF)**
SPOP	*Rabbit*	*Abcam*	*WB*	*ab137537*	*1:500-4000*
SPOP	*Rabbit*	*Santa-Cruz*	*WB*	*sc-377206*	*1:500-4000*
SPOP	*Rabbit*	*Proteintech*	*WB*	*16750-1-AP*	*1:1000*
FLAG M2	*Mouse*	*Sigma-Aldrich*	*WB*	*F3165.2MG*	*1:1000*
α-tubulin	*Mouse*	*Sigma-Aldrich*	*WB*	*T9026*	*1:1000*
GAPDH	*Mouse*	*Sigma-Aldrich*	*WB*	*G8795*	*1:1000*
CDC20	*Rabbit*	*Bethyl*	*WB*	*IHC-00232*	*1:1000*
Phospho-Ser139 γH2AX	*Mouse*	*Novus*	*IF*	*NB100-74435*	*1:500*
cGAS	*Rabbit*	*Cell signalling*	*IF*	*DID3G*	*1:400*
Donkey anti-mouse Alexafluor488	*Donkey*	*Invitrogen*	*IF*	*A-21202*	*1:500*
Swine anti-rabbit HRP	*Pig*	*Dako*	*WB*	*P0217*	*1:1000*
Rabbit anti-mouse HRP	*Rabbit*	*Dako*	*WB*	*P0260*	*1:1000*

3.4 RNA extraction

3.4.1 TRIzol® based total RNA extraction

Total RNA for use in miRNA detection was extracted from cells using TRIzol® Reagent (Life Sciences, Invitrogen) according to the manufacturer's protocol. Media was removed from cells and cells were washed with PBS before 1 mL of TRIzol ® Reagent was added directly to cells and homogenized. Samples were then transferred to a sterile Eppendorf tube and incubated at room temperature for 5 minutes. After this time, 200 µL of chloroform (Sigma Aldrich) was added to each sample followed by vigorous shaking, a 3-minute incubation at room temperature then centrifugation at 12,000 x g for 15 minutes at 4^{o}C. The upper aqueous phase was transferred to a fresh Eppendorf tube, 500 µL of isopropanol (Fisher Chemicals) was added, a further 10 minutes of incubation at room temperature followed by centrifugation at 12,000 x g for 10 minutes at 4^{o}C. The RNA pellet formed was washed in 75% ethanol and centrifuged at 7,500 x g for 5 minutes at 4^{o}C before the ethanol was removed and the pellet allowed to air dry. Once dry, the pellet was resuspended in 30-50 µL nuclease-free water (Invitrogen) and solubilized at 60^{o}C before RNA concentration was quantified using a Nanodrop spectrophotometer (Thermo Scientific). Once extracted and quantified RNA was stored at -80^{o}C.

3.4.2 Thermofisher GenElute RNA miniprep kit based extraction

For RNA used in RT-qPCR and RNAseq analysis, extractions were performed using the Thermofisher GenElute RNA miniprep kit (Thermo Scientific). Extractions were performed as detailed in the manufacturers protocol. In brief, cells were lysed in 500 µL lysis solution before addition of 500 µL 70% ethanol and vortexing. Samples were then run through the RNA binding columns, washed twice before being eluted in the provided elution solution before RNA concentration was quantified using a Nanodrop spectrophotometer (Thermo Scientific). Once extracted and quantified RNA was stored at -80^{o}C.

3.5 Reverse Transcription and real-time qPCR analysis

3.5.1 Reverse transcription

cDNA for use in qPCR reaction was generated using reagents listed in Table 3.8 alongside 1 µg of RNA incubated at 37°C for 1 hour. Following this, samples were incubated at 100 °C for 10 minutes to denature enzymes before diluting further with 130 µL RNase/DNase-free H_2O. The final diluted cDNA samples were stored at -20 °C until use in downstream real-time qPCR (RT-qPCR) analysis.

Table 3.8 – *Reagents used per Reverse transcriptase reaction for cDNA generation*

Reagent	**Volume/reaction (µL)**
M-MLV Reverse Transcription Buffer	*4*
4 nM dNTPs	*2*
Oligo (DT)18 primers (100 µg /ml)	*1*
M-MLV Reverse Transcriptase	*0.3*
1 µg / µL RNase/DNase-free H_2O	*12.7*

3.5.2 RT-qPCR

Gene expression from cDNA samples was quantified using a QuantStudio 7 Flex Real-Time PCR system (Thermo Scientific), using custom primers purchased from IDT/Sigma (Table 3.9) alongside PowerTrack™ SYBR™ Green Master Mix (Applied biosystems). 384 well plates (Applied biosystems) were used, loaded with 10 µL total reaction components as described in Table 3.10. Alongside experimental samples, no-template controls were set up using RNase/DNase-free H_2O in place of the cDNA sample to check for any potential primer contamination. 384 well plates were subsequently sealed with MicroAmp optical adhesive films, centrifuged at 1000 x g for 20 seconds then ran using qPCR parameters as described in Table 3.11. The relative gene expression was then calculated, normalised to the reference RPL13A using the equation:

$$[\Delta][\Delta]Ct = [\Delta]Ct^{Samples} - [\Delta]Ct^{Reference}$$

Table 3.9 - *Custom primers used in RT-qPCR analysis*

Gene target	**Forward primer (5'-3')**	**Reverse primer (5'-3')**
RPL13A	CCTGGAGGAGAAGAGGAAAGAGA	TTGAGGACCTCTGTGTATTTGTCAA
SPOP	CTGCTATCTTCCTATCAGGATC	CAAAAGTCCTGGCCTTCATGG
BRCA2	GAATGGCAGACTGACAGTTG	GAGCAGGCCGAGTACTGTTA
ATR	ACCTCAGCAGTAATAGTGA	GGCCACTGTATTCAAGGGA
RAD51	TGGAAGAAGAAAGCTTTGGC	TCTTCACATCGTTGGCATTT
CHK1	ATATGAAGCGTGCCGTAGA	TGCCTATGTCTGGCTCTATT

Table 3.10 - *Total well composition for RT-qPCR reactions*

Reagent	**Volume/well (µL)**
Forward primer (25 ng/µl)	*0.4*
Reverse primer (25 ng/µl)	*0.4*
RNase/DNase-free H_2O	*2.2*
PowerTrack™ SYBR™ Green Master Mix	*5*
cDNA	*2*

Table 3.11 - *RT-qPCR parameters*

Step	Cycles	Time	Temperature	Data collection
Enzyme activation	*1*	*10 minutes*	*95 °C*	-
Denature	*40*	*5 seconds*	*95 °C*	-
Anneal/extend		*30 seconds*	*60 °C*	*
Melt-curve analysis	-	-	*60 – 95 °C*	*

3.6 Statistical Analysis

Unless stated otherwise, all graphical data presented is representative of the mean of three individual biological replicates and error bars represent +/- standard error means (SEM). Statistical analysis was performed using GraphPad Prism (Ver. 10.0.2) software, with the type of statistical test indicated in figure legends, although typically are either unpaired t-tests, one-way ANOVA or two-way ANOVA. *,**,***,**** represent p-values of <0.0332, <0.0021, <0.0002 and <0.0001 respectively with non-significant results represented as NS.

Chapter 4 – Validation of SPOP targeting reagents and development of model systems to further interrogate effects of SPOP mutations within PC cell lines.

4.1 Introduction

Due to none of the commonly used PC cell lines harbouring *SPOP* mutations, previous studies have had to rely on alternate methods of SPOP manipulation to mimic the loss of function effects observed in patients. To better understand the current methods of SPOP manipulation being applied to PC cell lines, a review of the current literature was performed. Using a simple PubMed search for ((SPOP) AND (Prostate Cancer)) NOT (Review), 206 papers were returned, of which abstracts were reviewed leaving 45 relevant articles. The methods section and presented data from these articles was assessed and the methods used for SPOP manipulation were compiled and summarised in Figure 4.1. From this, the predominant methods used were (i) ectopic overexpression of either SPOP-WT or SPOP mutant protein; and (ii) depletion of SPOP protein through siRNA or shRNA gene silencing. A number of studies utilised CRISPR/Cas9-mediated knock-out or knock-in, but these were very much in the minority. Many of the papers reviewed utilised mouse models with altered SPOP levels although these were outside the scope of the review conducted and thus not investigated further.

Depletion of SPOP using either SPOP-targeting siRNA transfection or stable expression of SPOP-targeting shRNA containing plasmid vectors has been reported to phenocopy an SPOP loss of function. However, rarely do knockdown studies result in full depletion of SPOP protein and hence activity. Due to the previously mentioned dominant negative effects of mutant SPOP protein, ectopic, plasmid-driven overexpression of SPOP mutants can represent the loss of function effects observed by *SPOP* mutations in patient samples (Cuneo et al., 2023).

Of the 6 papers utilising CRISPR/Cas9-mediated *SPOP* knock-out, 2 used transient *SPOP* knock-out in short term experiments resulting in varied levels of SPOP protein depletion. The other 4 generated stable *SPOP* knock-outs in PC3, DU-145 and C4-2 PC cell lines before utilising these stable lines for downstream experimentation. Only one paper was able to facilitate an *SPOP* knock-in where the F102C mutations was introduced into *SPOP* wild-type LNCaP cells, which were subsequently sent for RNA-sequencing (RNAseq) and grafted into mouse models (Burleson et al., 2022). Ultimately this brief review demonstrated that a range of techniques have been used to manipulate SPOP within PC cell lines of which all attempt to model *SPOP* mutant patient samples to different degrees. Importantly, of the CRISPR/Cas9 manipulated

knockout models described in the literature, none represent an AR positive castrate resistant phenotype expressing AR-Vs observed in late stage PC, thus it may be of interest to generate knock-out/in models within a PC cell line that represents this stage of disease.

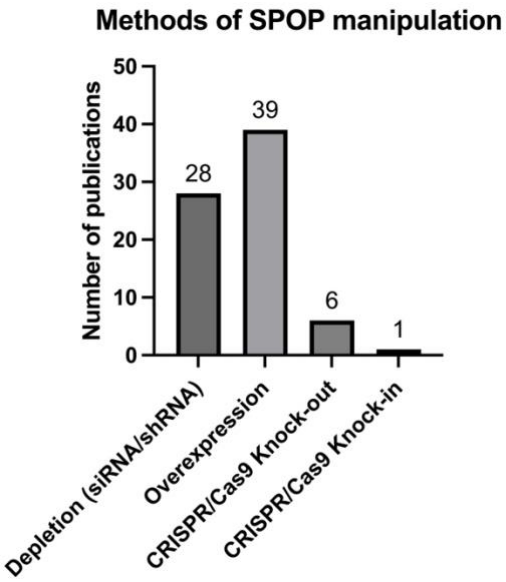


Figure 4.1 - Techniques used for manipulating SPOP levels from the literature. *Graph outlining methods of SPOP manipulation used across 45 total publications to alter levels of SPOP within cell lines.*

4.2 Aims

This project was the first in the host laboratory to investigate SPOP biology. It was therefore critical that comprehensive optimisation of several reagents, including SPOP targeting siRNAs, anti-SPOP antibodies, SPOP targeting sgRNAs and RT-qPCR primers was conducted as they would be extensively used throughout the study. As well as this, generation of new plasmid constructs to inducibly over-express wild-type or mutant SPOP, as well as new CRISPR/Cas9 manipulated models in cell lines not currently described in the literature, was necessary to allow for a full range of experimental studies of SPOP to be conducted in novel PC disease models.

This will be done by;

1. Validating SPOP-targeting reagents to allow transient and permanent depletion of SPOP protein expression within PC cell lines using siRNA- and CRISPR-based approaches, respectively.
2. Utilising these reagents to develop new PC cell line models which better recapitulate the effects caused by *SPOP* mutations within late stage PC patients.

Ultimately, completion of these aims will provide greater confidence in the reagents applied throughout the study and provide new models of PC which will potentially highlight novel SPOP biology with translational impact.

4.3 Materials and Methods

4.3.1 Bacterial transformation, plasmid DNA isolation and transfection

Bacterial transformation of plasmid DNA

Stbl3 *E.Coli* competent cells (ThermoFisher) were used for transformation and propagation of plasmid DNA. These cells were stored long-term at -80 ^{0}C. When required, cells were thawed on ice for 10 minutes and 1 μg of plasmid DNA was added to cells, gently mixed then incubated on ice for 30 minutes. Following this incubation, cells were heat shocked at 42^{o}C for 45 seconds before being placed on ice for another 2 minute incubation. 350 μL of LB culture media (10 g NaCl, 5 g Yeast Extract, 10 g Tryptone in 1 L diH_2O)) was added to the cells and transferred to a Thermatron incubator shaker (Infors HT) for 4-hour incubation at 37^{o}C with 220 rpm agitation. After this time, 50-100 μL of the cell suspension was spread onto pre-warmed LB agar plates (1% (w/v) NaCl, 1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1.5% (w/v) agar) containing 50 μg/mL appropriate antibiotic. Plates were incubated at 37^{o}C overnight.

Culture of transformed bacteria

For maxi-prep cultures, a single bacterial colony was incubated in 5 mL of LB medium which contained 50 μg/ml of appropriate antibiotic. The starter culture was incubated for 8 hours at 37°C with rotation at 200 rpm. The starter culture was transferred to a sterile conical flask containing 200 mL of pre-warmed antibiotic-containing LB medium and incubated at 37°C overnight with rotation.

Plasmid DNA extraction

PureLink™ HiPure Plasmid Maxiprep Kit kits (Life Sciences) were used for plasmid DNA extraction from cultured bacteria. In brief, these kits involved resuspending the bacterial cell pellets post-centrifugation, lysing cells, filtering of the cellular debris, precipitating the plasmid DNA with isopropanol and resuspending the resultant plasmid DNA in molecular grade water. The resultant DNA was resuspended in a suitable volume of molecular grade water. The DNA concentration was measured using a Nanodrop spectrophotometer (Thermo Scientific). DNA was stored at -20 ^{0}C for future use.

4.3.2 Genomic DNA purification

Genomic DNA purification was performed using the GenElute™ Mammalian Genomic DNA Miniprep Kit (Sigma Aldrich). Following cell detachment and cell suspension neutralisation as described in Section 3.1.2, cell suspensions were transferred to Eppendorf tubes and centrifuged at 400 x g for 5 minutes and supernatants removed. The resultant cell pellet was then lysed and passed through the DNA binding column as per manufacturer's instructions. The final elution step was performed using 200 µL RNase/DNase-free distilled water (Invitrogen). Final DNA concentration was measured using a Nanodrop spectrophotometer (Thermo Scientific).

4.3.3 Polymerase Chain Reaction

250 ng of the genomic loci of interest to be sent for sequencing or cDNA used in cloning were amplified by polymerase chain reaction (PCR). Primers at 0.3 µM (Table 4.1) were used alongside Terra™ PCR Direct Polymerase Mix (Takara) as per manufacturers instruction for 50 uL reactions. Reaction mixes were mixed and briefly centrifuged then transferred to a thermocycler (Applied Biosystems). PCR parameters for each run were set up as per described in Table 4.2, with products stored at -20 ^{O}C.

Table 4.1 – *PCR primers used for cloning and indel assesment. Red – AgeI restriction site. Green – SPOP coding sequence. Blue – BamHI restriction site.*

Primer	**Sequence (5'-3')**
TLCV2-SPOP-F	CCGGACCGGTTCTAGAGCGCTGCCACCATGTCAAGGGTTCCAAGTCCT
TLCV2-SPOP-R	GATCGGATCCGGATTGCTTCAGGCGTTTGCG
SPOP-KO-Amplicon-F	AACATAAGGTGAGAAAAACCTGAAAA
SPOP-KO-Amplicon-R	CAACAAACATTTAGTGAGTACCTATTAT
SPOP-KO-Amplicon-Seq	AAATGATAATCTGCCCTTCATCCTGTTT

Table 4.2 – *Reaction parameters used for PCR*

Step	**Temperature (°C)**	**Time**
Initial denaturation	98	30 seconds

30 Cycles	Denaturation	98	10 seconds
	Annealing	60	15 seconds
	Extension	68	1 minute
Final Extension		68	3 minutes
Hold		4	-

4.3.4 DNA gel electrophoresis

To assure the correct size of PCR amplicons, products were ran on 1% Agarose/TAE buffer (40 M Tris, 20 mM acetic acid, 1 mM EDTA) gels containing 1X GelRed (Biotium). Alongside PCR products, a Quick-Load™ 1 kb DNA ladder (New England Biolabs) were separated across the gel using a Mini-Sub Cell GT (Bio-Rad) system at 90V. Gel images were taken using the ChemiDoc™ Imaging system (Bio-Rad). If bands were required for downstream applications, such as Sanger sequencing or cloning, bands at the correct size were excised from the gel and transferred to Eppendorf tubes.

4.3.5 DNA gel extraction and Sanger sequencing

DNA was extracted from agarose gel slices using the Geneclean® Kit (MP Biomedicals) as per manufacturer's instruction. DNA was resuspended in RNase/DNase-free distilled water (Invitrogen) and total amount quantified using a Nanodrop spectrophotometer (Thermo Scientific). The product was diluted to 50 ng/uL and 20 uL was sent for Sanger sequencing (Genewiz/Eurofins). As per instructions provided, a separate Eppendorf containing 5 μL of 5 μM forward sequencing primer complementary to the region of interest was also sent.

4.3.6 Virus production and transduction

A 2nd generation system was used for lentiviral generation, involving two vectors encoding the packaging and envelope proteins along with a third vector, containing the desired cDNA to be expressed, known as the shuttling vector.

Viral Generation

HEK293T cells were seeded down in a 150 mm culture dish in low-antibiotic growth media (RPMI-1640 supplemented with 10% (v/v) FCS (Gibco) and 2 mM L-glutamate (Merck)) and incubated at 37°C before transfection. The following day, cells were transfected using the LT-1

transfection reagent (as described in Section 3.2.2)(Mirus Bio), with 6.75 μg pMD2.G (VSV-G envelope expressing plasmid; Addgene plasmid #12259), 2.5 μg psPAX2 (viral packaging vector; Addgene plasmid #12260) and 9 μg of desired plasmid vector to be transduced the culture dish of transfected HEK293FT cells was then incubated for 48 hours. After this incubation, the media was replaced and every 24 hours following this, virus was harvested by transferring the growth media into sterile Falcon tubes. In total, viral harvests took place 3 times before the virus containing media was centrifuged at 400 x g for 5 minutes, with the supernatant then filtered using a Low Protein Binding Durapore® (PVDF) 0.45 μm Membrane (Millex®). The viral containing media was divided into 1 mL aliquots in cryogenic tubes (Thermo Scientific) and stored at -80 ^{0}C.

Mammalian cell transductions

Cells were seeded into 90mm culture dishes in full media, to which 1-2 mL of the virus-containing media was added and returned to incubate at 37 ^{0}C for 48 hours. Following this, the full media was removed and replaced with full media containing 5 μg/μL puromycin to provide a continuous selection pressure for cells transduced with the desired plasmid. Following selection, colonies origionating from single cells were picked, trypsanised and transferred to 6-well culture dishes before being returned to incubate. Once confluent cells were treated with Doxycycline, to induce transgene expression, as described in Table 3.2 and GFP expression assessed via confocal fluorescent microscopy (Section 4.3.8 below) following 48-hour incubation. GFP positive cell poptilations were then passaged, with half added to culture flasks and the remaining cells pelleted and harvested using SDS sample buffer and subsequent western blot analysis of protein expression.

4.3.7 CRISPR/Cas9 mediated genetic manipulation and genomic edit analysis.

To facilitate CRISPR-mediated genomic editing of the *SPOP* gene, a previously generated in-house CWR22Rv1-AR-EK-iCas9 stable cell line derivative was used which express Cas9 under the control of a tetracycline promoter [Figure 4.2, 4.5]. CWR22Rv1-AR-EK-iCas9 cells were seeded onto 6-well culture plates for 24 hours then treated with Doxycycline as described in Table 3.1 for 48 hours. Following this, cells were transfected with either *SPOP*-targeting sgRNA

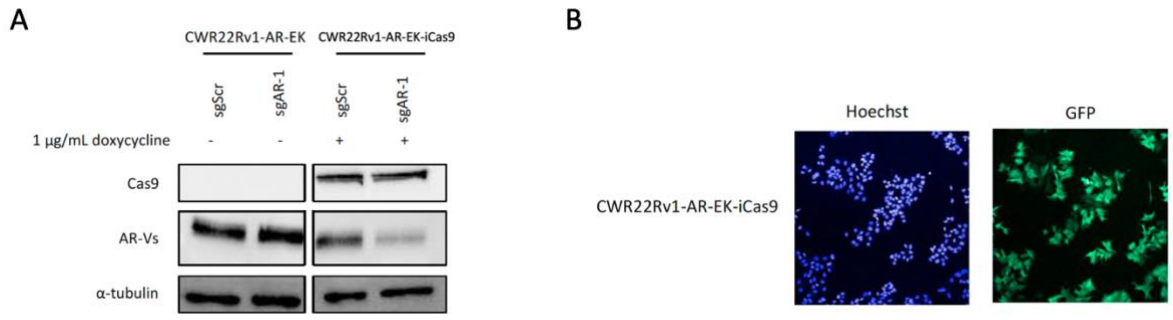


Figure 4.2 - Previous in-house development and validation of inducible Cas9 CWR22Rv1-AR-EK cells. A. *Cas9 and AR-V protein levels assessed within both CWR22Rv1-AR-EK-iCas9 and the parental CWR22Rv1-AR-EK cell lines by western blot analysis. Cells were treated with Dox or vehicle control 24 hours before transfection with sgScr or sgAR (AR Exon1) targeting synthetic guide RNA and 48-hour subsequent incubation. Data is representative of 3 individual experiments.* ***B.*** *Fluorescence microscopy of CWR22Rv1-AR-EK-iCas9 cells stimulated with Dox, as before, demonstrating GFP expression alongside nuclear Hoeschst staining. Data is representative of 3 individual experiments.*

or non-specific sgRNA as in Section 3.2.1 and returned to incubate for a further 48 hours. For analysis of sgRNA impacts on protein levels within cell populations, cells were harvested with SDS sample buffer following the 48 hour incubation and protein lysates underwent SDS-PAGE gel electrophoriesis as in Section 3.21. Analysis of insertion/deletion (INDEL) efficiency was analysed by ICE analysis (below) following sanger sequencing for PCR amplicons as detailed above in Section 4.3.3/4/5.

ICE Analysis

Once CRISPR-targeted *SPOP* amplicons had been generated, gel extracted and sent for Sanger sequencing as in Section 4.3.5, trace files generated from sequencing reactions were used to assess indel percentage and guide efficiency. To do this the ICE analysis tool provided by Synthego (available from ice.synthego.com) was used, using default parameters. Each experimental trace was compered to control trace files, generated from samples transfected with non-targeting sgRNA.

4.3.8 Live cell confocal imaging

To assess GFP fluorescence of cells following doxycycline induction of the pTLCV2-derived Cas9-T2A-GFP cassette stably-integrated into the CWR22Rv1-AR-EK-iCas9 derivative, live cell confocal imaging was used. Here, after 48-hour doxycycline induction, cells were observed for GFP induction which indicated expression of the pTLCV2 plasmid backbone [Figure 4.5].

Fluorescent output of cells was imaged using a NE2000 microscope (Nikon) and imaged in both brightfield and FITC channels.

4.4 Results

4.4.1 Validation of reagents to facilitate SPOP depletion in PC cells.

Anti-SPOP Antibody Validation.

One of the key reagents to be validated was an anti-SPOP antibody which could provide selective and consistent detection of SPOP via western blot. Many publications who have reported manipulating SPOP levels have utilised either expression of FLAG- or HA-tagged SPOP to reliably detect ectopic protein via anti- FLAG or -HA antibodies; or SPOP targeting antibodies. Of those used in the literature, common antibodies include Santa-Cruz (sc-377206), Abcam (ab137537) and Proteintech (16750-1-AP). To validate which of these would be suitable for use in this project, SPOP detection was attempted using these reagents from whole cell protein lysates from a range of PC cell lines [Figure 4.3].

The first Ab tested was the Santa-Cruz sc-377206 [Figure 4.3A-C]. As detailed by the manufacturer, banding was expected at 42 kDa [Figure 4.3A]. However, when run in-house no consistent banding was seen across a range of western blots run, with non-specific bands seen at 35 kDa [Figure 4.3B]. Only once was the expected banding seen however this was unable to be repeated consistently [Figure 3C]. Disappointingly, a similar trend of results was seen with both Abcam antibodies, with only non-specific bands seen from all of the western blotting methods used [Figure 4.3D&E]. However, use of the Proteintech anti-SPOP Ab was successful for SPOP detection at the predicted height seen [Figure 4.3F&G]. Although there were still non-specific bands seen across the membrane, this Ab was deemed the best of those tested and thus used throughout the rest of the project as it resulted in highly reproducible clear SPOP banding at the expected height. This was further validated with siRNA-mediated depletion of SPOP as described below.

SPOP targeting siRNA validation

To phenocopy the loss of function effects of SPOP mutations seen within patients, as applied in previous publications, knockdown of SPOP using SPOP-targeting siRNA transfection was used. Here, a pool of 4 SPOP-targeting siRNAs was purchased from Dharmacon and transfected into a range of PC cell lines to assess if this would facilitate a reduction in SPOP levels at both the transcript and protein level by qPCR and WB, respectively [Figure 4.4]. In all lines tested,

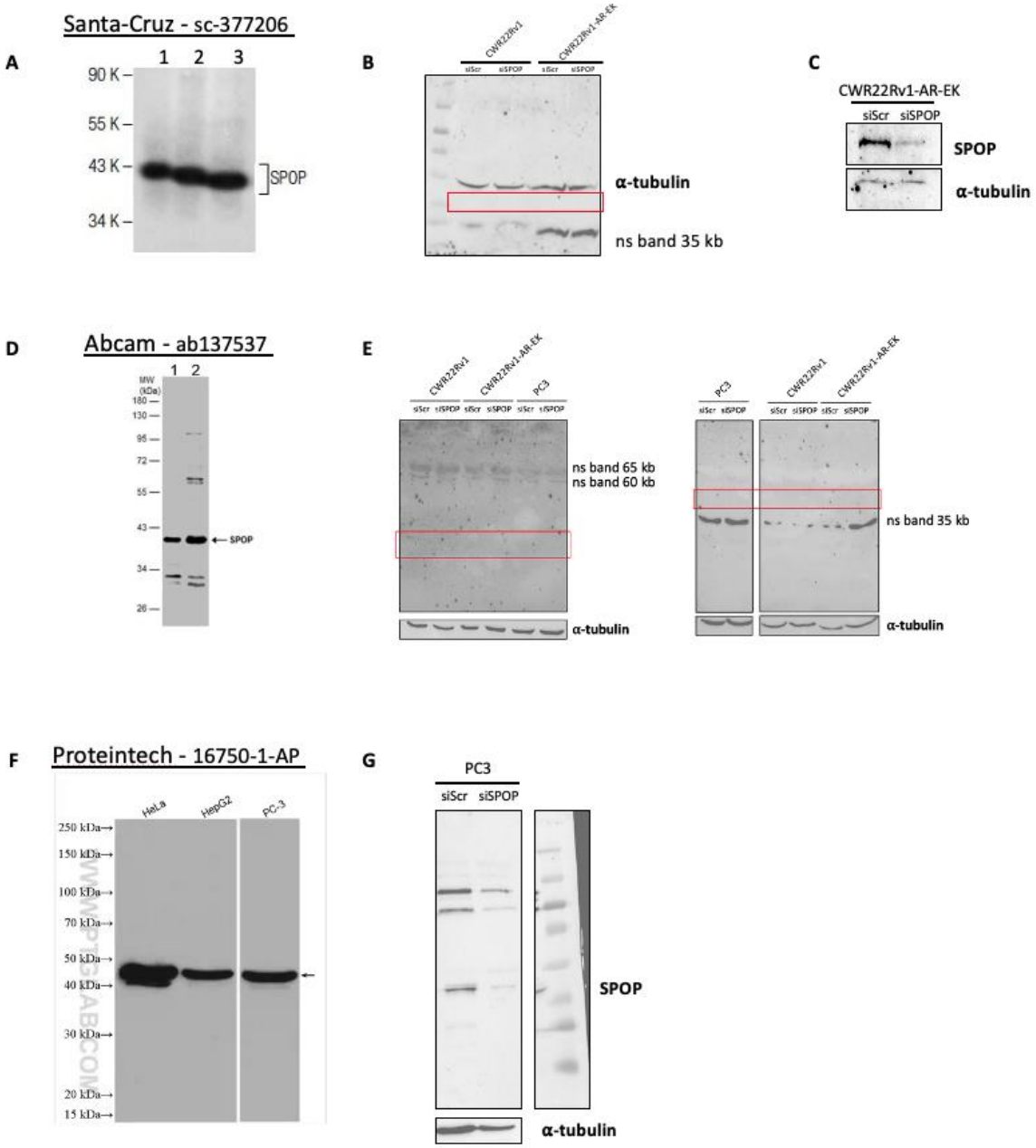


Figure 4.3 - Validation of SPOP targeting antibodies from various suppliers by western blot analysis. ***A***. *Image taken from Santa-cruz supplier website demonstrating expected banding of SPOP when using the sc-377206 antibody.* ***B***. *Experimental western blot results showing full membrane with non-specific (ns) banding seen at 35 kDa. Red box highlights expected band height expected for SPOP as outlined from supplier image.* ***C***. *SPOP protein levels assessed within both CWR22Rv1-AR-EK cells transfected with wither non-specific (siScr)or SPOP targeting (siSPOP) siRNA to diminish SPOP protein levels. Data is representative of 1 experimental/technical repeat.* ***D***. *Image taken from Abcam supplier website demonstrating expected banding of SPOP when using the ab137537 antibody.* ***E***. *Experimental western blot results showing full membrane with non-specific (ns) banding seen at 65, 60 and 35 kDa. Red box highlights expected band height expected for SPOP as outlined from supplier image.* ***F.*** *Image taken from Proteintech supplier website demonstrating expected banding of SPOP when using the 16750-1-AP antibody.* ***G.*** *Full membrane from experimental western blot results showing SPOP protein levels from CWR22Rv1-AR-EK cells treated with either siScr or siSPOP with non-specific (ns) banding seen at 65, 60 and 35 kDa.*

the level of SPOP protein level depletion observed was approximately 30-80%, depending on cell line, via western blot incorporating the Proteintech anti-SPOP antibody. Subsequent analysis of SPOP transcript levels by RT-qPCR indicated an approximate 45% reduction in mRNA levels across the cell lines tested. These results were seen 48 hours following transfection and

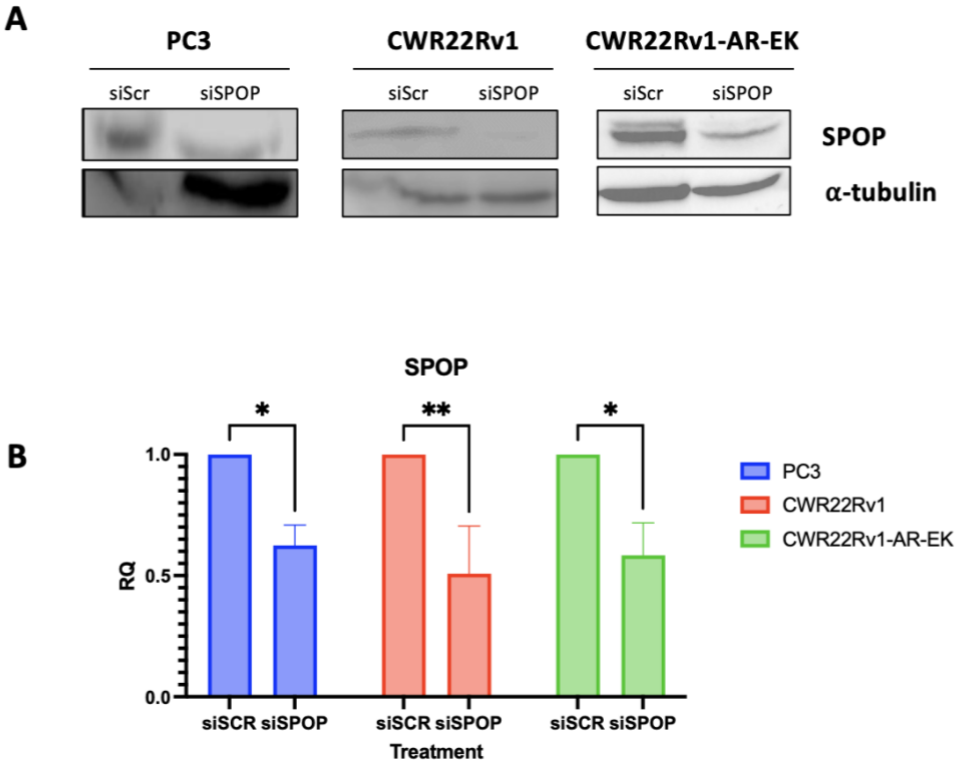


Figure 4.4 – Validation of SPOP targeting siRNA. *SPOP protein levels assessed vis western blot analysis (**A**) and mRNA levels assessed via RT-qPCR (**B**) from PC3, CWR22Rv1 and CWR22Rv1-AR-EK cells 48-hours post-transfection with either siScr or siSPOP siRNA. Western blot data is representative of 3 individual experiments. RT-qPCR data is representative of 3 individual experiments ± SEM. Statistical significance was determined by T-tests (*, **, *** and **** denote P-values <0.0332, <0.0021, <0.0002 and <0.0001 respectively, ns denotes not significant).*

thus a minimum 48 hour incubation post-transfection was used in all following experimental set-ups.

4.4.2 Generation of an inducible Streptavidin-tagged Cas9 prostate cancer CWR22Rv1-AR-EK cell line.

Whilst SPOP depletion via siRNA transfection will phenotypically replicate the loss of function conferred by *SPOP* mutations in patient samples the remaining ~30% population of SPOP protein may remain active masking some of the effects of SPOP loss. Thus, a potential better model to further investigate *SPOP* mutations would be to knock-in one of the physiologically relevant mutations using CRISPR/Cas9 machinery. Multiple factors can feed into the efficiency of generating these knock-in lines, including sgRNA targeting efficiency, template strand abundance and the chance of HDR being used to incorporate the donor template of interest. To increase the likelihood of efficient HDR, streptavidin tagged Cas9 can be used alongside

biotinylated template donor DNA. This will position the template donor sequence in close proximity to Cas9 for efficient donor DNA incorporation. Thus, efforts were made to generate a stable CWR22Rv1-AR-EK cell line derivative which can, upon tetracycline induction, inducibly express a FLAG-tagged streptavidin-Cas9 using a pTLCV2-Cas9 backbone (Addgene plasmid #87360) derivative [Figure 4.5A]. Following long-term puromycin selection for transduced cell populations and single colony clonal selection, as outlined in Section 4.3.6, cells were induced to assess streptavidin-Cas9 expression using an anti-FLAG antibody. As shown in Figure 4.5, each of the screened clones expressed substantial amounts of streptavidin-Cas9 upon Dox induction, allowing for potential use in downstream knock-in cell line generation attempts.

4.4.3 Generation of inducible wild-type and mutant SPOP overexpression plasmids

Cloning strategy and plasmid validation.

To generate PC cell line derivatives which stably and inducibly overexpress either $SPOP^{WT}$ or mutant $SPOP^{W131G}$, the pTLCV2-StrepCas9 vector was used as a recipient backbone for both wild-type and mutant SPOP cDNAs [Figure 4.6A]. Using *Bam* H1 and *Age* 1 restriction sites

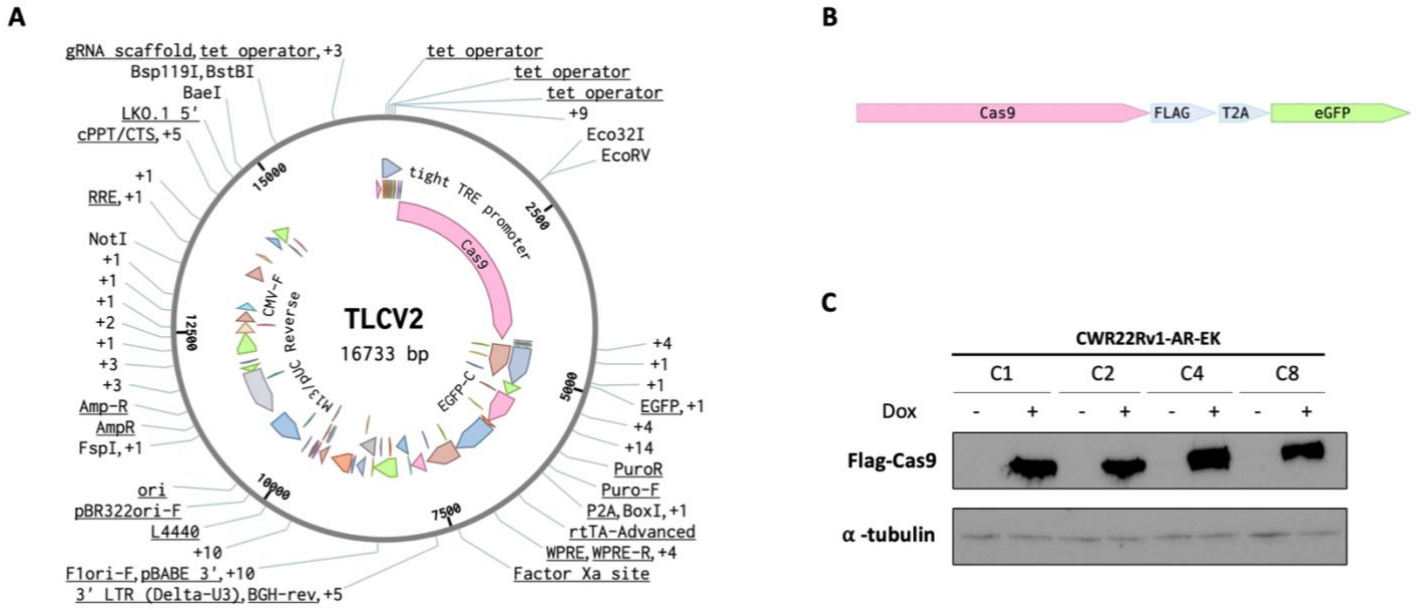


Figure 4.5 - Validation of inducible Strep-Cas9 expressing CWR22Rv1-AR-EK cells. A. *Schematic map of pTLCV2 plasmid transfected into cells.* ***B.*** *Linear map of mRNA expressed upon Dox induction of plasmid backbone.* ***C.*** *FLAG-Cas9 protein expression on CWR22Rv1-AR-EK single cell clones transfected with pTLCV2 plasmid by western blot analysis. Cells were treated with Dox or vehicle control and returned to incubate before protein lysates were harvested. Data is representative of 3 individual experiments.*

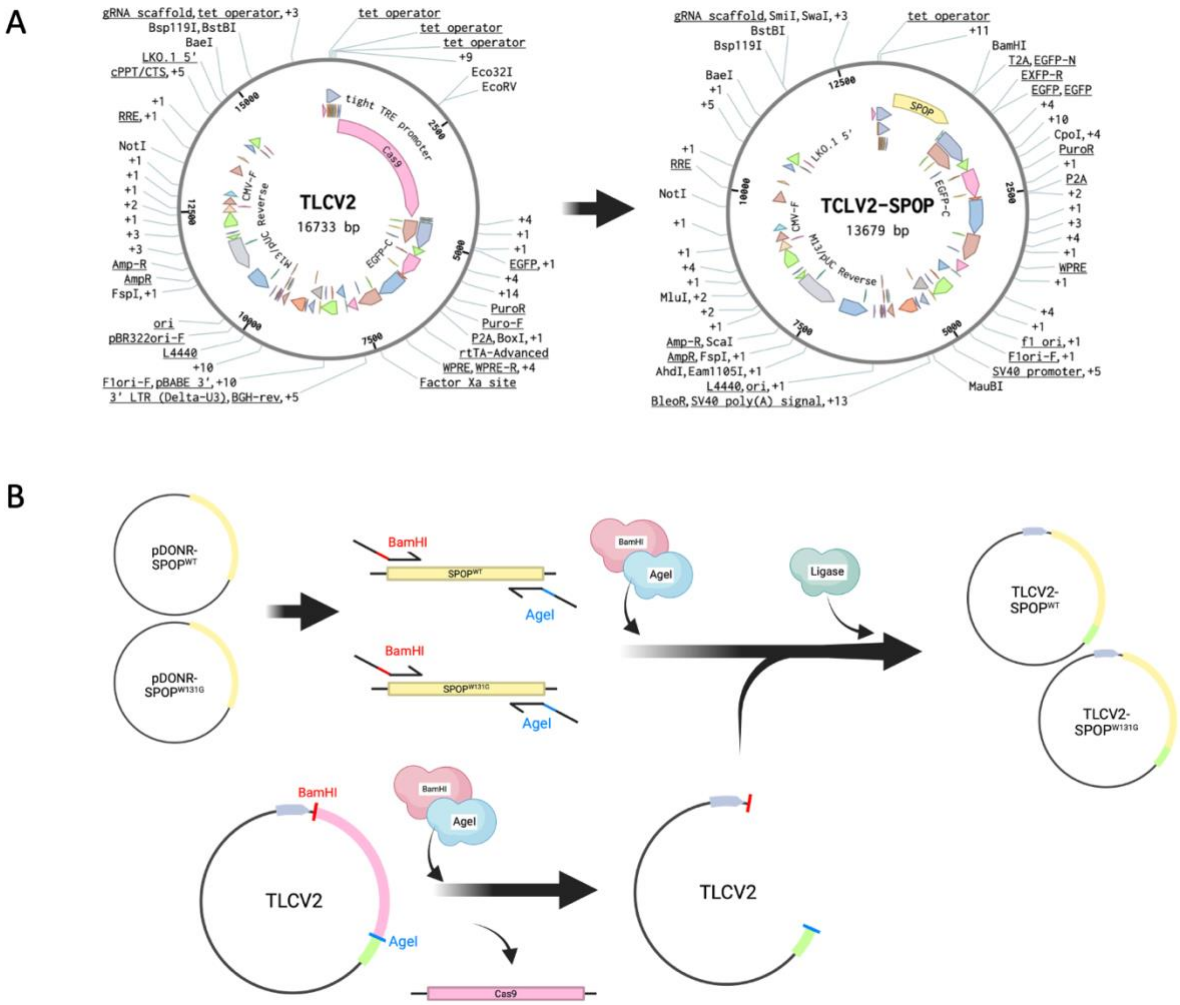


Figure 4.6 -Cloning strategy for generation of inducible SPOP-WT/MT plasmids. A. *Schematic plasmid maps of desired final plasmid maps.* ***B.*** *Cloning methodology used to remove Cas9-FLAG from the original TLCV2 vector and relace with the SPOPWT/SPOPW131G from pDONR vectors.*

flanking the FLAG-tagged-Cas9 within pTLCV2, Cas9 was excised from the backbone leaving pTLCV2 ready for receiving SPOP cDNAs [Figure 4.6B]. $SPOP^{WT}$ or $SPOP^{W131G}$ PCR amplicons were generated using the SPOP cDNA-containing pDONR223_SPOP_WT (Addgene plasmid #81856) and pDONR223_SPOP_p.W131G (Addgene plasmid #81632) as templates, with primers incorporating 5' *Bam* H1 and 3' *Age* 1 sites, to allow for ligation into the digested pTLCV2 vector upstream and in-frame with the T2A-GFP cDNA. The resultant amplicons were digested with both the *Bam* H1 and *Age* 1 enzymes and ligated into the digested pTLCV2 backbone.

To validate that the SPOP amplicons had ligated into the backbone, the undigested pTCLV2-SPOP plasmid was run on an agarose gel alongside pTCLV2-SPOP digested with both *Bam* H1 and *Age* 1, and the uncut and digested parental pTLCV2-Cas9 plasmid. As shown in Figure 4.7A,

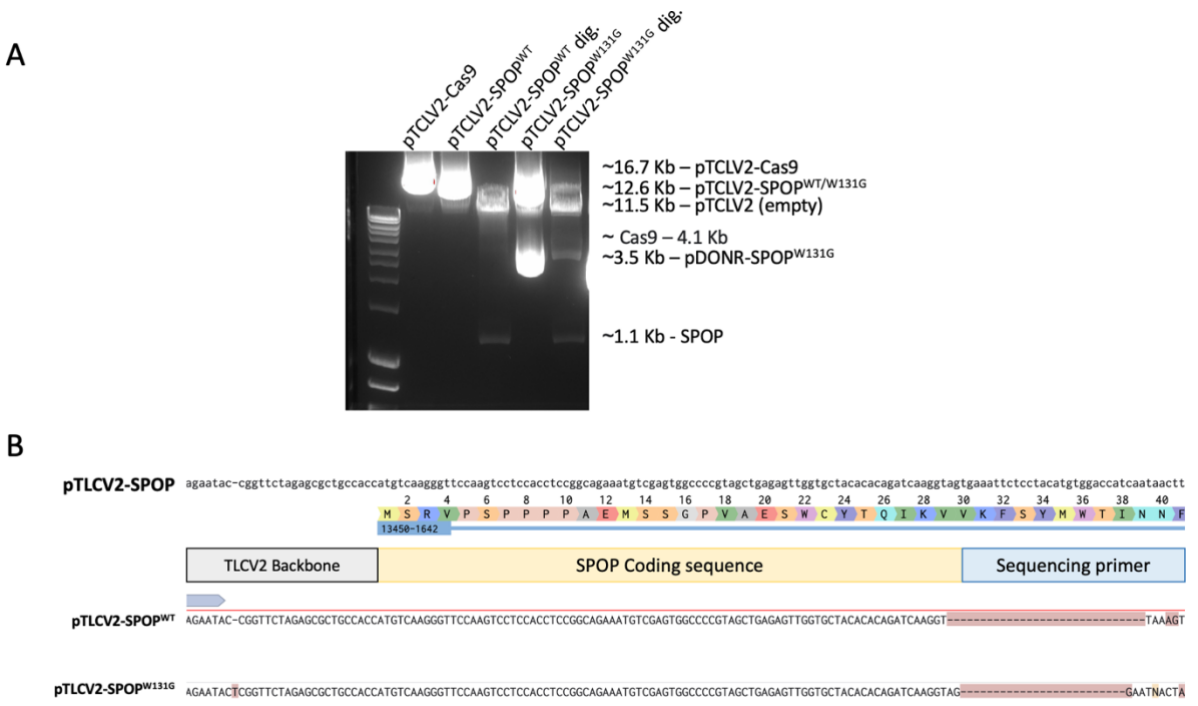


Figure 4.7 - Validation of new pTLCV2-SPOP plasmid constructs. A. *1% Agarose DNA electrophoresis of diagnostic digests of pTLCV2-Cas9, pTLCV2-SPOPWT and pTLCV2-SPOPW131G plasmid constructs with BamH1 and AgeI restriction enzymes. Suspected plasmid constructs and resultant digested products noted on right with kb nucleotide lenghs based of DNA ladder.* ***B.*** *Sanger sequencing results for the pTLCV2-SPOPWT and pTLCV2-SPOPW131G plasmid constructs aligned against the SPOP coding sequence generated from in silico cloning. Sequencing primer (blue) was positioned within the SPOP Coding sequence (yellow) running 3' to 5' into the pTLCV2 plasmid backbone to further validate the SPOP coding sequence had successfully been cloned into the pTLCV2 backbone.*

undigested pTLCV2-Cas9 plasmid is observed at 16.7 Kb while dual *Bam* H1 and *Age* 1 digestion results in generation of 11.5 Kb and 4.1 Kb fragments representing the digested backbone and released Cas9 sequence respectively. Importantly, for the newly generated pTLCV2-SPOP^{WT} and pTLCV2-SPOP^{W131G} plasmids, bands can be seen at 12.6 Kb representing the new plasmid constructs, which upon dual *Bam* H1 and *Age* I digestion released SPOP cDNA at 1.1 Kb. As a second form of validation, preps of the pTCLV2-SPOP^{WT} and pTCLV2-SPOP^{MT} were sent for sanger sequencing using a sequencing primer that annealed within SPOP Exon 1 and ran 3' to 5' through the ligation site into the pTLCV2 backbone [Figure 4.7B]. Of note, sequencing results from the pTLCV2-SPOP^{W131G} plasmid looks to have a single nucleotide insertion in the TLCV2 backbone. Downstream analysis of protein expression from this vector (below) will confirm whether this insertion will effect the expression of functional SPOP^{W131G} or is an artefact from sequencing runs. Taken together these show that both the full SPOP coding sequence had been successfully cloned into the pTLCV2 backbone.

Transfection of new constructs into HEK293T cells for functional validation via western blot and confocal microscopy.

To further validate these newly constructed plasmids were facilitating the expression of either WT or MT SPOP within cell lines, HEK293T cells were transfected with both plasmids. As described above, the expression of the pTCLV2-SPOP$^{WT/W131G}$ constructs are under the control of a tetracycline promoter, thus addition of doxycycline should trigger expression of these plasmids. As such, transfected HEK293T cells were induced with 1 μg/ml doxycycline for 24 hours prior to confocal microscopy imaging of GFP which is a surrogate marker of SPOP expression as a consequence of the SPOP-T2A-GFP cassette [Figure 4.8A]. Encouragingly, GFP expression in response to the doxycycline addition was observed, although some fluorescence was seen in the uninduced population suggestive of leaky transgene expression. In parallel, protein lysates were harvested from these cell populations and subject to western blot analysis to assess doxycycline-induced SPOP protein expression. Confirming the immunofluorescence observations, expression of SPOP was found to be markedly elevated in cells treated with doxycycline [Figure 4.8B]. Together, by virtue of reliable induction of SPOP expression as indicated by immunofluorescence, albeit indirect detection of GFP, and immunoblotting, these new SPOP expression constructs have been validated. They can now be implemented in downstream experimental procedure utilising either transient SPOP over-expression, or the generation of stable PC cell line derivatives which, upon induction, overexpress either SPOP^{WT} or SPOP^{W131G}.

4.4.4 Generation of inducible SPOP-expressing prostate cancer cell line derivatives

To further the repertoire of available models for downstream experimental set-ups, attempts to generate models of SPOP^{WT} or mutant overexpression were made, utilising the previously generated inducible SPOP^{WT} or SPOP^{W131G} constructs (Section 4.4.3). Inducible model generation was conducted within the PC3, CWR22Rv1 and CWR22Rv1-AR-EK PC cell lines however models only in the PC3 cell line were validated and used for downstream experiments [Figure 4.9]. As described in Section 4.3.6, cells underwent viral transduction using the pTLVC2-SPOP^{WT} or pTLCV2-SPOP^{131G} plasmids, and subsequent clonal selection. Initial clone screening was performed via western blot analysis to assess the cells ability to reliably induce SPOP

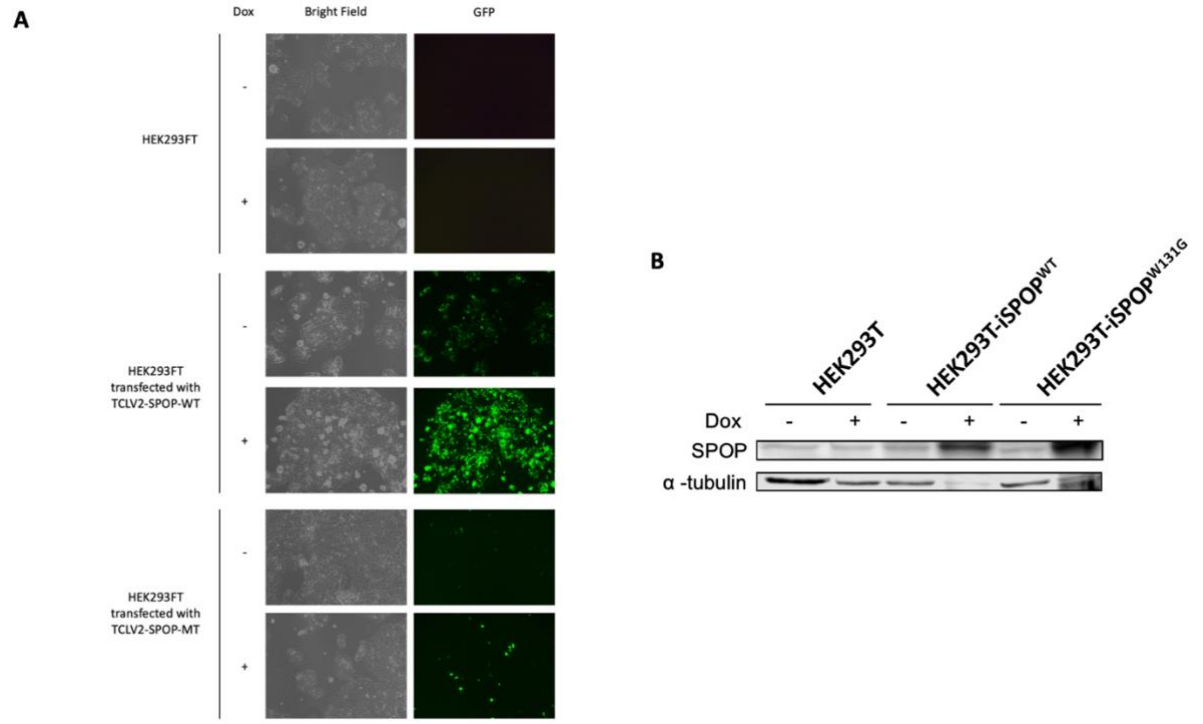


Figure 4.8 – Validation of pTLCV2-SPOPWT/W131G constructs ability to inducibly express SPOP within mammalian cells. A. *Confocal microscopy images of HEK293T cells transfected with either pTLCV2-SPOPWT and pTLCV2-SPOPW131G constructs, then treated with Dox or vehicle control and returned to incubate for 24 hours before GFP expression was assessed as a surrogate marker for SPOP protein expression. Data is representative of 3 individual experiments.* ***B.*** *Following confocal imaging protein lysates were harvested from cell populations and underwent SPOP protein level analysis by western blot. Data is representative of 3 individual experiments.*

protein expression following Dox induction [Figure 4.9A,B]. From the clonal populations transfected with the pTLCV2-SPOP^{WT} plasmids, clones 1, 4, 5 and 6 demonstrated clear induction of SPOP expression upon Doc induction, and as such stocks were taken and stored for each, with Clone 1 being selected for use in downstream experiments and referred to here-in as PC3-iSPOP^{WT}.

In the clonal populations assessed following pTLCV2-SPOP^{W131G} transfection, only clone 1 displayed notable induction of SPOP overexpression following Dox treatment, thus was taken forward for downstream experiments and here-in referred to as PC3-iSPOP^{W131G}. Further validation was conducted on these new cell derivatives by assessing GFP expression in response to Dox induction, which in both the PC3-iSPOP^{WT} and PC3-iSPOP^{W131G} cells showed clear GFP induction in the presence of Dox and no GFP expression in the absence of Dox, which reassuringly correlated with SPOP induction when these populations were harvested and SPOP protein levels assessed via western blot [Figure 4.9C,D]

With plans to develop prostate cancer cell line derivatives with capabilities of overexpressing endogenous SPOP in mind, a method of facilitating depletion of endogenous SPOP levels via siRNA transfection is desirable. Due to the SPOP expression vectors only containing the CDS and thus lacking the 3'UTR, a siRNA targeting a region within SPOP's 3'UTR would allow for

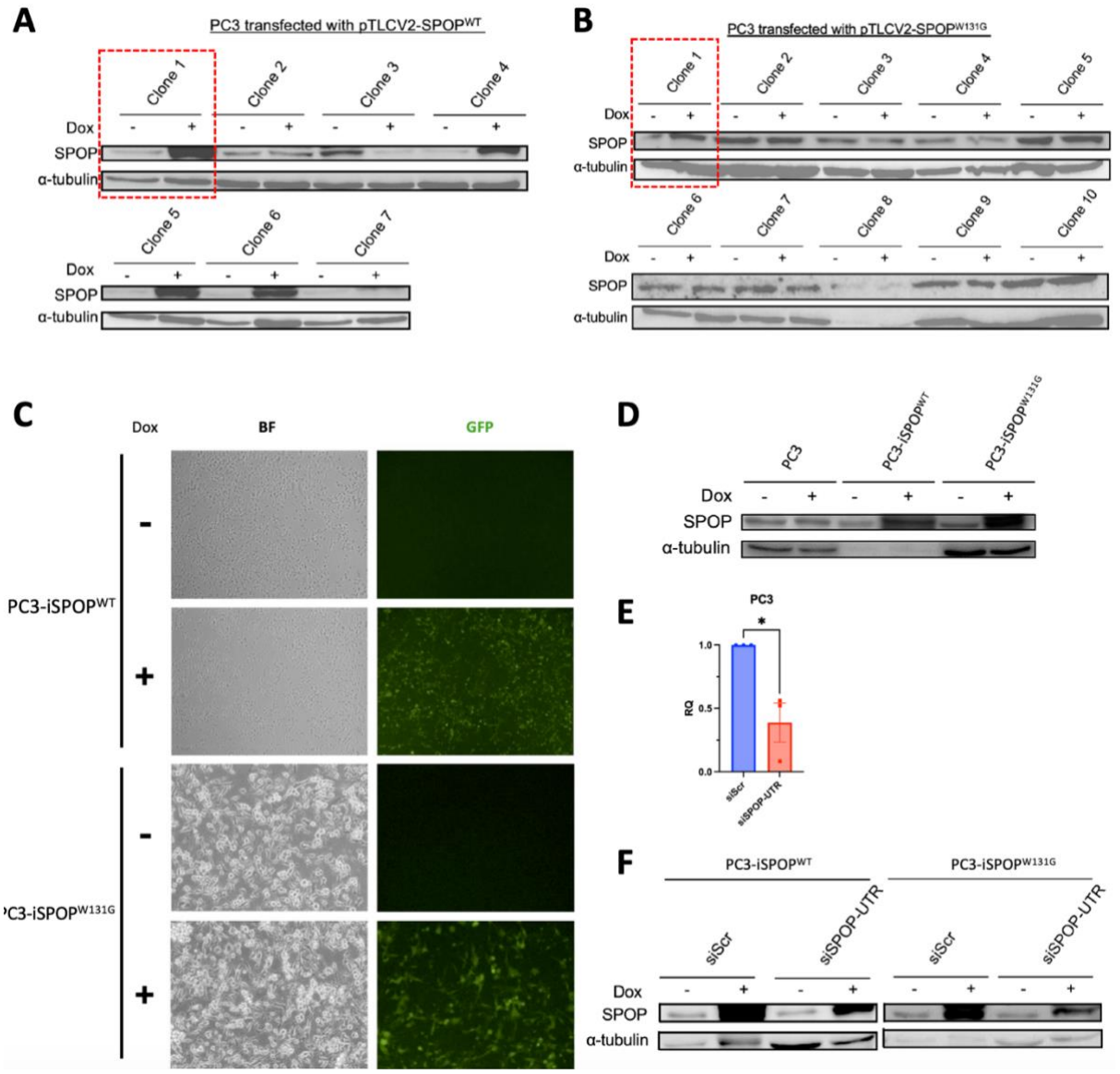


Figure 4.9 – Generation of PC3 cells capable of indelibly expressing SPOPWT or SPOPW131G using pTLCV2-SPOPWT/W131G plasmid constructs. A&B. *SPOP protein levels from multiple single cell clones transfected with either pTLCV2-SPOPWT or pTLCV2-SPOPW131G plasmid constructs and treated as before with Dox/vehicle control by western blot analysis. Red boxes highlight clones selected for further downstream experiments. Data is representative of 3 individual experiments.* ***C.*** *Confocal fluorescence microscopy of PC3-iSPOPWT and PC3-iSPOPW131G cells treated with Dox/vehicle control to assess GFP expression. Data is representative of 3 individual experiments.* ***D.*** *SPOP protein levels from protein lysates from C assessed via western blot analysis. Data is representative of 3 individual experiments.* ***E.*** *RT-qPCR analysis of SPOP mRNA levels PC3 cells transfected with either siScr or siSPOP-UTR and left to incubate for 72 hours. Data is representative of 3 individual experiments ± SEM. Statistical significance was determined by T-tests (*, **, *** and **** denote P-values <0.0332, <0.0021, <0.0002 and <0.0001 respectively, ns denotes not significant).* ***F.*** *SPOP protein levels following 72-hour siRNA transfection and 24hour Dox/vehicle control treatment assessed by western blot analysis. Data is representative of 3 individual experiments.*

knock-down of endogenous SPOP levels without affecting the exogenously generated SPOP levels. As before PC cell lines were transfected with this 3'UTR targeting siRNA (siSPOP-UTR) and the effects on transcript and protein levels were assessed [Figure 4.9E]. A robust depletion of both SPOP mRNA was seen 48 hours post transfection, providing confidence in this siRNA

showing similar effects to the pool used previously and allowing future experiments looking into rescuing phenotypes seen due to SPOP depletion to be conducted.

Finally, within the newly developed PC3-iSPOP^{WT} and PC3-iSPOP^{W131G} cell line derivatives an example rescue experimental set up was conducted to see if SPOP protein levels can be recovered following siSPOP-UTR mediated depletion [Figure 4.9F]. Promisingly within both models the siSPOP-UTR was able to reduce SPOP expression in the absence of Dox, with proteins levels exceeding endogenous levels upon addition of Dox, regardless of siRNA transfection. With these cell line derivatives and reagents validated, more thorough investigation into the effects of SPOP mutations within PC cell lines is capable with the ability to see if phenotypes observed from SPOP knock-down can be recovered by overexpressing either SPOP^{WT} or SPOP^{W131G}.

4.4.5 Generation of CWR22Rv1-AR-EK-SPOP$^{-/-}$ cell line derivatives using the CRISPR/Cas9 machinery

SPOP targeting sgRNA validation and effects on CWR22Rv1-AR-EK cells

sgRNAs targeting SPOP were designed using IDT, synthego and the BROAD institute (CRISPick) online tools using the *SPOP* gene as an input. From these, top design hits were collated and compared across the 3 tools used, with those common across the 3 being selected and ordered for validation. 3 sgRNA designs (here-in named sgSPOP-KO1, sgSPOP-KO2 and sgSPOP-KO3) all targeted *SPOP* exon 3; specifically sgSPOP-KO1/2 targeting the same PAM site, either annealing 5' or 3' to the PAM site, respectively, with sgSPOP-KO3 targeting a PAM site 24 bp downstream [Figure 4.10].

sgRNA transfection was performed as described in Section 3.2.1 using either each individual guide or a pool of all 3 in the CWR22Rv1-AR-EK-iCas9 cell line [Figure 4.2] [Figure 4.10], alongside non-specific sgRNA (sgSCR). Following 48 hours incubation, whole cell lysates were harvested to assess the effectiveness of these sgRNA designs in depleting SPOP protein levels via western blot. As shown in Figure 4.10B, all individual guides resulted in an appreciable reduction in SPOP levels; with sgSPOP-KO1 showing the greatest SPOP depletion of those tested. With the effects of SPOP knock-out on cell viability within CWR22Rv1-AR-EK cell line at the time unclear, SRB growth assays were performed, using this transient guide transfections to assess the effects of SPOP knock-out on cell proliferation. 120 hours after guide transfection,

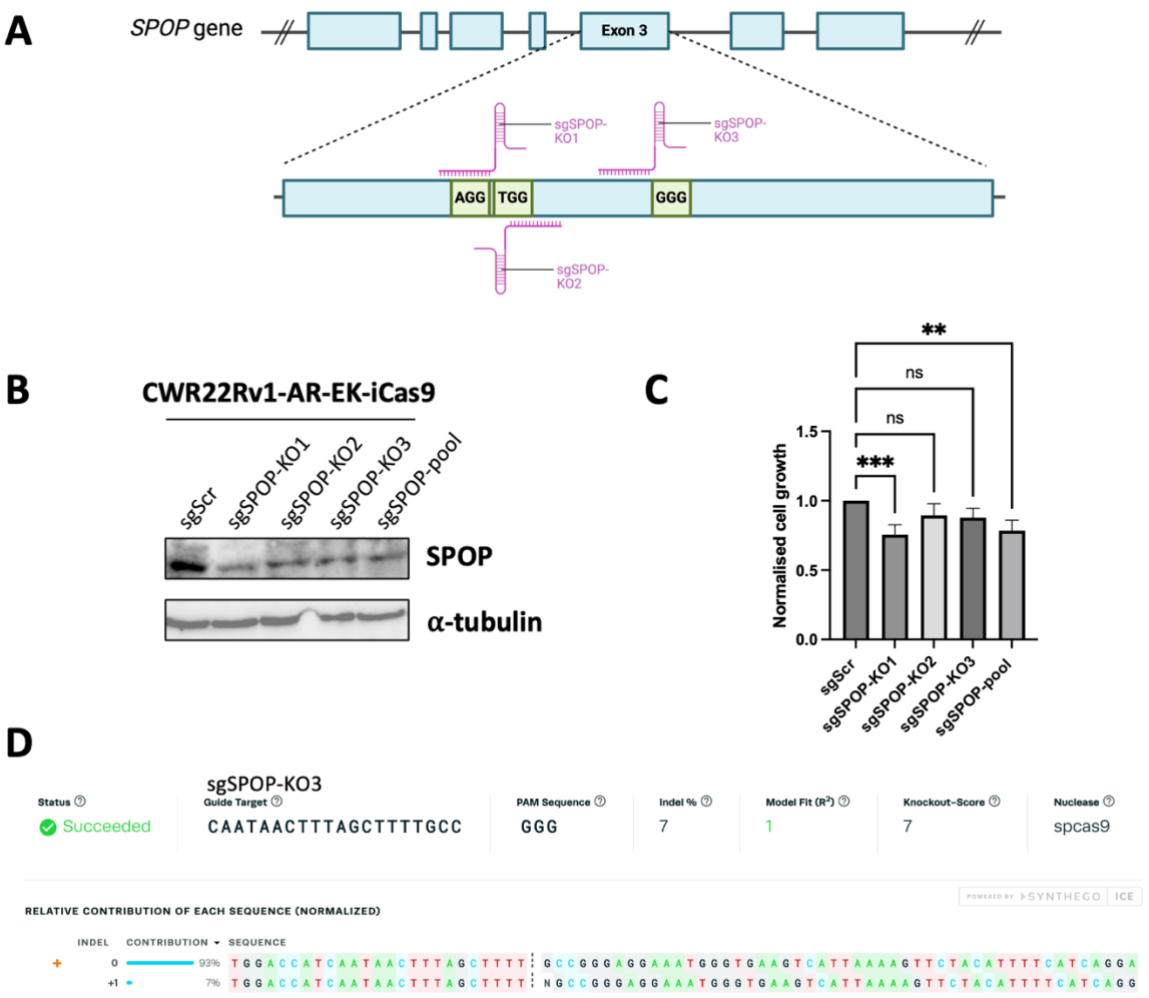


Figure 4.10 – SPOP targeting sgRNA validation and cell growth effects. A. *Schematic representation of target PAM sites for SPOP targeting sgRNA designs.* ***B.*** *SPOP protein levels of CWR22Rv1-AR-EK-iCas9 cells transfected for 48 hours with SPOP targeting sgRNA following 24 hour Dox induction via western blot. Data is representative of 3 individual experiments.* ***C.*** *Relative cell growth from CWR22Rv1-AR-EK-iCas9 transfected as above assessed by SRB assay. Data is representative of 3 individual experiments ± SEM. Statistical significance was determined by One way ANOVA (*, **, *** and **** denote P-values <0.0332, <0.0021, <0.0002 and <0.0001 respectively, ns denotes not significant). D. Representative ICE analysis for individual SPOP sgRNAs with associated indel percentage.*

CWR22Rv1-AR-EK-iCas9 cells were fixed and stained then relative cell density was quantified relative to sgScr cells [Figure 4.10C]. In contrast to sgSPOP-KO2 and -KO3 which had no effect on cell growth, sgSPOP-KO1 reduced proliferation by approximately 24.5% while the pooled guides reduced cell proliferation by 22.2% compared to the parental lines. Finally, using the same sgRNA transfection strategy as described previously, genomic DNA was harvested 48 hours after guide transfection. To simplify the process only the sgSPOP-pool and sgSCR treatments were used. Amplicons for SPOP-exon 3 were generated from the resultant genomic DNA preparations, and sent for sequencing and ICE analysis [Figure 4.10D]. Here an indel frequency of 22% was calculated for the sgSPOP-pool samples compared to the sgScr control, which is likely responsible for the reduced SPOP protein level observed via western blot.

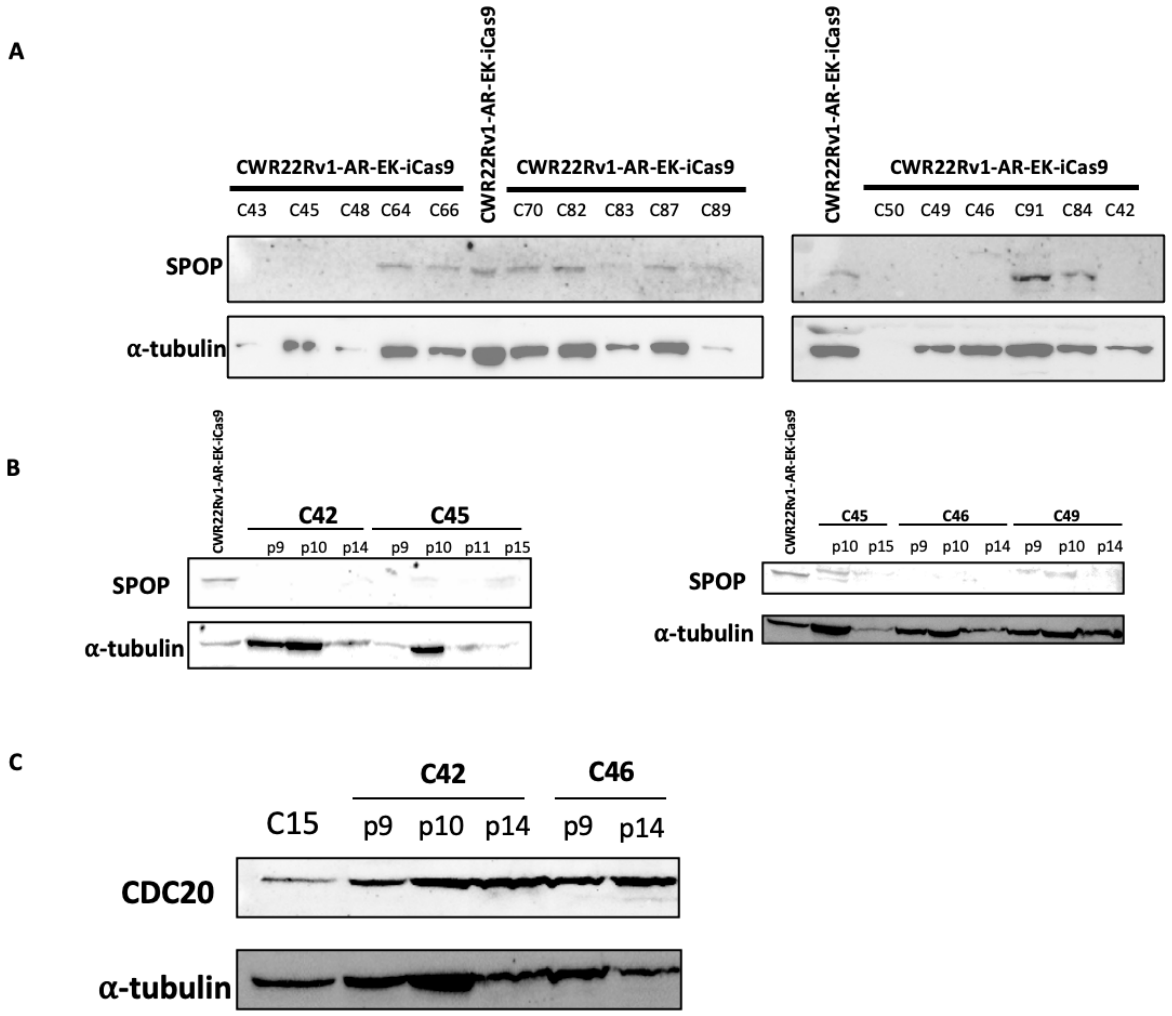


Figure 4.11 – Generation of SPOP knock-out CWR22Rv1-AR-EK cells. A&B *SPOP protein levels assessed within both CWR22Rv1-AR-EK cells transfected with SPOP targeting sgRNA and parental CWR22Rv1-AR-EK-iCas9 cells by western blot analysis. Data is representative of 3 individual experiments.* ***C.*** *CDC20 protein levels assessed within two CWR22Rv1-AR-EK-SPOP-/- clones and parental CWR22Rv1-AR-EK-iCas9 cells by western blot analysis. Data is representative of 3 individual experiments.*

Importantly, gRNAs to target SPOP have been validated to effectively induce genomic edits at the desired locus to impact SPOP protein abundance.

Stable CWR22Rv1-AR-EK-SPOP$^{-/-}$ cell line derivative generation using CRISPR-Cas9 machinery.

With data indicating that the 3 *SPOP*-targeting gRNAs were working to reduce SPOP protein levels after transient transfection, the next steps were to generate single cell clonal populations harbouring a *SPOP* knock-out. CWR22Rv1-AR-EK-iCas9 cells were transfected with pooled SPOP sgRNA as before and left to incubate for 48 hours. Following incubation, cells

were then seeded into larger 150 mm culture dishes at low density and left until colonies visible by eye had formed on the plates. These colonies were then picked using a sterile pipette tip and placed in individual wells, allowed to grow before screening via western blot. In total 144 clones were screened for reductions in SPOP protein levels; with those observed to have substantial SPOP protein reduction passaged twice more prior to SPOP protein analysis to assess if SPOP depletion was stable [Figure 4.11A]. Results revealed that although most clones showed a reduction in SPOP levels, only a small number of screened clones looked to have complete loss of SPOP, including C42, C45, C46, and C49, which were subsequently kept in culture and SPOP protein levels assessed over subsequent passages [Figure 4.11B]. Further analysis revealed that while C45 and C49 showed only partially depleted SPOP protein levels, C42 and C46 demonstrated complete loss of SPOP protein. As a result, C42 and C46 were taken on for further validation and will here-in be referred to as CWR22Rv1-AR-EK-SPOP$^{-/-}$-C1 and CWR22Rv1-AR-EK-SPOP$^{-/-}$-C2 respectively. Finally, protein lysates were ran via western blot on the CWR22Rv1-AR-EK-SPOP$^{-/-}$-C1 and CWR22Rv1-AR-EK-SPOP$^{-/-}$-C2 clones to assess if the effects of loss of SPOP protein within these lines correlated to substrate accumulation using CDC20, a previously reported ubiquitylated substrate of the SPOP-containing E3 ligase complex (Wu et al., 2017)[Figure 4.11C]. In both clones an accumulation of CDC20 can be seen across the range of passages tested when compared to the parental CWR22Rv1-AR-EK-iCas9, further confirming loss of SPOP expression.

To further validate these knock-out clones genomic DNA was harvested and amplicons generated around SPOP-Exon3, which were then sent for sequencing, specifically targeting SPOP sgRNA binding region and PAM site to assess the indels generated via ICE analysis as described in Section 4.3.5 [Figure 4.12]. Here both clones returned a 100% indel generated, with predominant contributions to the resultant indels formed by sgSPOP-KO3. CWR22Rv1-AR-EK-SPOP$^{-/-}$-C1 saw secondary contributions from sgSPOP-KO1 whereas CWR22Rv1-AR-EK-SPOP$^{-/-}$-C2 displayed secondary indel contributions from sgSPOP-KO2.

Figure 4.12 – ICE analysis from CWR22Rv1-AR-EK-SPOP-/- clones. *Ice analysis comparing sanger sequencing results from CWR22Rv1-AR-EK-SPOP-/--C1 (**A**) and CWR22Rv1-AR-EK-SPOP-/--C2 (**B**) against CWR22Rv1-AR-EK-iCas9 genomic DNA. g1, g2 and g3 represent sgSPOP-KO1, sgSPOP-KO2 and sgSPOP-KO3 respectively. Black dashed lines indicate cut sights where indels have been introduced from sgRNA.*

Together these results provide evidence towards the generation of stable CWR22Rv1-AR-EK cell derivatives harbouring complete and robust knock-out of SPOP at the protein level, which display the expected characteristics regarding substrate protein accumulation.

Finally, to assess due to the downstream assays being conducted on these cells, the effects of SPOP knockout on CWR22Rv1-AR-EK growth rates was determined, as if there were any substantial difference each line would have to be normalised to its own internal control as opposed to the parental control [Figure 4.13]. Here the parental CWR22Rv1-AR-EK-iCas9 displayed a doubling time of ~32 hours while the CWR22Rv1-AR-EK-SPOP$^{-/-}$-C1 cells had a doubling time of ~39 hours, with the parental showing significantly increased total growth by the experimental endpoint at 120 hours than CWR22Rv1-AR-EK-SPOP$^{-/-}$-C1 cells. As a result this will be considered in downstream experimental set-up and data normalisation.

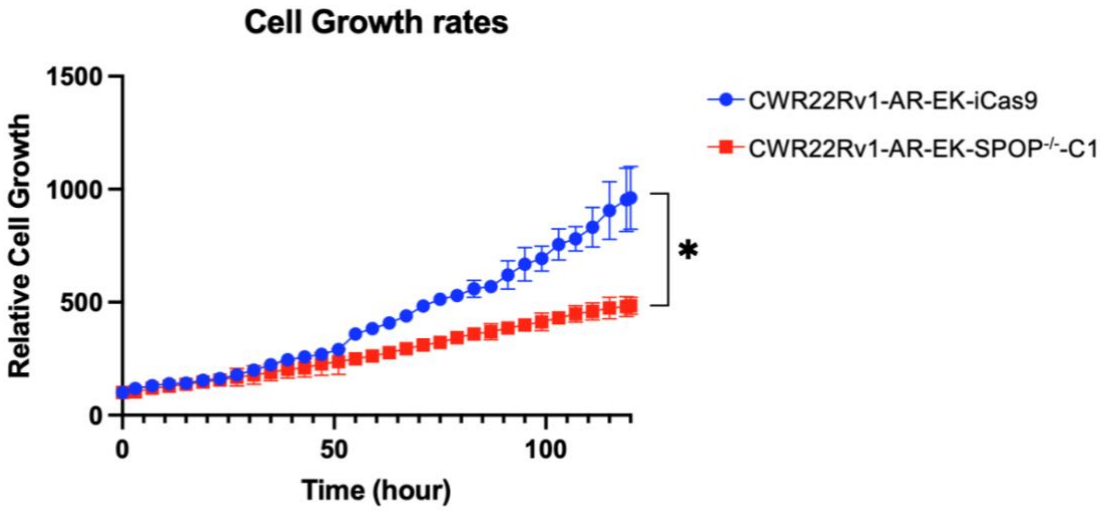


Figure 4.13 – Determination of cell growth rates of CWR22Rv1-AR-EK-iCas9 and CWR22Rv1-AR-EK-SPOP-/--C1 cells. *Cell growth from both CWR22Rv1-AR-EK-SPOP-/--C1 and parental CWR22Rv1-AR-EK-iCas9 cells across 120-hour time-course normalized relative to the cell density at 0-hours. Data is representative of 3 individual experiments ± SEM. Statistical significance was determined by T-tests between endpoint values (*, **, *** and **** denote P-values <0.0332, <0.0021, <0.0002 and <0.0001 respectively, ns denotes not significant).*

4.5 Discussion

The SPOP protein has been found to play key roles in normal cell homeostasis, notably cell proliferation, survival and DNA damage repair. Due to high prevalence of loss of SPOP function in PC patient cases, much work has been conducted to generate a deeper understanding into the specific intracellular effects that this loss causes and how it contributes to cancer progression. Ultimately these studies utilise a range of techniques to mimic this loss of function within PC cell lines, with the most predominant seen in published literature being overexpression of either wild-type or mutant SPOP protein, and siRNA or shRNA mediated SPOP depletion. These methods look to bring about loss of SPOP function within the described experiments, although there are caveats to each. With regards to overexpression of wild-type or mutant SPOP protein, overexpression of wild-type SPOP is commonly used as a control, however the effect of increased SPOP levels over that deemed steady-state/physiologically relevant may perturb normal cell homeostasis. Further, there may be roles of the SPOP protein that are outside its direct substrate interactions, thus SPOP depletion may not be representative of the effects caused by loss of function mutations alone. siRNA/shRNA mediated SPOP depletion may also not be representative of complete loss of function caused by SPOP mutations as these methods may not reduce SPOP protein expression completely, leaving functional SPOP within cells.

Development of new PC cell line models utilising alternative methods of SPOP manipulation could assist in generation of deeper understanding into the loss of function mutations seen in PC. With this in mind, work enclosed within this chapter looked to achieve the aims of validating a variety of SPOP targeting reagents and the generation of new models not previously described within the literature that would be extensively utilised in downstream experiments. Ultimately, development of experimental pipelines and new models for investigation of SPOP functional loss within PC cell lines that combine siRNA knock-down, overexpression of wild-type or mutant SPOP proteins, and through CRISPR-Cas9 mediated knock-in of relevant SPOP mutations and complete knock-out of the SPOP protein have been validated for future use.

One of the key reagents necessary for the project was a reliable anti-SPOP antibody that could be used to validate SPOP protein levels via western blot in response to a range of potential

treatments, however at the beginning of the project most of the published literature had utilised FLAG- or HA-tagged SPOP protein as surrogate markers for assessment of ectopic protein levels. In this study, three antibodies from Santa-Cruz, Abcam and Proteintech were tested with multiple whole protein lysates from PC cell lines. Through extensive testing with many modifications to protocol used, only the Proteintech antibody was able to provide reliable experimental results that persisted from batch-to-batch and was thus used for the remainder of western blot analysis of SPOP protein levels. Although the SPOP protein displayed clear banding at the molecular weight, there was also non-specific banding present between 60-65 kDa suggesting there may be off target interactions with the antibody making it unreliable for other applications, such as immunofluorescence and flow cytometry. Early experiments conducted also used SPOP targeting siRNA to reduce SPOP protein levels to phenocopy *SPOP* loss of function mutations seen within patient samples. The siRNA pool tested showed robust but not complete depletion of SPOP protein levels 48 hours post-transfection across the range of cell lines tested, meaning that the effects seen in response to SPOP knock-down may not be representative of complete loss of function.

With the aim of generating knock-in models, a range of sgRNA was tested within the CWR22Rv1-AR-EK-iCas9 cell line targeting PAM sites proximal to the clinically-relevant and commonly mutated Y87 and F133 sites within SPOP. Across the guides tested, very low INDEL percentages were observed and with no other guide designs available, alternate methods of increasing the effectiveness of these current guides would be necessary if PC cell lines harbouring *SPOP* mutations are to be created. To this end, a new cell line derivative in the CWR22Rv1-AR-EK background expressing Streptavidin-tagged Cas9 was generated, so that biotinylated sgRNA could be used to enhance Cas9 positioning and subsequent indel generation. Cell lines capable of doxycycline inducible streptavidin-tagged Cas9 were generated and validated by western blot. However further validation was not conducted due to time constraints and the development of alternative models.

An alternative method of introducing SPOP mutations, that has been extensively used within the literature, is the plasmid-based overexpression of either wild-type or mutant forms of the SPOP protein within PC cell lines. To create inducible plasmid-based expression systems $SPOP^{WT}$ or $SPOP^{W131G}$ coding sequences were cloned into the existing pTCLV2-Cas9 backbones in place of Cas9. Subsequently the protein expression within cell lines and surrogate GFP marker expression were validated for these plasmids. An issue noted from the diagnostic

digests is the presence of an unknown band is the presence of some pDONR-SPOP^{W131G} plasmid in stock of the pTCLV2-SPOP^{W131G}, this was deemed to ultimately insignificant as it would play no issue in downstream model development as they would be generated from clonal selection and the pDONR lacks the puromycin selection marker. This does, however, have an effect on the expression of both SPOP and GFP when transfected into the HEK293T cells when validating the inducible expression, due to the 1 µg of total plasmid transfected into the cells being split between the two plasmids, resulting in seemingly less expression of both proteins from the pTLCV2-SPOP^{W131G} when compared to the pTLCV2-iSPOP^{WT}.

With the low chance of generating knock-in mutations with the available repertoire of guide designs, introducing the inducible expression of either wild-type or mutant forms of the SPOP within the knock-out background may represent an alternative means of generating PC cell lines that express only mutant SPOP. Although this work was unable to be carried out within the timescale of the project, validation of rescue experimental set-ups using SPOP-3'UTR targeting siRNA were validated using the inducible systems in the PC3 cell background. Clonal populations were generated following transfection with either the pTLCV2-iSPOP^{WT} or pTLCV2-SPOP^{W131G} plasmids and inducible expression of the SPOP protein was validated by western blot. Furthermore, a new siRNA design was validated that was capable of targeting the 3'-UTR region on SPOP RNA allowing for knockdown of endogenous SPOP protein whilst not affecting exogenous protein expression driven from the plasmid constructs. By validating this rescue-based experimental system there is potential for much deeper interrogation of any phenotypic or mechanistic effects seen in response to SPOP siRNA knock-down or knock-out.

Seeing as the INDEL efficiency of the guides used to generate knock-in mutations were low, new guides were tested targeting exon 3 of *SPOP* in attempts to knock-out SPOP protein expression. Individual guides were able to reduce protein expression and cause minimal loss in cell proliferative ability and were taken further to generate clonal populations of CWR22Rv1-AR-EK-iCas9 cells that harbour a complete SPOP protein knock-out. As such, two independent clones were generated that displayed both complete loss of SPOP protein expression and independent indels introduced in each, but importantly showed accumulation of one of the known SPOP substrate proteins CDC20 (Wu et al., 2017). Minimal information about the effects of complete SPOP loss within PC cell lines has been provided from the literature, with the technique mainly being used to assess effects on specific mechanisms and protein interactions rather than looking at broader effects on cell homeostasis as a whole. With the

expected increase in CDC20 seen in these cells, it can be assumed that mechanistically the cells are behaving as expected in response to SPOP loss, and the same accumulation can be seen in other SPOP substrates. With the AR being a key driver of prostate cancer progression, and a known SPOP substrate it would be ideal look at AR accumulation in an SPOP knock-out background. However, due to the CWR22Rv1-AR-EK cell line expressing only AR-Vs which lack the hinge region containing the SPOP binding motif, investigation into the effects on AR signalling are unavailable within this model. Future knock-out models in alternate PC background would allow for deeper investigation into the contributions of SPOP mutations into AR biology within PC however none were able to be generated within the timeframe of this project.

Finally, the effects of SPOP loss had on the proliferation rates of cells was assessed, with significant reductions seen in cell growth across a 120-hour time course. Ultimately this will have effects when it comes to using these two cell lines for proliferative assays as the effects of SPOP loss may mask the actual effects caused by treatments to cells. As such, data will have to be normalised to the vehicle control in each cell line downstream only for proliferation-based assays, meaning any significant reductions in cell viability are due to prescribed treatments rather than the cells slower growth rates caused by SPOP loss.

Through answering the initial aims of this chapter, together the reagent validation and model development conducted here will be of considerable use in the next stages of the project. As evidenced through the later chapters, the use of these new models has allowed much deeper investigation into the experimental effects of drug treatments in a range of SPOP 'mutant' backgrounds, providing much greater confidence in the data generated. Alongside this, use of the new CWR22Rv1-AR-EK-SPOP$^{-/-}$ models has allowed for great deals of information to be generated about the effects of SPOP loss in response to DNA-PKcsi treatment, as well as analysis of general transcriptomic effects informing future studies into potential new drug susceptibilities.

Chapter 5 – SPOP mutations as a prostate cancer biomarker for radiosensitivity following DNA-PKcs inhibitor treatment.

5.1 Introduction

5.1.1 SPOP and DNA damage repair

SPOP is essential for the mRNA expression of many DNA repair related genes (e.g., ATR, BRCA2, ChK1, RAD51), which was supported by proteomics data (Hjorth-Jensen et al., 2018). This study also found SPOP-WT interacted with proteins involved in many other cellular processes such as splicing (Hjorth-Jensen et al., 2018). SPOP was initially implicated in DNA damage repair signalling in vitro by Zhang et al., 2014 and was later confirmed to be functional in this pathway in vivo (Bernasocchi et al., 2021). SPOP foci are seen to accumulate in an ATM dependent manner around sites of DNA damage with SPOP depletion being linked with decreased RAD51 foci formation and reduced Chk1 activation (Zhang et al., 2014a). Further, in response to DNA damage, ATM has been found to phosphorylate S119 of SPOP resulting in SPOP-mediated degradation of cell division-related proteins and non-degrative ubiquitination of 53BP1. As such, this promotes the use of HR for DSB repair over NHEJ (Xiao et al., 2021). Ultimately, SPOP has been found to be a key regulator in HR of DNA damage, and mutations within the SPOP MATH domain resulted in cells displaying a 'BRCA-like' phenotype (Boysen et al., 2015).

One of the key findings in the work presented by Boysen et al., 2015 was from analysis of SPOP mutant cancers compared to wild type populations. Here significant increased abundance of DNA rearrangements were observed in SPOP mutant patients as well as higher total copy number alterations (Boysen et al., 2015). These findings are consistent with the transcriptomic changes seen in SPOP mutant cells promoting error-prone DNA repair as the principle means of DSB repair, leading to the increases in genomic alterations.

SPOP mutations confer sensitivity to DDR targeting agents

In *BRCA1*-deficient breast and ovarian cancer, cells are sensitive to poly (ADP-ribose) polymerase 1 (PARP) inhibition. This interaction is a clear demonstration of the concept of synthetic lethality, where combined aberrant activity within two genes or proteins results in

cell death, where alterations in activity individually do not. Here, loss of PARP1 through its inhibition causes induction of the DDR through stalled replication forks, which in *BRCA1*-deficient cancer cells which have resultant HR deficiency, which often leads to cell death (Lord and Ashworth, 2017). Similarly, cells with compromised HR activity, as a result of *SPOP* mutations, share a *BRCA1* mutant phenotype and, as a consequence, are also sensitive to PARP inhibition (Boysen et al., 2015). This was validated in PC3, LNCaP and CWR22Rv1 PC cell lines depleted of SPOP by siRNA, to phenocopy SPOP loss-of-function mutations, and treated with either PARP inhibitors, Olaparib or veliparib. In each cell line, down-regulated SPOP expression resulted in sensitivity to both PARP inhibitors in both proliferative and colony forming assays. This was assessed further in CWR22Rv1 cells overexpressing the loss-of-function mutant $SPOP^{F133V}$. Consistent with the SPOP knockdown studies, Olaparib treatment resulted in reduced cell viability in cells ectopically-expressing the SPOP mutant as measured by clonogenics assays. Crucially, these findings contrasted with observations from PC cells over-expressing SPOP-WT, which found the efficacy of Olaparib was diminished (Boysen et al., 2015). Together, this data provides evidence that the presence of a single point mutation within the MATH domain of SPOP is sufficient to confer loss of SPOP activity and compromise resultant DNA damage repair. As such, SPOP mutation may present a therapeutic vulnerability in PC cells that could potentially be pharmaceutically exploited for individualised PC therapy in the clinical setting.

5.1.2 SPOP mutations co-occur with CHD1 mutations and confer PARP sensitivity

In normal tissue chromodomain helicase DNA-binding protein 1 (CHD1) plays roles in chromatin remodelling in an ATP-dependant manner localising to promoters of genes marked by H3K4 tri-methylation, as well as regulating histone turnover and nucleosome assembly, ultimately playing key roles in regulating gene transcription (Li et al., 2023a). However, CHD1 deletion is found in around 15% of PC cases (Barbieri et al., 2012, TCGA, 2015).

CHD1 contributes to the regulation of AR occupancy and for *ERG* rearrangement in PC (Blattner et al., 2014). CHD1 was found to be coenriched on chromatin alongside AR, whilst also found in complex with multiple nuclear receptor-specific cofactors (Augello et al., 2019). When CHD1 is lost in PC, changes in the AR transcriptome are observed due to restriction of the AR to sites enriched for HOXB13, ultimately driving pro-oncogenic pathways (Grbesa et al., 2021). CHD1

has also been reported to be essential for efficient HR by regulating 53BP1 stability and CtIP recruitment to sites of DNA damage; with loss causing DNA repair defects and enhanced sensitivity to PARP inhibitors (Kari et al., 2016). Interestingly *SPOP* mutations are highly correlated with CHD1 deletions, potentially explaining the mutual exclusivity between *SPOP* mutations and *ETS* rearrangements, such as the common *ERG* fusion, in PC (Blattner et al., 2014).

As would be expected based on their roles in DNA repair, defects in both *SPOP* and *CHD1* genes leads to a synergistically higher level of DNA damage in both mouse and human prostate epithelial cells (Zhu et al., 2021). Defects in both CHD1 and *SPOP* have a combinatory effect on increasing DNA damage in human prostate epithelial cells. Deletion of *CHD1* also causes DNA repair defects and sensitivity to PARP inhibitors (Kari et al., 2016).

5.1.3 SPOP mutations and AR signalling axis targeting therapies

A growing body of evidence suggests that PC-associated *SPOP* mutations result in increased effectiveness of AR signalling axis-targeted therapeutics, such as enzalutamide and abiraterone, likely due to the AR being a SPOP substrate and has shown both increases in AR protein levels and AR driven gene expression. In LAPC4 PC cells lines ectopically over-expressing of SPOP mutants confers elevated susceptibility to enzalutamide empty vector controls; suggesting *SPOP* mutant patients may respond better to AR-targeting agents (Bernasocchi et al., 2021). In support of this, an 89-patient study conducted by Boysen *et al*. (2018) demonstrated a strong correlation between *SPOP* mutation and loss of *CHD1* in mCRPC patients, and this potentiated increased sensitivity to abiraterone treatment (Boysen et al., 2018). Furthermore, a recent retrospective analysis of a Phase III efficacy study of enzalutamide in advanced PC provided evidence that *SPOP* mutant mCRPC patients are more likely to demonstrate an elevated response to enzalutamide treatment (Swami et al., 2022).

5.1.4 SPOP Immune microenvironment and anti-PD-L1 immunotherapy

Outside of its role in controlling DNA damage repair and the AR signalling cascade, SPOP has also been shown to be involved in regulating the prostate immune microenvironment; with *SPOP* mutations being identified to negatively impact immune surveillance of the prostate

niche which ultimately drives evasion of anti-tumour immunity. Firstly, SPOP has been shown to directly interact with the immune checkpoint protein PD-L1 and target it for degradation. As such, *SPOP* mutations lead to accumulation of PD-L1, diminishing $CD3^+$ tumour-infiltrating T lymphocyte levels and cytotoxic T-cell-mediated tumour cell clearance and immune evasion in PC mouse xenograft models and human PC primary samples (Zhang et al., 2018a). As a result, *SPOP* mutant PC leads to a disordered immune microenvironment and ultimate accumulation of PD-L1, which may potentiate susceptibility to immune checkpoint blockade therapy through inhibition of PD-L1. This interaction has been demonstrated in response to CDK4/6 inhibition which results in accumulation of PD-L1 which, upon anti-PD-L1 immunotherapy treatment in xenograft models, increased tumour regression and improved overall survival (Zhang et al., 2023). Further information into the mechanistic nature of this increased PARP inhibitor sensitivity within SPOP mutant prostate cancer has been provided through RNAseq analysis of SPOP mutant PC patient data alongside overexpression PC associated SPOP mutants within PC cell lines. Here, evidence suggests SPOP mutations destabilising STING1 leading to increased non-canonical STING-NF-KB signalling ultimately bringing about immunosuppressive effects. Upon PARP inhibition, there is a shift to canonical cGAS-STING signalling which has antitumour effects (Geng et al., 2023).

5.1.5 SPOP and DNA hypermethylation

Another key characteristic of SMPC is DNA hypermethylation, due to the stabilisation of the SPOP substrate GLP, responsible for the recruitment of DNA methyltransferase to chromatin. As such, through the use of PC xenografts and PDX models harbouring SPOP mutations, evidence displayed sensitivities to single agent treatment with the DNA methylation inhibitor 5-AzaC alongside in combination with docetaxel (Zhang et al., 2021).

Also of note, SPOP itself has been shown to susceptible to small-molecule inhibition. With such a well characterised SPOP-binding motif, computational screens were performed to return designs of small molecule inhibitors that were able to disrupt SPOP activity and block PTEM and DUSP7 substrate degradation in clear cell renal carcinoma (ccRCC) leading to repression of the Akt/ERK signalling and ultimately suppression of proliferation (Guo et al., 2016). With *SPOP* mutations showing sensitivities to therapeutics such as PARPi, new synthetic lethality

relationships with current PC therapeutics may be revealed in SPOP-WT cancers in response to dual treatment with an SPOP inhibitor.

5.1.6 DNA-PKcs inhibition and current clinical candidates

With SMPC having increased dependency on NHEJ for the repair of DSBs, potentially through targeting of this pathway via inhibition of a key NHEJ protein, DNA-PKcs, new sensitivities in-line with those previously described may be discovered. There have been a number of compounds developed targeting DNA-PKcs, namely NU7441, NU5455, AZD7648 and peposertib (amongst others) which have seen varying success as potential clinical candidate compounds.

NU7441 is a small molecule inhibitor found to be highly selective for DNA-PK with an IC50 of 14 nM (Leahy et al., 2004). It has seen use in a range of studies across multiple cancer backgrounds all looking to enhance the current effectiveness of either radio- or chemotherapy via blocking DNA damage repair ability (Ciszewski et al., 2014, Yang et al., 2016). NU5455 is a more selective DNA-PKcs inhibitor with less off target inhibitory effects on other PIKK family members (Willoughby et al., 2020). Here, an IC50 of 8.2 ± 2 nM was described with a 228-fold increase in selectivity for DNA-PKcs over NU7441. Unfortunately, both have not been taken forward for clinical use. The AstraZenica compound was discovered through a compound library screen and boasts an IC50 of 0.6 nM. Further, Fok et al., 2019 describes decreased A549 cell viability in clonogenic assay experiments when combining AZD7648 and ionising radiation exposure (Fok et al., 2019). AZD has been used in phase I clinical trials in combination with the chemotherapeutic Doxorubicin, as well as a monotherapy (NCT03907969). M3814, now known as peposertib, has since been developed by Merck and is the current leading clinical candidate for DNA-PKcs inhibition with an IC50 of <3 nM. There has been evidence of potentiation of radiotherapy and tumour regression in mouse models and has shown great promise in phase I clinical trials in human advanced solid tumours (Zenke et al., 2020, van Bussel et al., 2021).

5.2 Aims

Given the role of SPOP in HR-mediated DNA repair, and the findings that loss of function *SPOP* mutations result in a *BRCA*-like cellular phenotype, it is likely that PC-associated *SPOP* gene aberrations represent a therapeutic vulnerability. Consistent with this, elevated sensitivities to PARP inhibition in loss of function SPOP PC models *in vitro* has been reported. However aside from this interesting demonstration of synthetic-lethality between SPOP and PARP, there remains major knowledge gaps in whether *SPOP* mutation affords other therapeutic vulnerabilities in relation to DNA damage repair. As evidenced by the higher rate of genetic abnormalities in *SPOP* mutant PC patient samples, loss of SPOP function drives an over-reliance on the error-prone NHEJ DNA repair pathway. As such, the hypothesis to test in this study is that key proteins within the NHEJ repair pathway are attractive potential therapeutic targets. More specifically, consistent with its role in NHEJ, and the availability of potent inhibitors, DNA-PKcs blockade in *SPOP* mutant-like cellular backgrounds will be tested to assess if cells lacking SPOP function are more sensitive to DNA-PKIs, as single agent treatment and in combination with IR-induced DNA damage, as a consequence of failure to effectively repair DNA damage.

Applying SPOP knockdown and knockout models utilised in Chapter 4, the effects of DNA-PKcs inhibition within SPOP mutant-like PC cell lines will be investigated by:

1. Assessing cellular growth and viability of SPOP depleted and CRISPR/Cas9 mediated *SPOP* knock-out PC cell lines upon single agent DNA-PKI treatment and in combination with IR exposure.

2. Providing mechanistic insight into the effects of single-agent and combined DNA-PKcs inhibition/IR treatments in wild-type and SPOP-defective models.

Through completion of these aims, it is hoped that evidence of enhanced sensitivity to DNA-PKIs in *SPOP* mutant backgrounds is generated providing a platform for further work to validate these findings in more complex and patient representative models.

5.3 Methods

5.3.1 Compounds

List of compounds used in cell proliferation assays can be found in Table 5.1. Cells were treated with the compound of choice or representative vehicle control (DMSO) and returned to incubate. For experiments requiring IR exposure, cells were treated as described below following 1-hour incubation post-drug/vehicle treatment.

Table 5.1 – *Compounds used for DNA-PKcsi*

Compound	**Supplier**	**Stock concentration**	**Treatment concentration**	**Storage**
NU7441	Newcastle University Drug Discovery	5 mM in DMSO	2 nM - 30µM	-20/-80 °C Short/long term
NU5455	Newcastle University Drug Discovery	5 mM in DMSO	2 nM - 30µM	-20/-80 °C Short/long term
AZD7648	Sellekchem	5 mM in DMSO	2 nM - 30µM	-20/-80 °C Short/long term

5.3.2 Ionising radiation treatment

Cells were transferred from incubation following drug/vehicle control pre-treatment (above) and doses of ionising radiation (0.5–8 Gray (Gy)) were administered using R5320 irradiation system (Gulmay Medical) set at 320 kV and 10 mA. Resultant Gy treatment dose was calculated

dependant on both depth of solution within cell culture vessel with lid and time of exposure. Following IR treatment cells were returned to incubate for designated lengths of time as described below.

5.3.3 Cell proliferation assays

Cell counts

Following prescribed DNA-PKI/IR treatments, cells were returned to incubate for 120 hours after which time they were gently washed with PBS, then incubated with 1 x Trypsin-EDTA to detach cells. Trypsin-EDTA was neutralised in full media and cells were transferred to sterile Eppendorf tubes and centrifuged at 400 x g for 5 minutes after which the supernatant was removed, and cells were resuspended in 1 mL full media. 10 µL of this suspension was used for the cell count using a haemocytometer. Total cells within a 4x4 square were counted twice for each treatment arm and an average taken as the final result. Total averages for all repeats were then normalised against the control arm.

Clonogenic assay

Clonogenic assays were performed as described in Frankern *et al*., (2006). In brief, 500 cells were seeded into 6-well cell culture plates and once adhered, were treated as necessary before being returned to incubate for 4 weeks, with media changes performed each week. At the end of week 4, culture media was removed and cells were gently washed with 1 mL PBS twice before fixation with 4% PFA in PBS (v/v) for 30 minutes. Once fixed, cells were again washed with 1 mL PBS twice before colonies were stained with crystal violet (contents) for 30 minutes and subsequently washed with tap water and left to air dry. Resultant colonies were counted for each treatment and normalised to the plating efficiency (PE) calculated from the control (untreated) wells using the equation:

$$PE = \frac{no.colonies\ formed}{no.colonies\ plated} \times 100$$

Using the calculated plating efficiency for each cell line, the surviving fraction (SF) of each treatment was calculated using the equation:

$$SF = \frac{no.colonies\ formed\ after\ treatment}{no.\ colonies\ seeded\ \times\ PE}$$

5.3.4 Sulforhodamine B (SRB) growth assay

1000 cells were seeded into the inner 60 wells of a 96-well cell culture plate with 100 µL of PBS added to the peripheral wells to minimise edge effect. Upon adhesion, cells were treated as required in a total volume per well of 100 mL, and returned to incubate for 120 hours. Following this incubation, cells were fixed by direct addition of ice cold 50% trichloroacetic acid (TCA) (w/v) (to a 10% (v/v) final concentration) at 4°C for at least 1 hour. Once fixed, cells were washed with tap water and allowed to air dry before addition of 100 mL of 4% SRB dissolved in 1% glacial acetic acid (v/v) to stain for 30 minutes. Plates were then washed with 1% glacial acetic acid (v/v) to remove unbound dye and left to air dry before solubilisation of the stain via addition of 100 uL 10 mM Tris pH 10.8 for 15 minutes with intermittent agitation. Resultant absorbance was measured at 570 nm using a microplate reader (Bio-Rad).

5.3.5 Immunofluorescence and γH2AX quantification

Immunofluorescence (IF) was used to indirectly visualise target proteins within fixed cell samples. Cells were seeded to an appropriate density onto sterile 20 mm x 20 mm glass cover slips (ThermoFisher) and 24 hours were allowed for cell adherence. After this, cells were transfected with either scrambled (siScr) or SPOP siRNA (siSPOP; see *siRNA Transfection, Chapter 3.2.1*) and allowed to incubate for another 48 hours before control or drug treatment and incubated for a further 1 hour. Cells were then exposed to 2 Gray (Gy) IR and returned to incubate for 1-24 hours. Culture medium was then removed, and cells were gently washed twice using PBS for 5 minutes each. 1 mL of 4% paraformaldehyde in PBS was used to fix cells for 20 minutes at room temperature before washing twice in PBS for 5 minutes each. 1 mL of 0.1% triton X-100 (Merck) in PBS was added to permeabilise cells for 10 minutes. Permeabilization reagent was removed, and samples were washed twice in PBS for around 5

minutes per wash. A further 1 mL of 4% bovine serum albumin (BSA) (Merck) in PBS was added to block cells for 30 minutes at room temperature. Samples were then incubated at 4°C overnight with primary antibody (Table 3.7) diluted in 4% BSA in PBS. Secondary antibody only samples were incubated overnight in 4% BSA in PBS only. Samples were washed twice in PBS for around 5 minutes per wash prior to incubation for 1 hour at room temperature in the dark with fluorochrome-conjugated secondary antibody, followed by two 5-minute washes in PBS. Samples were mounted onto slides using a drop of mounting media with DAPI (Abcam). Slides were imaged using the Leica DM6 B widefield microscope using SPOT Imaging to capture images. Representative images of DAPI and γH2AX staining generated from this protocol and used for quantification can be seen in

Images (as represented in Figure 5.1) were then analysed using the ImageJ software (available from: https://imagej.nih.gov/ij/). Nuclear regions of cells were selected based on DAPI staining and taken as total cell number image. Subsequent γH2AX foci quantification was then conducted through maxima counts with the set threshold (>15) remaining constant across all images analysed. For each image total maxima from γH2AX staining were divided by the cell number quantified from DAPI nuclear staining, giving the foci/cell count. This performed in for two images per experimental arm, with the average being taken to represent a single biological replicate. Foci/cells were then normalised to internal control arms giving relative foci/cell for each experimental arm.

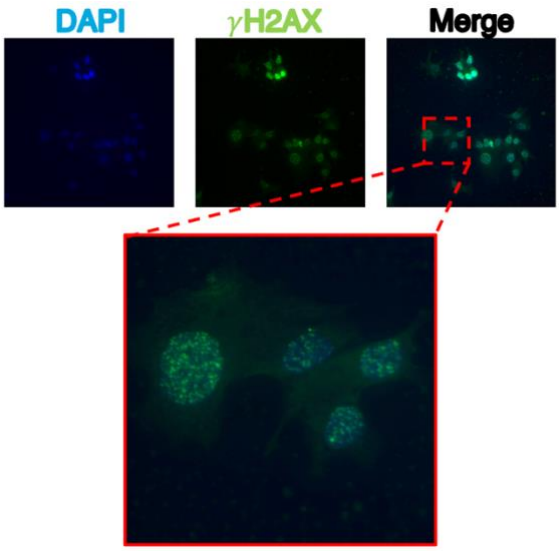


Figure 5.1- Representative immunofluorescence microscopy image for Phospho-Ser139 γH2AX staining used for γH2AX foci/cell quantification.

5.3.6 Cell cycle analysis

Cells were seeded into 6-well cell culture plates, treated as necessary then returned to incubate. For harvest at required time-points, cells were washed and trypsinised then transferred to a falcon tube where they were then centrifuged at 400 x g for 5 minutes. The resultant pellets were then washed again in PBS by resuspension and centrifugation at 400 x g for 5 minutes before the pellet was resuspended in 100 mL citrate buffer (250 mM sucrose, 40 mM sodium citrate, pH7.6). Following this, 400 mL of propidium iodide (PI) stain in buffer (20 µg/mL (PI), 0.5 mM EDTA, 0.5% NP40, 10 µg/µL RNase A) was added to each sample excluding a non-stained control sample to which 400 µL of PBS was added in place of the PI stain. Samples were left to incubate covered at $4^{O}C$ for 1 hour, before being mixed and loaded onto an Attune™ NxT flow cytometer (Invitrogen) for data acquisition from a minimum of 10,000 events. Gating for single cell populations to remove cell debris and aggregates was performed using the non-stained and stained control samples, with final data analysis was conducted using FCS Express (DeNovo Software).

5.4 Results

5.4.1 siRNA-mediated SPOP depletion renders cells more sensitive to a combination of DNA-PKcs inhibition and ionising radiation.

To investigate the anti-proliferative effects of both single agent NU7441 treatment or 2 Gy IR exposure and the effects of a combination of the two, cell count experiments were conducted in the CWR22Rv1-AR-EK cell line [Figure 5.2A]. This cell line was chosen initially to test the proposed hypothesis as it represents an aggressive, late-stage treatment resistant disease, and as such positive results here may indicate the potential of a new therapeutic option for patients where current available treatments are limited. In this experimental set-up, using previous in-house data calculating the IC50 dose for NU7441, a dose of 2 nM NU7441 was used, markedly below the 2 µM IC50 as would allow for detection of more subtle differences in response to the wide range of treatments present within each experimental arm. Here, the effects of SPOP depletion alone on cell proliferation can be seen, no significant reduction in cell number after the 120-hour period. In the siScr cells, a 25.48% and 49.34% reduction was observed in response to single agent NU7441 and IR treatment, respectively. Depletion of SPOP further increased these reductions, with a significant further reduction to 48.34% in response to NU7441 treatment and non-significant 60.73% reduction in response to IR exposure compared to the single agent treatments in the siScr experimental arms, suggesting that loss of SPOP can sensitise cells to these single agent DNA-PKI and IR treatments. As expected, combining NU7441 and IR treatment in the siScr control cells resulted in a 64.24% reduction in cell number. Crucially, in cells depleted of SPOP, this reduction was significantly increased to 85.45% indicating cells lacking SPOP function are considerably more sensitive to dual DNA-PKcs blockade and IR treatment.

These initial experiments provided evidence of a synthetic-lethality relationship between SPOP and DNA-PKIs which is exacerbated by introducing exogenous IR-induced DNA damage. To study this further, colony forming (clonogenic) assays were next undertaken in the CWR22Rv1-AR-EK cell-line using the same single-agent and dual NU7441 and IR treatments as above. Resultant surviving fractions were quantified and compared to the untreated control experimental arm [Figure 5.2B]. Here the impact of 2 Gy IR can be seen when looking at the control arm with a significant 57% reduction in surviving fraction seen in response to the IR

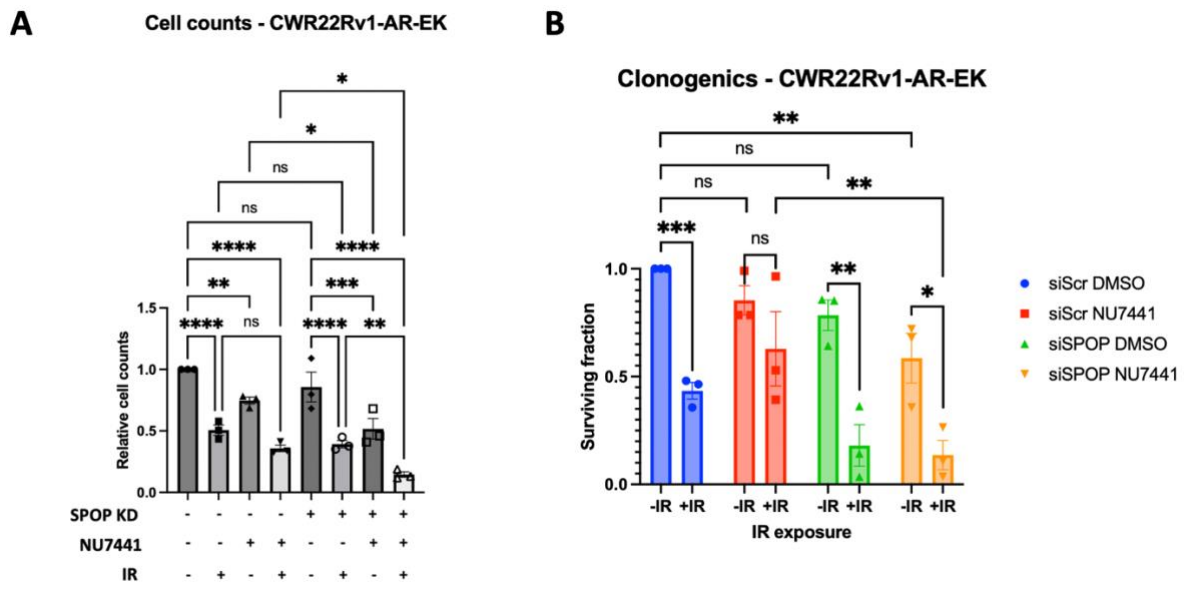


Figure 5.2 – Combination treatment of DNA-PKi and IR cause reduced growth in SPOP depleted CWR22Rv1-AR-EK cells. A. *Cell count experiments comparing cell counts from CWR22Rv1-AR-EK cells transfected with siScr/siSPOP siRNA for 48 hours, 1-hour pre-treatment with 2 nM NU7441 or vehicle control pretreatment before 2 Gy IR exposure, with final cell counts conducted 120 hours post IR exposure. Data is normalised relative to the untreated control and represents 3 individual experiments ± SEM. Statistical significance was determined by a One-way ANOVA with Dunnett's multiple comparison test (*, ** and *** denote P-values <0.05, <0.01 and <0.001, respectively, ns denotes not significant).* ***B.*** *Clonogenic assay of cells treated as in A, and left to incubate for 3 weeks before endpoint processing. Data is normalised relative to the untreated control and represents 3 individual experiments ± SEM. Statistical significance was determined by a One-way ANOVA with Dunnett's multiple comparison test (*, **, *** and **** denote P-values <0.0332, <0.0021, <0.0002 and <0.0001 respectively, ns denotes not significant).*

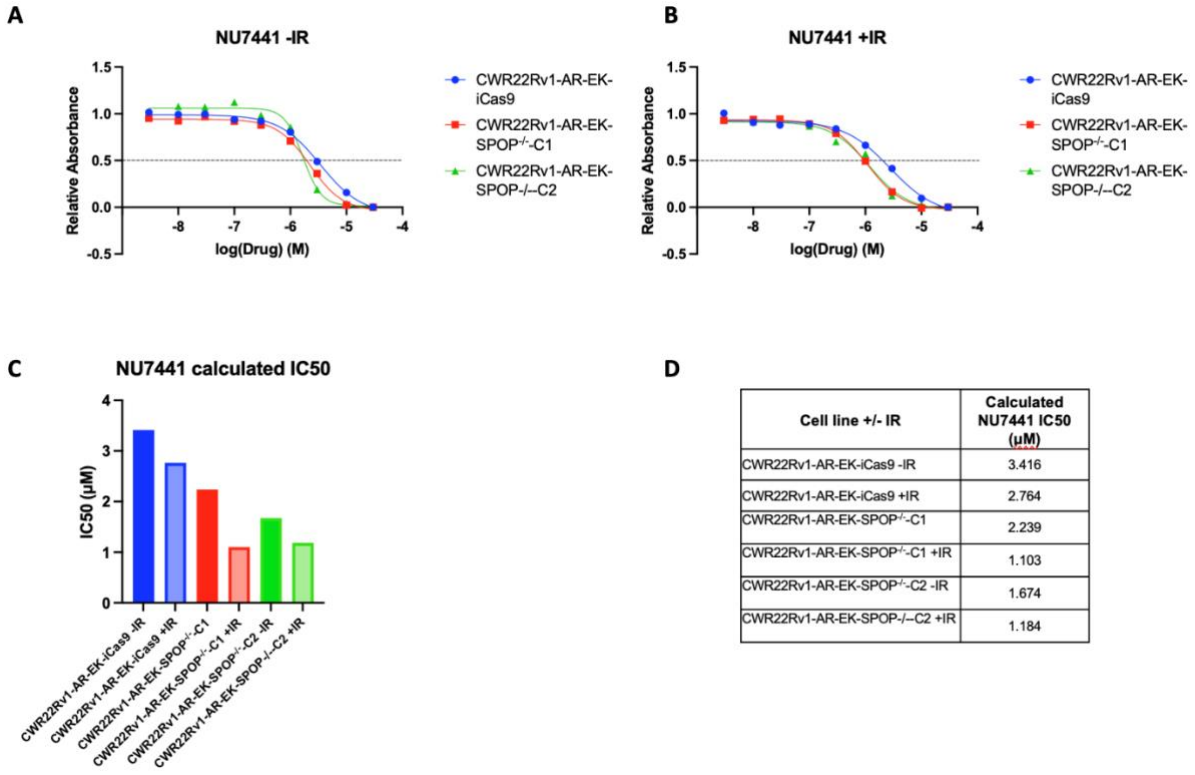
treatment. Upon addition of 2 nM NU7441, no significant decrease in the survival fraction was observed, with survival fraction decreasing by 37% with the addition of IR exposure. Furthermore, SPOP depletion alone resulted in a non-significant 22% reduction in survival fraction, with this dropping substantially by 82% following IR exposure. IR exposure significantly further reduced surviving fraction in SPOP depleted cells treated with NU7441 to 14%. Combination of SPOP depletion and NU7441 treatment showed a non-significant reduction in surviving fraction in the absence of IR, from 85% in the siScr experimental arm to 59% in the siSPOP arm.

Of note, the evidence of the synthetic-lethality relationship between SPOP depleted cells and NU7441 treatment can be seen through comparisons of the siScr NU7441 and siSPOP NU7441 experimental arms in the presence of IR. Here, the combination of NU7441 +IR resulted in a significant reduction in surviving fraction from 63% in the siScr experimental arm to 13% in response to the same treatment in SPOP depleted cells. Although there was a reduction in surviving fraction seen as a result of SPOP knockdown with IR exposure and NU7441 treatment

compared to the siScr cells, this was not significant. Ultimately this cell count and clonogenic data indicate the identification of a potential synthetic lethality caused by SPOP depletion within the CWR22Rv1-AR-EK cell line through treatment with NU7441 which is further enhanced with IR exposure, whilst supporting published data describing SPOP depletion alone resulting in heightened radiosensitivity.

5.4.2 SPOP knock-out CWR22Rv1-AR-EK models show evidence of synthetic lethality relationship in response to DNA-PKcs inhibition which is further potentiated by IR treatment

To further investigate the sensitivity of SPOP-depleted cells, via siRNA knockdown, to DNA-PKcs inhibition, different DNAPKIs were tested in the CWR22Rv1-AR-EK-SPOP$^{-/-}$ model (Section 4.4.5). To do this, SRB assays were set-up using a wide dose range of different DNAPKcs inhibitors with the aim to compare IC_{50} values between the parental CWR22Rv1-AR-EK-iCas9 and two clonal derivative CWR22Rv1-AR-EK-SPOP$^{-/-}$ cell lines both in the presence and absence of 2 Gy IR. The first inhibitor tested in these models was NU7441 as previous data had already displayed a sensitivity to a combination treatment in response to siRNA-mediated SPOP depletion [Figure 5.3]. In the absence of ionising radiation, the regression fits did not return a clear shift in IC_{50} between wild-type cells and CWR22Rv1-AR-EK-SPOP$^{-/-}$ clones 1 (C1) and 2 (C2), although the values returned from the calculations showed a subtle shift in IC_{50} values when comparing the SPOP$^{-/-}$ models (IC_{50} values of 2.239 µM and 1.674 µM for SPOP$^{-/-}$-C1 and SPOP$^{-/-}$-C2 clones, respectively) to the parental control (IC_{50} = 3.416 µM)(Figure 5.3A). Combining NU7441 and 2 Gy IR treatment resulted in a clear alteration in IC_{50} values in both



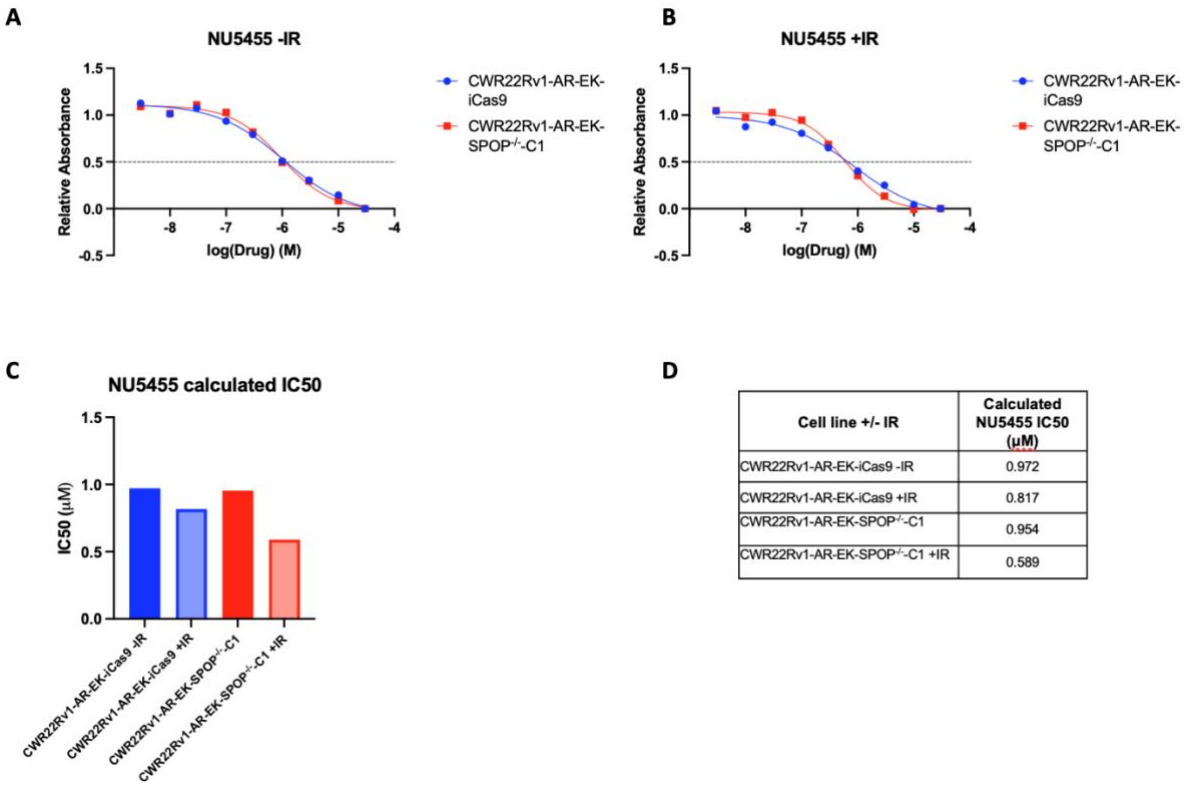
Cell line +/- IR	Calculated NU7441 IC50 (μM)
CWR22Rv1-AR-EK-iCas9 -IR	3.416
CWR22Rv1-AR-EK-iCas9 +IR	2.764
CWR22Rv1-AR-EK-SPOP$^{-/-}$-C1	2.239
CWR22Rv1-AR-EK-SPOP$^{-/-}$-C1 +IR	1.103
CWR22Rv1-AR-EK-SPOP$^{-/-}$-C2 -IR	1.674
CWR22Rv1-AR-EK-SPOP-/--C2 +IR	1.184

Figure 5.3 - CWR22Rv1-AR-EK-SPOP-/--C1 and CWR22Rv1-AR-EK-SPOP-/--C2 cells show sensitivity to NU7441 treatment which is heightened upon IR exposure compared to CWR22Rv1-AR-EK-iCas9 cells. *Response to dose range of NU7441 in the absence (**A**) and presence (**B**) of 2 Gy IR. Resultant IC50 calculated for NU7441 in each of the cell lines +/- IR represented in graphical format (**C**) as well as actual calculated values ttabularised (**D**). Data is normalised relative to each treatment arms vehicle control and is represents 3 individual experiments.*

SPOP knock-out clones when compared to the control; with outputs from the regression fits demonstrating clear separation in drug response with calculated IC_{50} of 1.103 µM and 1.184 µM for SPOP$^{-/-}$-C1 and SPOP$^{-/-}$-C2 compared to

2.764 µM for wild-type cells [Figure 5.3B-D].

The next DNAPKcs inhibitor tested was NU5455 a derivative of NU771 (Section 5.1.6)[Figure 5.4]. In contrast to findings with NU7441, single agent NU5455 treatment IC_{50} values demonstrated minimal difference between control and SPOP$^{-/-}$-C1 cell line derivatives; with IC_{50} values of 0.9719 µM and 0.954 µM respectively. Combining NU5455 and IR treatment in wild-type cells resulted in a slight reduction in IC_{50} to 0.817 µM, while in the *SPOP*$^{-/-}$-C1 clone, IR exposure decreased the NU5455 IC_{50} to 0.589 µM [Figure 5.4B-D]. Taken together these results confirm the previously demonstrated increased anti-proliferative effects of the combination of IR and DNAPKcs inhibition in cells compromised for SPOP function. It has also



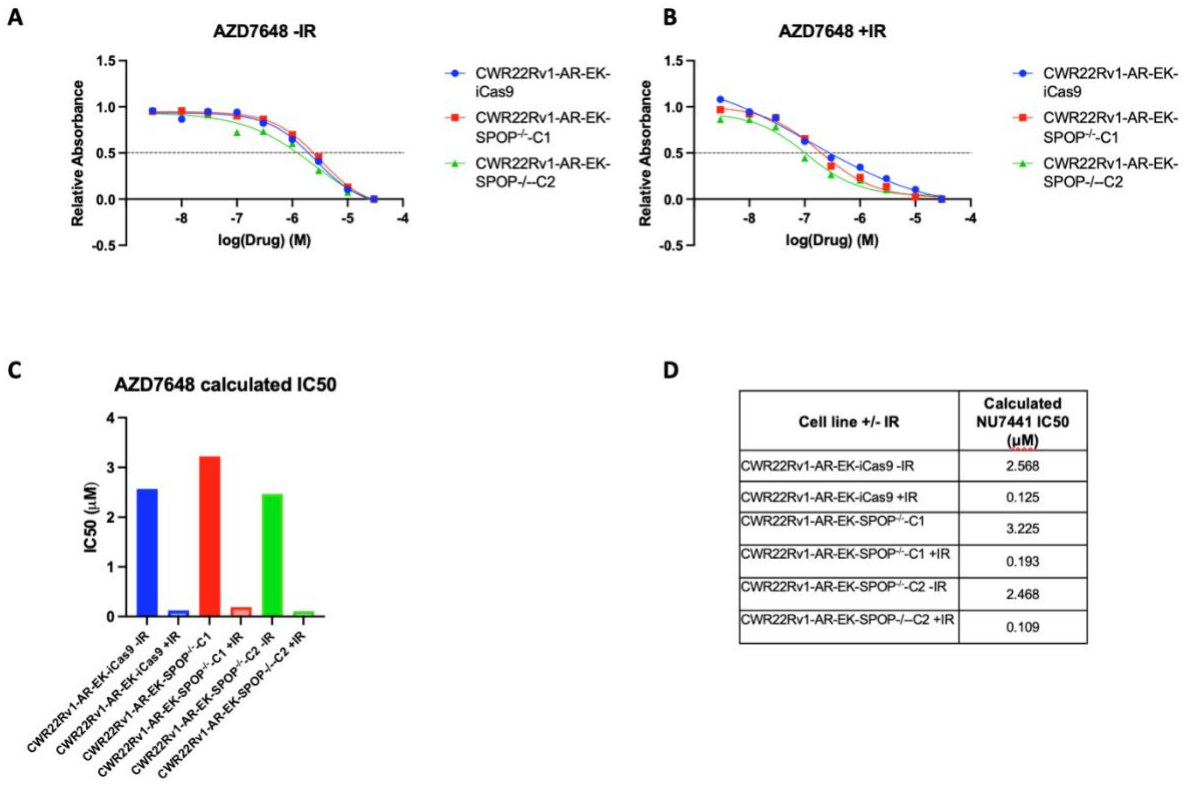
Cell line +/- IR	Calculated NU5455 IC50 (µM)
CWR22Rv1-AR-EK-iCas9 -IR	0.972
CWR22Rv1-AR-EK-iCas9 +IR	0.817
CWR22Rv1-AR-EK-SPOP$^{-/-}$-C1	0.954
CWR22Rv1-AR-EK-SPOP$^{-/-}$-C1 +IR	0.589

Figure 5.4 - CWR22Rv1-AR-EK-SPOP-/--C1 and CWR22Rv1-AR-EK-SPOP-/--C2 cells show sensitivity to NU5455 treatment which is heightened upon IR exposure compared to CWR22Rv1-AR-EK-iCas9 cells. *Response to dose range of NU5455 in the absence (**A**) and presence (**B**) of 2 Gy IR. Resultant IC50 calculated for NU5455 in each of the cell lines +/- IR represented in graphical format (**C**) as well as actual calculated values tabularised (**D**). Data is normalised relative to each treatment arms vehicle control and is represents 3 individual experiments.*

highlighted the increased sensitivity of SPOP depleted cells to single agent NU7441 treatment, but not to NU5455 suggesting subtleties in drug action may give rise to these differences.

Finally, the more clinically-relevant DNAPKcs inhibitor AZD7648 was assessed, the results of which contrasted with those observed for the previous compounds tested [Figure 5.5]. In the absence of IR exposure, wild-type cells demonstrated an IC_{50} of 2.568 µM which increased to 3.225 µM in the SPOP$^{-/-}$-C1 and decreased to 2.468 µM in the SPOP$^{-/-}$-C2. In combination with 2 Gy IR exposure, the IC_{50} values for all 3 cell lines tested decreased drastically to 0.125 µM in the parental line and to 0.193 µM and 0.109 µM in SPOP$^{-/-}$-C1 and SPOP$^{-/-}$-C2 models respectively.

In keeping with our SPOP knockdown experiments, further clonogenic assays were conducted utilising the CWR22Rv1-AR-EK-SPOP$^{-/-}$ model to provide further support for a synthetic vulnerability between SPOP and DNA-PKcs blockade [Figure 5.6]. Here, the original 2 nM dose



Cell line +/- IR	Calculated NU7441 IC50 (µM)
CWR22Rv1-AR-EK-iCas9 -IR	2.568
CWR22Rv1-AR-EK-iCas9 +IR	0.125
CWR22Rv1-AR-EK-SPOP$^{-/-}$-C1	3.225
CWR22Rv1-AR-EK-SPOP$^{-/-}$-C1 +IR	0.193
CWR22Rv1-AR-EK-SPOP$^{-/-}$-C2 -IR	2.468
CWR22Rv1-AR-EK-SPOP-/--C2 +IR	0.109

Figure 5.5 – IC50 calculations for AZD7648 treatment in CWR22Rv1-AR-EK-iCas9, CWR22Rv1-AR-EK-SPOP-/--C1 and CWR22Rv1-AR-EK-SPOP-/--C2 cells. *Response to dose range of AZD7648 in the absence (**A**) and presence (**B**) of 2 Gy IR. Resultant IC50 calculated for AZD7648 in each of the cell lines +/- IR represented in graphical format (**C**) as well as actual calculated values tabularised (**D**). Data is normalised relative to each treatment arms vehicle control and is represents 3 individual experiments.*

was used (as in the original cell count experiments) alongside a 1 µM dose which was more representative of the IC50 calculated from the SPOP$^{-/-}$ clones following IR treatment. First, the plating efficiency was determined using the untreated controls for each cell line tested as the effects of complete loss of SPOP on growth rates (Section 4.4.5) was known, although effects on cell adherence and viability was unknown [Figure 5.6C]. The plating efficiency for the parental CWR22Rv1-AR-EK-iCas9 was 70.3%, which was greater than the 44% plating efficiency calculated for the SPOP$^{-/-}$-C1 derivative. Due to this disparity, the calculated survival fractions and resultant normalisation for wild-type and SPOP$^{-/-}$ derivatives has been conducted by comparing to the untreated control for that respective cell line.

In the absence of NU7441 treatment, the effects of the 2 Gy IR exposure on the colony forming potential of each line were substantial [Figure 5.6D]; with IR exposure resulting in a statistically significant reductions in the surviving fraction of the parental control line and a similar significant reduction in the SPOP$^{-/-}$-C1 clone. Clear differences in surviving fraction can be seen following the addition of 2 nM NU7441 [Figure 5.6E]. NU7441 treatment caused the surviving

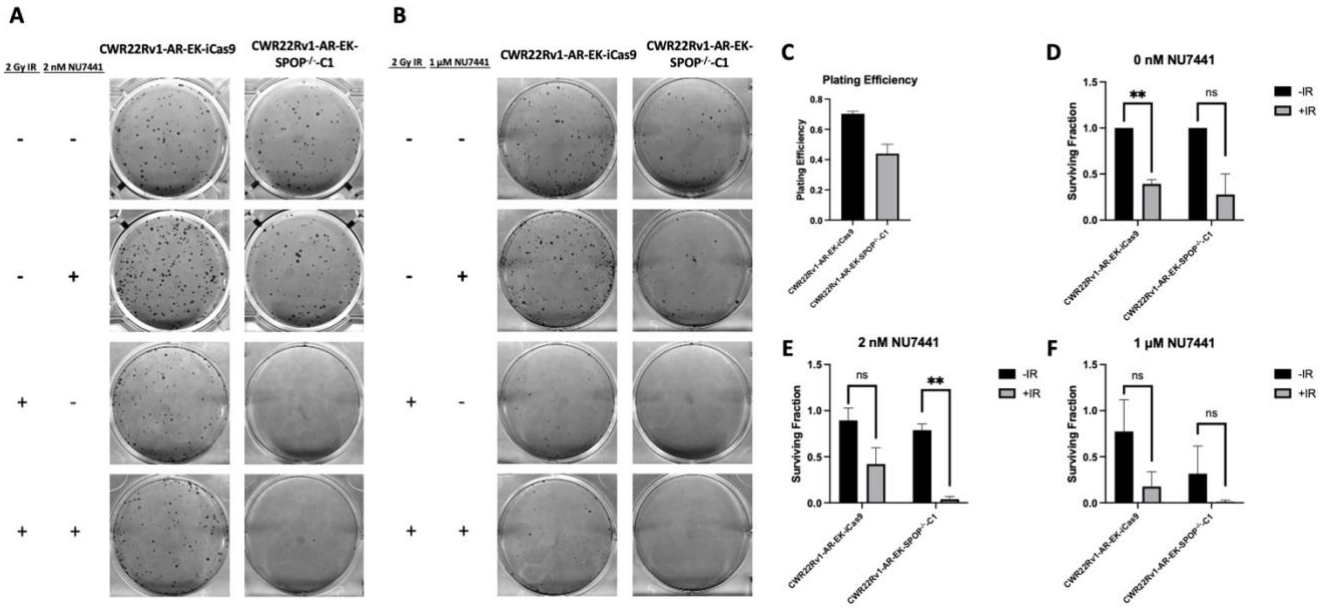


Figure 5.6 – NU7441 treatment shows reduced survival fraction in CWR22Rv1-AR-EK-SPOP-/--C1 cells compared to CWR22Rv1-AR-EK-iCas9 in the presence and absence of IR. A&B. *Representative endpoint images from clonogenic assays in CWR22Rv1-AR-EK-iCas9 cells and CWR22Rv1-AR-EK-SPOP-/--C1 cells. Data is representative 3 individual experiments.* ***C.*** *Plating efficiency calculated from the CWR22Rv1-AR-EK-iCas9 and CWR22Rv1-AR-EK-SPOP-/--C1 clones used to normalise downstream surviving fraction calculations. Data represents 3 individual experiments ± SEM.* ***D-F.*** *Survival fractions of CWR22Rv1-AR-EK-iCas9 and CWR22Rv1-AR-EK-SPOP-/--C1 clones in the presence and absence of IR at vehicle control (**D**), 2 nM (**E**) or 1 μM (**F**) concentrations of NU7441. Data represents 3 individual experiments ± SEM. Statistical significance was determined by T-tests (*, **, *** and **** denote P-values <0.0332, <0.0021, <0.0002 and <0.0001 respectively, ns denotes not significant).*

fraction of the parental line to decrease to 89%, with a further reduction to 79% seen in the SPOP$^{-/-}$-C1. The surviving fraction of the parental line reduced to 42% upon IR exposure in combination with 2 nM NU7441, however a much greater significant decrease to 0.4% was seen in the SPOP$^{-/-}$-C1 clone. A similar trend was observed when the NU7441 dose was increased to 1 µM [Figure 5.6F]. Here, clear effects of the single agent drug treatment on the surviving fraction can be seen with the parental cells seeing a surviving fraction of 77% and a significant reduction to 31% in the SPOP$^{-/-}$-C1 cells. These reductions are further heightened upon IR exposure with surviving fractions of 18% and 0.1% returned from the control cells and SPOP$^{-/-}$-C1 cells respectively.

Having observed a clear single agent 2 Gy IR exposure resulting in reduced survival of cells, final clonogenic assays were set up with an expanded range of IR doses ranging from 0.5 Gy to 8 Gy [Figure 5.7]. The response for this IR dose range was also tested in the background of 0.55 µM and 1.15 µM NU7441, representing IC_{25} and the IC_{50} values calculated previously, alongside an equivalent DMSO control. As before, plating efficiencies were calculated from the control

experimental arms for each line tested and used to calculate surviving fractions. In the IR alone arm, a clear reduction in survival fraction was observed from both the SPOP$^{-/-}$-C1 and -C2 clones across all IR doses tested when compared to the control line [Figure 5.7A]. In all experimental arms tested, 8 Gy IR resulted in a complete loss in cell viability. Reductions in surviving fraction were seen in both knock-out clones tested as low as 0.5 Gy IR exposure when compared to the parental control cells, with a similar pattern observed across all IR doses up to 4 Gy; where no colonies were formed apart from a small number in the control arm. A similar general trend was seen in the cells treated with 0.55 µM NU7441, in that a reduction in survival fraction was seen from both SPOP$^{-/-}$-C1 and -C2 clones when compared to the control [Figure 5.7B]. With increases in IR dosage, this general trend continued up to 2 Gy where there were few colonies formed in the control cells and no colonies formed in the SPOP knock-out clones. When NU7441 dose was increased to the IC50 value of 1.15 µM, there was a much greater effect on the surviving fraction in the SPOP knock-out clones [Figure 5.7C].

In the absence of IR and 0.55 µM NU7441 treatment, parental control cells returned a surviving fraction of 104%, significantly higher than the 76% and 72% returned for the SPOP$^{-/-}$-C1 and SPOP$^{-/-}$-C2, respectively [Figure 5.7D]. Furthermore, parental control cells demonstrated a 83% surviving fraction in response to the 1.15 µM NU7441 dose, significantly higher than the 29% and 34% observed for the SPOP$^{-/-}$-C1 and -C2 clones, respectively. At the lowest IR dose tested of 0.5 Gy, no significant changes in surviving fraction were seen in the absence of NU7441 [Figure 5.7E]. Upon treatment with 0.55 µM NU7441, control cells displayed a surviving fraction of 62%, while the SPOP$^{-/-}$-C1 and -C2 clones showed a significantly reduced surviving fraction of 29% and 27%, respectively. In response to the higher NU7441 dose of 1.15 mM, the difference between parental control and the SPOP knockout clones was extended further, whereby the parental control cells returned a surviving fraction post treatment of 54%, with a significantly reduced surviving fraction of 10% for SPOP$^{-/-}$-C2 and a further reduced surviving fraction of 0.05 for the SPOP$^{-/-}$-C1 clone [Figure 5.7E]. Overall, this data provides evidence that combining NU7441-mediated DNA-PKcs blockade with varying intensities of IR dose is a more cytotoxic treatment in cells representing *SPOP* mutation.

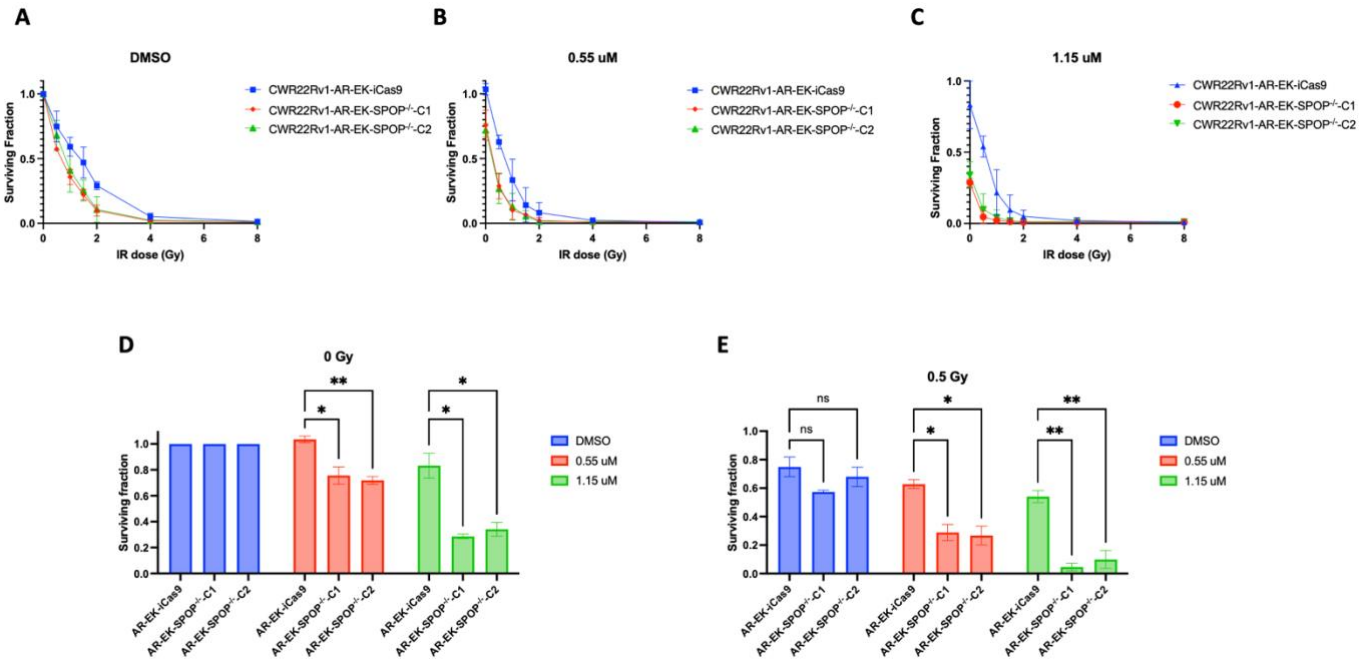


Figure 5.7 - CWR22Rv1-AR-EK-SPOP-/--C1 and CWR22Rv1-AR-EK-SPOP-/--C2 cells show increased radiosensitivity compared to CWR22Rv1-AR-EK-iCas9 cells, heightened by NU7441 treatment. A-C. *Survival fractions from clonogenic assays of CWR22Rv1-AR-EK-iCas9, CWR22Rv1-AR-EK-SPOP-/--C1 and CWR22Rv1-AR-EK-SPOP-/--C2 cells in response to a 0-8 Gy dose range of IR exposure. Data is normalised respective to each cell lines plating efficiency and is representative of 3 individual repeats ± SEM. Graphical representation of surviving fractions at 0 (**D**) and 0.5 (**E**) Gy IR dose. Data is normalised respective to each cell lines plating efficiency and is representative of 3 individual repeats ± SEM. Statistical significance was determined by a One-way ANOVA with Dunnett's multiple comparison test (*, **, *** and **** denote P-values <0.0332, <0.0021, <0.0002 and <0.0001 respectively, ns denotes not significant).*

5.4.3 Combination treatment with Doxorubicin and NU7441 does not show a synergistic effect in the presence of SPOP knock-down.

Given the treatment of NU7441 was more efficacious in cells subject to IR-induced DNA damage in both wild-type and SPOP defective cells, it was important see if these effects could be recapitulated using the DNA damaging chemotherapeutic agent doxorubicin. Utilising the same 2 nM concentration of NU7441 as used previously, SRB assays were set-up to test the effects of a range of doxorubicin doses on CWR22Rv1-AR-EK cells transfected with either siScr or siSPOP siRNAs to provide an appropriate treatment dose for this cell-line [Figure 5.8A]. As shown in Figure 5.8A, a marked reduction in cell number was observed between 0 - 100 nM in both experimental cell line arms; with greater doses not further impacting cell growth. Ultimately, this indicated that an appropriate doxorubicin dose range for downstream experiments in the CWR22Rv1-AR-EK cell line was between 0 and 100 nM. Using this identified treatment range subsequent SRB assays were conducted utilising 0 - 100 nM doxorubicin and

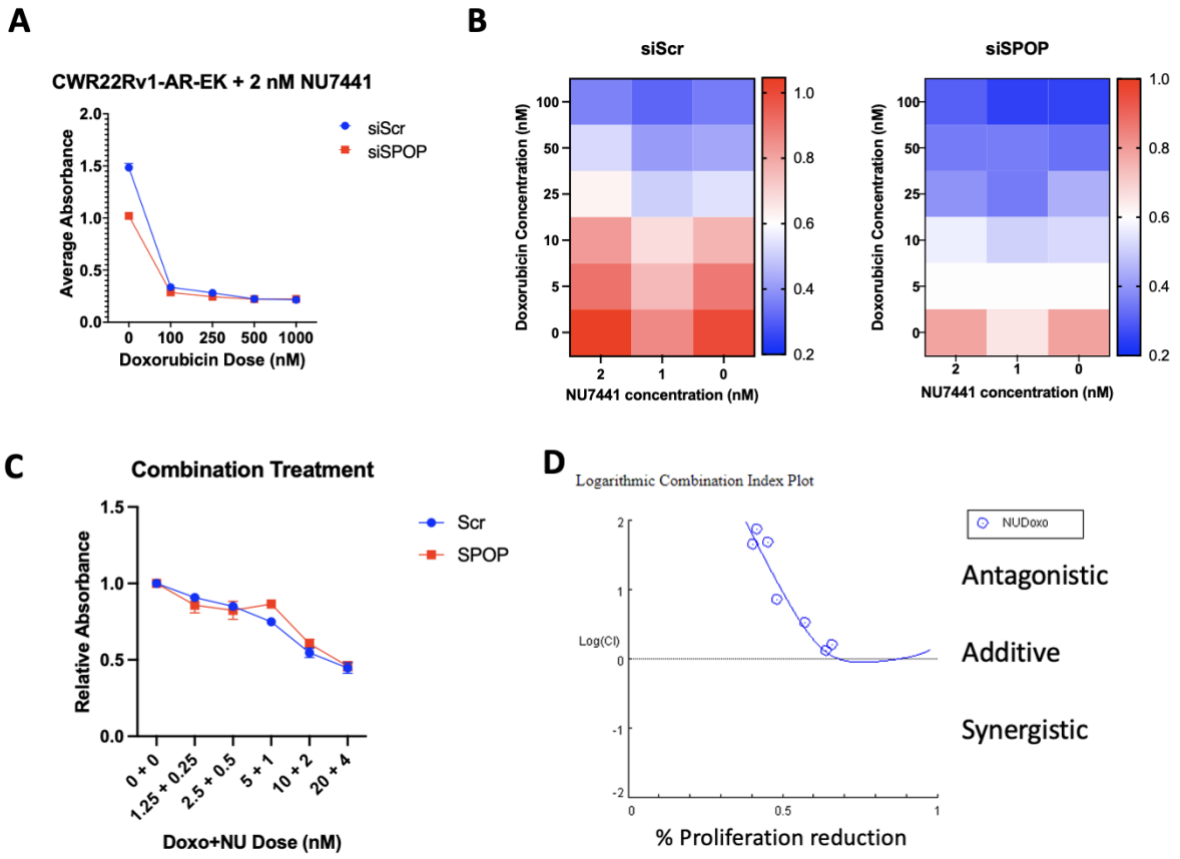


Figure 5.8 – SPOP depleted CWR22Rv1-AR-EK cells shows sensitivity to doxorubicin treatment but no additive effect with NU7441 treatment. A. *SRB assay for 0-1000 nM Doxo treatment dose range in CWR22Rv1-AR-EK cells transfected with either siScr or siSPOP siRNA in the presence of 2 nM NU7441. Data is representative of 3 individual repeats ± SEM.* ***B.*** *Heat maps for SRB assays using 0-100 nM Doxo dose range in the presence of vehicle control, 1 or 2 nM NU7441 in CWR22Rv1-AR-EK transfected with siScr or siSPOP siRNA. 1 (RED) – 0.2 (Blue) scale represents relative cell confluence at SRB endpoint. Data is representative of 3 individual repeats.* ***C.*** *SRB assay for combination treatments with Doxo and NU7441 at 5:1 ratio in CWR22Rv1-AR-EK cells transfected with either siScr or siSPOP siRNA in the presence of 2 nM NU7441. Data is representative of 3 individual repeats ± SEM.* ***D.*** *Compusyn graphical output for logarithmic combination index plot for combination of Doxo and NU7441. Data is representative of 3 individual repeats.*

2 nM and 1 mM NU7441 doses compared to a DMSO control. This experimental set-up was performed in CWR22Rv1-AR-EK cells, pre-treated with siScr or siSPOP siRNA to facilitate SPOP depletion prior to drug application [Figure 5.8B]. In the absence of NU7441, doxorubicin reduced siScr control cell number by 50% at 50 - 100 nM doses, while SPOP depletion saw this reduction between 25-50 nM. With 1 and 2 nM NU7441 treatments, cells continued to see this reduction between 50-100 nM of Doxo. In response to SPOP knockdown this 50% reduction took place between 10-25 nM, indicating that potentially a combination of Doxo and NU7441 resulting enhanced sensitisation within 'SPOP-MT' cells. Seeing this sensitisation to Doxo induced by SPOP depletion, the combination treatment was further investigated to assess whether this dual drug treatment was having a synergistic effect. This was done using CompuSyn software (Ver 1.0, available from https://www.combosyn.com) utilising the suggested set-up in which the two drugs are applied at a set ratio. The combination treatments

saw a steady reduction in proliferation as doses of both drugs increased, however there was no effects seen due to SPOP depletion [Figure 5.8C]. Data from these experiments was compiled then analysed using the CompuSyn software, showing that the combination treatment of NU7441 and doxorubicin (NUDoxo) yielded an antagonistic relationship [Figure 5.8D].

5.4.4 SPOP depletion combined with NU7441 treatment results in delayed resolution of γH2AX foci in response to IR induced DNA damage.

Given that cells depleted of SPOP demonstrated elevated sensitivity to combined NU7441 and IR treatment, it was important to study mechanistically what was the cause of the reduction in cell proliferation and viability. With SPOP implicated as a regulator of DNA double strand break repair, the ability of control and SPOP defective cells to repair DSBs was investigated using phosphor-γH2AX Ser139 foci resolution profiling by immunofluorescence as a surrogate marker for DNA repair competency. Here, CWR22Rv1, CWR22Rv1-AR-EK and PC3 cells were pre-treated with either siScr or siSPOP siRNAs, and 2 nM NU7441 or DMSO, prior to 2 Gy IR exposure before γH2AX foci per cell were quantified at 1-, 4- and 8-hour post post-IR [Figure 5.9]. In this set-up, for each experimental arm within each cell line γH2AX foci per cell was normalised relative to the siScr DMSO treated experimental arm 1-hour post IR exposure, essentially providing information as to the cells ability of foci resolution over the 8-hour window. In the CWR22Rv1 line, siScr control cells displayed resolution of foci over the 8-hour incubation period, with the relative foci/cell reducing from 1 to 0.57 [Figure 5.9A]. NU7441 treatment did not result in a decreased ability to resolve γH2AX foci when compared to the control arm, although did result in less initial foci/cell when compared to the control. siSPOP cells in the absence NU7441 treatment demonstrated resolution of γH2AX foci, initially starting with 1.02 foci per cell at the 1-hour time point which reduced past the control arm to 0.29 at the 8-hour time point. Crucially, however, across all time points tested, SPOP-depleted cells treated with NU7441 showed no resolution of γH2AX foci, with foci/cell increasing across the 8-hour course tested from 0.90 at 1 hour to 1.05 at the endpoint.

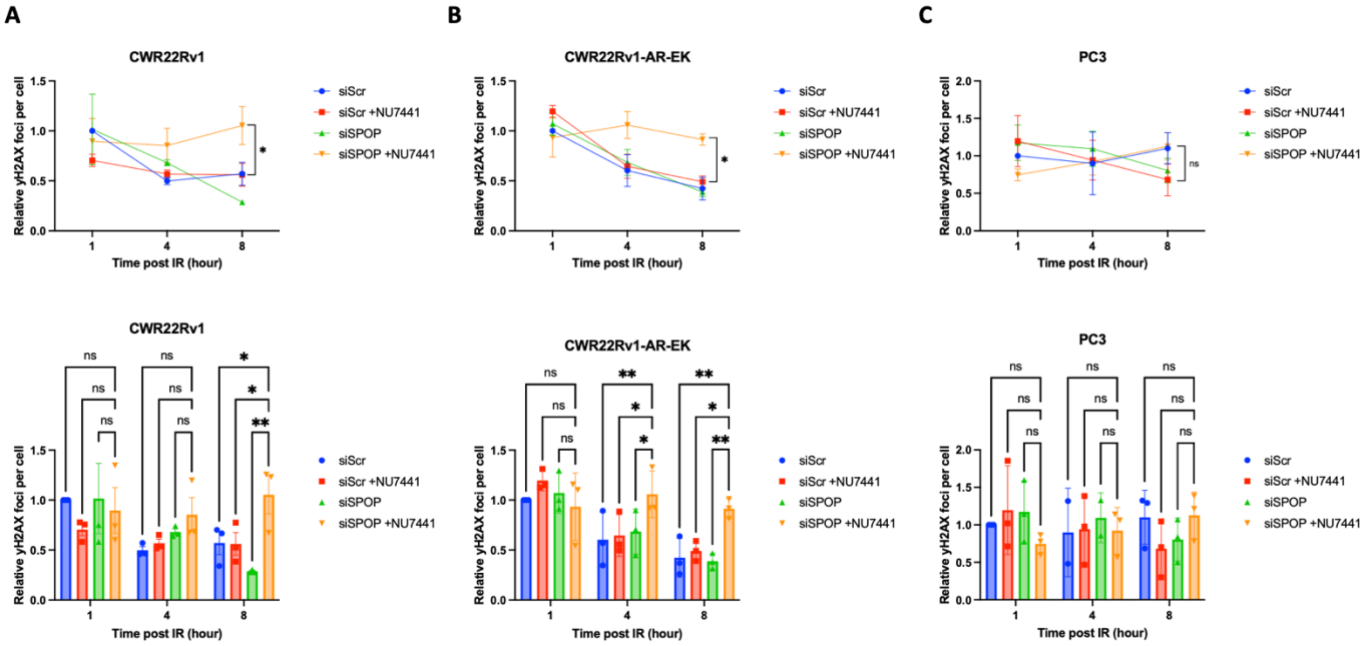


Figure 5.9 – SPOP knockdown alongside NU7441 treatment causes delayed γH2AX foci resolution 8-hours post IR exposure. *Immunofluorescent staining of γH2AX foci quantified per cell in CWR22Rv1 (**A**), CWR22Rv1-AR-EK (**B**) and PC3 (**C**) cells transfected with siScr/siSPOP siRNA, pre-treated with 2 nM NU7441 or vehicle control then exposed to 2 Gy IR before quantification of γH2AX foci quantified per cell at 1-, 4- and 8-hours post IR exposure. Foci per cell normalised siScr transfected cells treated with vehicle control at 1 hour for each line. Data is representative of 3 individual experiments ± SEM. Statistical significance was determined by a One-way ANOVA with Dunnett's multiple comparison test (*, **, *** and **** denote P-values <0.0332, <0.0021, <0.0002 and <0.0001 respectively, ns denotes not significant).*

A similar trend of results was demonstrated in the CWR22Rv1-AR-EK cell line [Figure 5.9B]. Here, the siScr control arm demonstrated clear resolution of γH2AX foci, with the relative foci per cell reducing to 0.42 after the 8-hour incubation. Consistent with CWR22Rv1 cells, both the siScr +NU7441 and SPOP-depleted experimental arms displayed the same trend of foci resolution across the 8-hour time course. Specifically, the siScr+NU7441 and siSPOP arms demonstrated 1.20 and 1.07 relative foci/cell at the 1-hour timepoint, respectively, which was resolved to a respective 0.49 and 0.39 relative foci/cell by the end of the 8-hour incubation. In contrast, in response to combined SPOP depletion and NU7441 treatment, no resolution of γH2AX foci observed. Here the relative foci/per cell 1 hour after IR was 0.93 and remained at this level to the end of the 8-hour time course.

Interestingly, these results were not mirrored in the PC3 cell line [Figure 5.9C]. Over the entire 8-hour time course for each of the experimental cell lines, there was no significant difference in γH2AX resolution and a high degree of variance observed within each treatment arm over the duration of the experiment. It is possible that PC3 cells may require longer to resolve DNA damage. There is evidence from Shaheen et al., (2011) to suggest that PC3 cells display

resolution of γH2AX foci within 4 hours in response to 10 Gy IR, thus a potential increase in IR dose in further experimental work within the PC3 cell line may result in greater damage and more noticeable differences in γH2AX foci resolution (Shaheen et al., 2011).

To expand this experiment further, and to potentially remedy the lack of γH2AX resolution in the PC3 cell-line, cells were transfected with siRNA, treated with or without NU7441 and irradiated before incubation for 24 hours prior to γH2AX foci quantification. For comparison between experimental arms, the number of foci/cell was normalised to that of the control experimental arm at 24 hours [Figure 5.10]. In the CWR22Rv1 line, relative foci/cell was found to be higher in the siScr +NU7441 and siSPOP arm compared to the siScr control arm, with a further increase seen upon SPOP depletion and NU7441 treatment [Figure 5.10A]. The CWR22Rv1-AR-EK cell line displayed similar results to the 8-hour timepoints [Figure 5.10B]. The control+NU7441 returned 1.77 relative foci/cell whereas the siSPOP treated experimental arm returned 2.96 relative foci/cell. Interestingly the siSPOP+NU7441 treatment arm returned the highest number of unresolved foci with 5.73 relative foci/cell following the 24-hour incubation period supporting the previous findings that SPOP depletion and DNA-PKcs blockade diminish cell viability as a consequence of failure to repair DNA breaks.

For PC3 cells, both the siScr, and siSPOP minus NU7441, experimental arms displayed no significant increased γH2AX foci number over the control. However, in contrast to data acquired at the 8-hour time-point, PC3 cells depleted of SPOP and subject to DNA-PKcs blockade demonstrated a statistically significant increase in γH2AX foci at the 24-hour time point, suggesting SPOP knockdown in combination with NU7441 compromises IR-induced DNA damage repair across several PC backgrounds.

5.4.5 Overexression of SPOPWT can overcome delayed resolution of γH2AX foci in response to SPOP-knockdown and DNA-PKi treatment in PC3 cells.

Utilising the doxycycline-inducible PC3-iSPOP^{WT} cell lines generated previously, similar experiments were conducted to establish if the delayed resolution of γH2AX foci can be

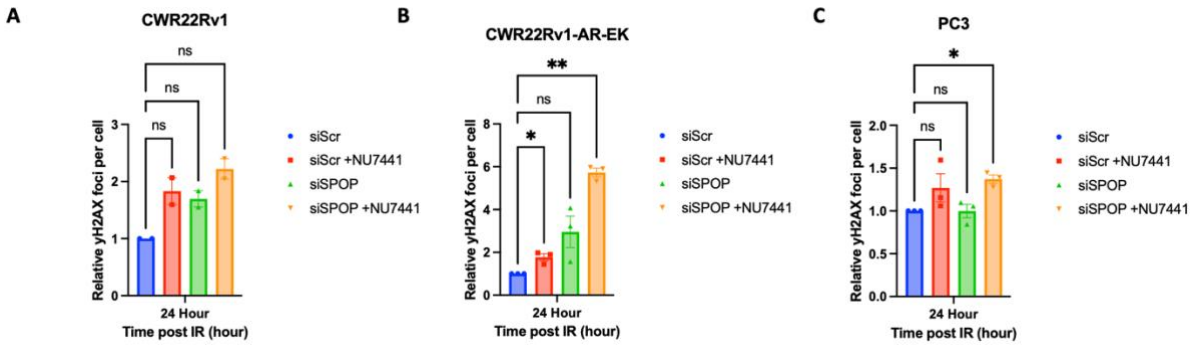


Figure 5.10 – SPOP knockdown alongside NU7441 treatment causes delayed γH2AX foci resolution 24-hours post IR exposure. *Immunofluorescent staining of γH2AX foci quantified per cell in CWR22Rv1 (**A**), CWR22Rv1-AR-EK (**B**) and PC3 (**C**) cells transfected with siScr/siSPOP siRNA, pre-treated with 2 nM NU7441 or vehicle control then exposed to 2 Gy IR before quantification of γH2AX foci quantified per cell at 24-hours post IR exposure. Foci per cell normalised siScr transfected cells treated with vehicle control at 1 hour for each line. Data is representative of 2 individual experiments ± SEM in the CWR22Rv1 line and 3 individual experiments ± SEM in the CWR22Rv1-AR-EK and PC3 lines. Statistical significance was determined by a One-way ANOVA with Dunnett's multiple comparison test (*, **, *** and **** denote P-values <0.0332, <0.0021, <0.0002 and <0.0001 respectively, ns denotes not significant).*

rescued by over-expression of wild-type SPOP. Due to the previous observation that 24-hours post IR exposure PC3 cells demonstrated elevated γH2AX foci only with siSPOP and NU7441 treatment, potentially by overexpressing $SPOP^{WT}$ protein in this experimental set-up the lack of foci resolution would be overcome and return to normal levels, ultimately resolving the effects of SPOP knock-down and NU7441 treatment. As such, cells were either transfected with siScr or siSPOP-UTR siRNAs, as in Section 4.4.4, treated with Dox or vehicle with and without 2 nM NU7441 prior to quantification of γH2AX foci per cell 1-, 8- and 24-hours following 2 Gy IR exposure [Figure 5.11]. Alongside these experimental arms, a vehicle-treated, siScr-transfected control arm was set-up; the average foci per cell of which was used to normalize the data from each other experimental arm. After 1-hour post-IR treatment, an increase in γH2AX foci ratio over the untreated control was observed for each of the experimental arms, but no significant difference was demonstrated. After 8 hours, clear resolution of the γH2AX foci can be seen in all arms with the exception of the siUTR +NU7441 treated arm, which saw a reduction in γH2AX foci from 2.00 after 1 hour to 1.55 following 8 hours post-IR exposure. Interestingly when comparing this arm to the siUTR +Dox +NU7441 experimental arm, the effect of $SPOP^{WT}$ overexpression is appreciable, in which a decrease from 2.57 relative foci per cell at 1-hour post-IR to 0.33 8-hours post-IR, ultimately overcoming the delayed foci resolution and returning to the foci levels seen in in the control arm. Furthermore, following 24 hours, this trend continues, where in the siUTR +NU7441 treated arm, relative foci per cell was 0.37, significantly greater than the 0.03 relative foci per cell seen in the siUTR +Dox +NU7441 where WT-SPOP has been overexpressed. Interestingly, although not as apparent in the SPOP

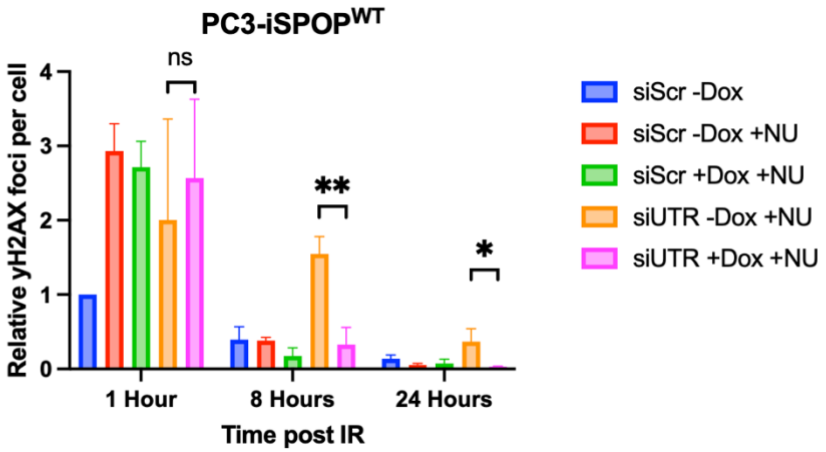


Figure 5.11 – Induction of SPOP is able to rescue phenotypes delayed resolution caused by SPOP knockdown and NU7441 treatment in PC3 cells. *Immunofluorescent staining of γH2AX foci quantified per cell in PC3-iSPOPWT cells transfected with siScr/siSPOP-UTR siRNA, pre-treated with 2 nM NU7441 or vehicle control then exposed to 2 Gy IR before quantification of γH2AX foci quantified per cell at 1-, 8- and 24-hours post IR exposure. Foci per cell normalised siScr transfected cells treated with vehicle control at 1 hour. Data is representative of 3 individual experiments ± SEM. Statistical significance was determined by T test (*, **, *** and **** denote P-values <0.0332, <0.0021, <0.0002 and <0.0001 respectively, ns denotes not significant).*

depleted experimental arms, overexpression of SPOP can be seen to increase the effectiveness of foci resolution when comparing the siScr +NU7441 and siScr +Dox and +NU7441 treated arms.

Finally, to confirm these effects seen as a result of siRNA-mediated SPOP depletion, this experiment was repeated within the newly developed CWR22Rv1-AR-EK-SPOP$^{-/-}$-C1 model. With previous experiments showing unresolved γH2AX foci 24-hours post IR exposure, the experimental set-up was adapted, looking at γH2AX foci levels 1- and 24-hours post IR treatment whilst also capturing γH2AX foci in the absence of IR exposure. Following pre-treatment with 2 nM NU7441, cells were harvested pre-IR, 1- and 24-hours post IR [Figure 5.12]. This allowed for detection of any increases in basal γH2AX foci levels in response to NU7441 treatment or SPOP knock-out alongside capturing the ability of cells to resolve gγH2AX foci caused by 2 Gy IR exposure. In the absence of IR exposure, CWR22Rv1-AR-EK-iCas9 cells showed no increase in γH2AX foci in response to NU7441 treatment, with 0.82 foci/cell seen relative to the control experimental arm. In CWR22Rv1-AR-EK-SPOP$^{-/-}$-C1 cells, there were basal levels of 2.43 relative foci/cell which increased further upon NU7441 addition to 3.56 relative foci/cell. As seen previously, all experimental arms saw a drastic increase in γH2AX foci abundance upon 1 hour post-IR exposure, which returned to the basal levels by the 24-hour

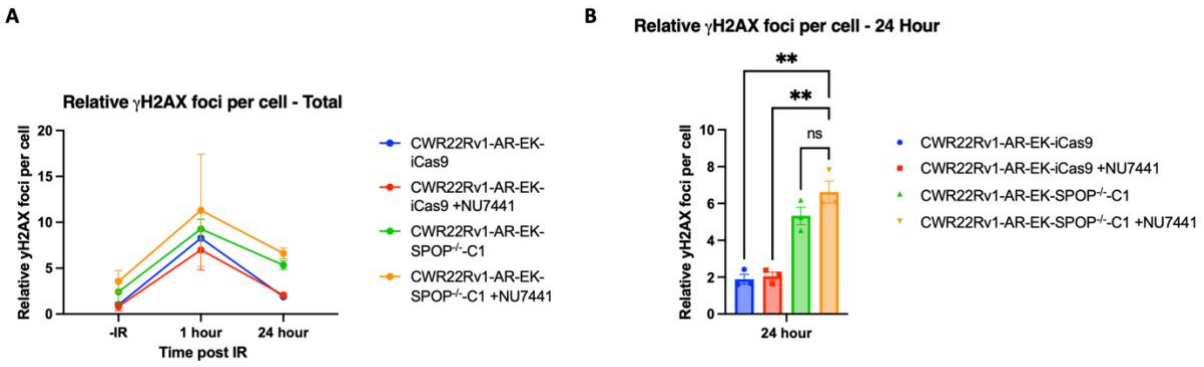


Figure 5.12 – CWR22Rv1-AR-EK-SPOP$^{-/-}$-C1 cells show higher γH2AX foci per cell at 24-hours compared to parental cells regardless of NU7441 treatment. *Immunofluorescent staining of γH2AX foci quantified per cell of CWR22Rv1-AR-EK-iCas9, CWR22Rv1-AR-EK-SPOP$^{-/-}$ -C1 pretreated with 2 nM NU7441 or vehicle control, in the absence of IR alongside 1- and 24- hours post 2 Gy IR exposure **(A)** alongside endpoint statistical analysis at 24 hours (**B**). Foci per cell normalised –IR vehicle control in parental cells. Data is representative of 3 individual experiments ± SEM. Statistical significance was determined by a One-way ANOVA with Dunnett's multiple comparison test (*, **, *** and **** denote P-values <0.0332, <0.0021, <0.0002 and <0.0001 respectively, ns denotes not significant).*

time-point. For both CWR22Rv1-AR-EK-iCas9 vehicle control and NU7441-treated experimental arms, the relative foci/cell increased to 8.28 and 7.00, respectively, 1 hour after IR exposure and reduced to respective 1.89 and 2.06 at the 24-hour time-point. This trend was not observed in the CWR22Rv1-AR-EK-SPOP$^{-/-}$-C1 derivative which saw a higher amount of foci/cell after the 24-hour incubation regardless of NU7441 treatment, with 5.33 relative foci/cell seen in the CWR22Rv1-AR-EK-SPOP$^{-/-}$-C1 experimental arm at 24-hours and a further 6.62 relative foci/cell seen with the addition of NU7441 [Figure 5.12B].

Taken together these results provide evidence that in response to SPOP depletion and NU7441 treatment, a range of PC cell lines demonstrate an impaired ability to resolve DSBs induced by IR. Furthermore, upon complete loss of SPOP function by gene knockout, CWR22Rv1-AR-EK cells show a higher basal level of DSBs and show delayed DSB resolution which is further enhanced by NU7441 treatment.

5.4.6 Cell cycle effects of NU7441 treatment and Ionising radiation treatment in the presence of SPOP knock-down.

Having demonstrated that SPOP-deficient PC cell lines show compromised resolution of DNA DBSs in response to DNAPKcs inhibitors and IR treatment, further experiments were conducted in an attempt to better understand the cell cycle impact of these cell manipulations. As before, CWR22Rv1-AR-EK and PC3 cells were transfected with scrambled control or SPOP-targeting siRNA, treated with either 2 nM NU7441 or vehicle, and exposed to 2 Gy IR, before analysing cell cycle by PI-based flow cytometry after a further 48 hour incubation [Figure 5.13]. Here only two replicates were performed for each treatment arm, meaning only indications of cell cycle effects in response to prescribed treatments can be inferred from this data, with final replicates necessary to confirm and assess statistical significance of effects seen. Examining the effect of just SPOP depletion upon cell cycle profile of CWR22Rv1-AR-EK cells, minimal changes were detected; with slight increase in S-phase at the expense of G2/M [Figure 5.13A]. In response to single agent NU7441, IR and combination treatments within the siScr experimental arms, again minimal change were observed [Figure 5.13B]. However, these same treatments of SPOP-depleted cells were found to impact cell cycle profile. Specifically, in the IR alone and NU+IR combination treatments, decreases in G1 and G2/M phases and a concurrent increase in S-phase was observed [Figure 5.13C]. Finally, comparing the combination treatments in both the siScr and siSPOP arms, a slight decrease in G1 and G2/M with a slight increase in S phase was found [Figure 5.13D].

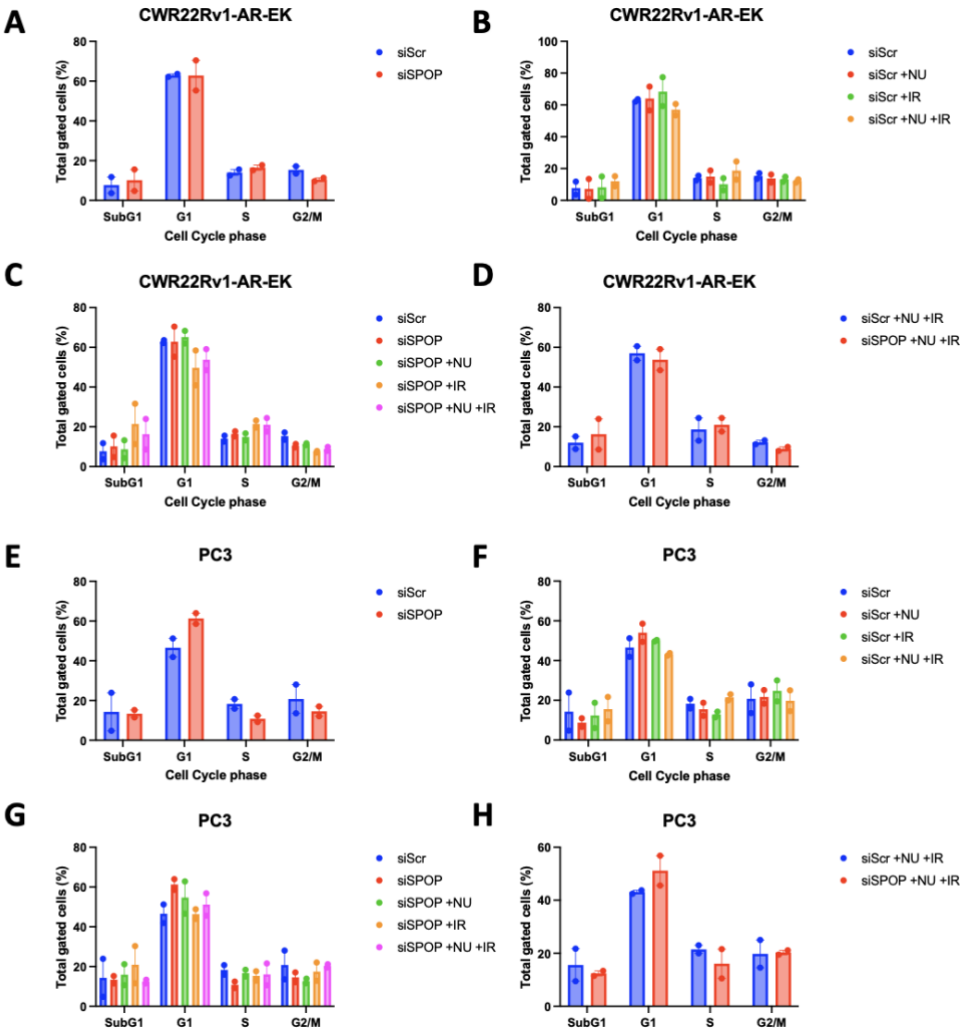


Figure 5.13 – Cell cycle effects of SPOP depletion, NU7441 treatment and IR exposure in CWR22Rv1-AR-EK and PC3 cells 48 hours after IR exposure. *CWR22Rv1 (**A-D**) and PC3 (**E-H**) cells were transfected with siScr/siSPOP siRNA, pretreated with 1 μM NU7441 or vehicle control before necessary treatment arms were exposed to 2 Gy IR and returned to incubate for 48 hours before PI staining cells. Cells were gated based of cell cycle stage with proportion of total gated cells in each respective phase represented graphically. Data is representative of 2 individual experiments ± SEM.*

For the PC3 cell line experiments, more pronounced cell cycle effects were seen in response to SPOP depletion alone; with an observed increase in G1 and decrease in S and G2/M phase [Figure 5.13E]. Minor changes were observed in response to NU7441, IR and the combination treatment within the siScr arms. However, in SPOP-depleted cells, there was a general decrease in G1-phase in response to individual and combined NU7441 and IR treatment [Figure 5.13F-G]. For clarity, and as shown in Figure 5.13H, combined NU7441 and IR treatments in SPOP-depleted PC3 cells respectively increased and decreased G1- and S-phases

of the cell cycle compared to siScr control cells albeit in two independent experimental repeats.

To extend findings, similar experiments were conducted with the elevated dose of 1 μM NU7441 which is more representative of the previously calculated IC_{50} dose and harvested after a longer incubation period of 120 hours [Figure 5.14]. For the CWR22Rv1-AR-EK cells, SPOP knockdown alone showed modest changes in cell cycle distribution which were consistent with the previous cell cycle experiment [Figure 5.14A]. In siScr-transfected CWR22Rv1-AR-EK cells, IR exposure reduced both G1- and S-phase concurrent with an increase in the G2/M phase [Figure 5.14B]. Furthermore, G1 looked to decrease and an increase in S phase was observed in the SPOP-depleted cells treated with single agent or combined NU7441 and IR treatments, with the combination showing increases in G2/M [Figure 5.14C]. Comparing the effect of the combination of NU7441 and IR exposure between siScr and siSPOP cells, depletion of SPOP resulted in a subtle S-phase increase but the N=2 nature of this experiment diminishes the strength of the conclusions [Figure 5.14D].

In the PC3 line, depletion of SPOP alone had no marked effect on cell cycle profile which is in contrast to that seen for the lower dose NU7441 and 48 hour experiment. In siScr cells, a sequential reduction in G1 was demonstrated with respective single and combined NU7441 and IR treatment which correlated with an elevation in subG1, although there were no clear effects on S or G2/M [Figure 5.14F]. In the background of SPOP knock-down, a G1 decrease

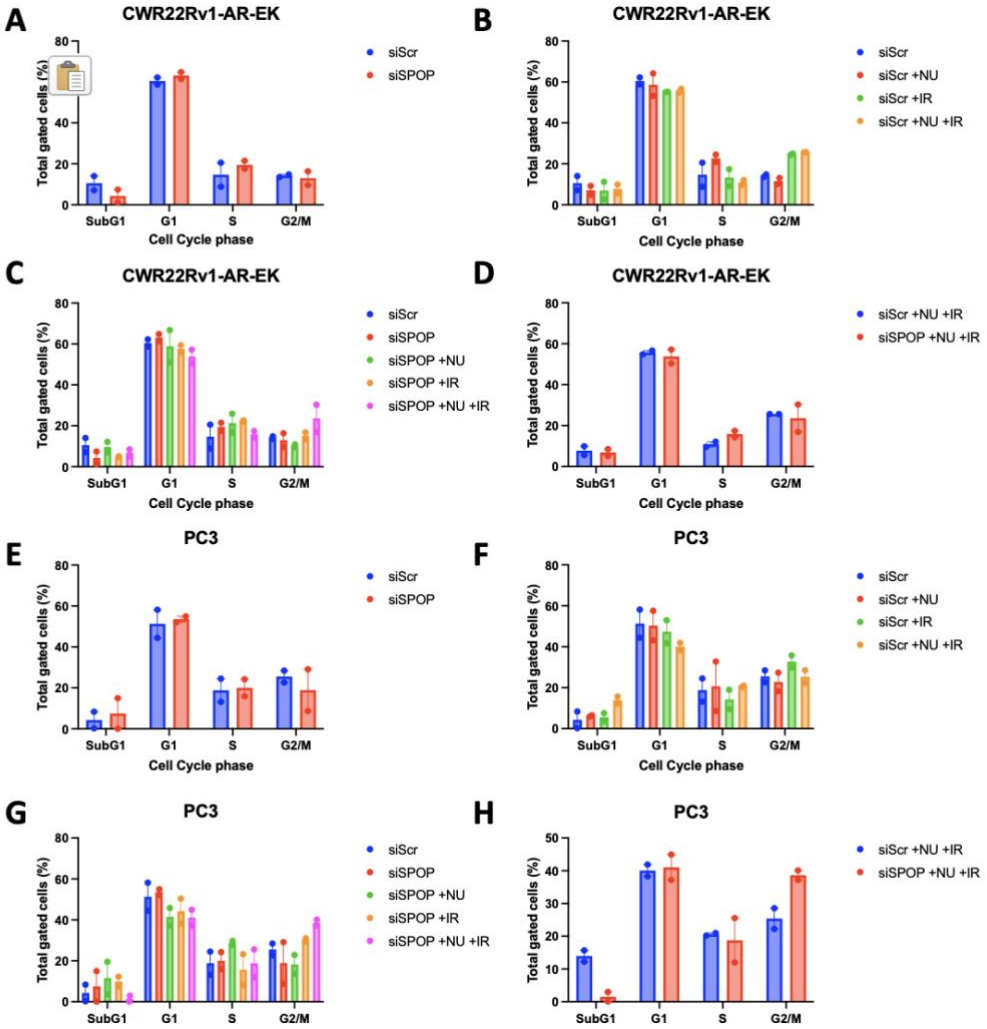


Figure 5.14 - Cell cycle effects of SPOP depletion, NU7441 treatment and IR exposure in CWR22Rv1-AR-EK and PC3 cells 120 hours after IR exposure. *CWR22Rv1 (**A-D**) and PC3 (**E-H**) cells were transfected with siScr/siSPOP siRNA, pretreated with 1 μM NU7441 or vehicle control before necessary treatment arms were exposed to 2 Gy IR and returned to incubate for 120 hours before PI staining cells. Cells were gated based of cell cycle stage with proportion of total gated cells in each respective phase represented graphically. Data is representative of 2 individual experiments ± SEM.*

was seen in all experimental arms apart from the knockdown alone, with increases in G2/M observed upon IR exposure and a combination of NU7441 and IR [Figure 5.14G]. Finally, comparing the combination both in the presence and absence of SPOP knockdown a clear decrease in SubG1 and increase in G2/M can be seen [Figure 5.14H].

5.5 Discussion

SPOP mutations have been shown to have drastic effects on normal cell homeostasis, with one of the key pathways perturbed being DDR. Observed in a PC setting, *SPOP* mutations confer repression to a range of known HR-associated genes leading to an increased rate of accumulation of genetic abnormalities (Boysen et al., 2015, Hjorth-Jensen et al., 2018). With synthetic-lethality relationships seen in response to inhibition of other DNA repair pathways, and the hypothesised over-reliance on the NHEJ, potentially through inhibition of alternate key NHEJ-mediated DSB repair proteins could reveal new synthetic lethal relationships in *SPOP* mutant cells. One such protein playing a key role in NHEJ is DNA-PKcs, which fortunately also has a range of inhibitors available; some of which are promising clinical candidates. Through inhibition of DNA-PKcs, in the presence of SPOP mutations, cells may be unable to reliably repair DNA damage and ultimately drive them towards an apoptotic cell fate.

Work in this chapter looked to investigate the presence of this hypothesised synthetic-lethality relationship through assessing the effects on cell growth of SPOP depleted cells through a range of cellular assays as well as gain mechanistic insight into the underlying cell events causing any of the phenotypes observed. Also investigated was how cells would respond in the presence of these combination treatments to exogenous sources of DNA damage, either through IR exposure or Doxorubicin treatments. The described synthetic-lethality was confirmed within PC cell lines using a range of growth assays alongside either SPOP depletion, through siRNA transfection, or complete SPOP protein knockout using CRISPR/Cas9 gene editing. Also observed was that this lethality was enhanced in the presence of IR, with the mechanism behind the reductions in proliferation and cell viability being caused by high amounts of unrepaired DNA damage.

The key finding from this project was the discovery of the synthetic-lethal relationship between SPOP depleted cells and DNA-PKcs inhibition, and that this can be enhanced through the introduction of exogenous sources of DNA damage. When compared to the synthetic lethal relationships in CWR22Rv1 cells described between SPOP knock-down and PARPi in Boysen et al., with an IC_{50} shift from 1.8 μM in control cells to 0.71 μM in the SPOP depleted cells in response to Olaparib (Boysen et al., 2015). This 2.5-fold increase in sensitivity to Olaparib is comparable to the 1.5-fold and 2.1-fold increase in sensitivity to NU7441 seen in the

CWR22Rv1-AR-EK-SPOP$^{-/-}$ clones when compared to parental controls. Also conducted within this paper was similar investigations into response to Olaparib in PC3 and LNCaP cells, which both displayed enhanced sensitivity than that seen in the CWR22Rv1 cell line (Boysen et al., 2015). Potentially, if experiments were conducted in these alternate cell backgrounds similar trends of enhanced sensitivity to DNA-PKIs would be seen.

With a substantial amount of data suggesting the presence of interplay between SPOP and sensitivity to DNA-PKIs, it was necessary to confirm this within alternate cellular backgrounds. One of the key benefits to the data within the CWR22Rv1-AR-EK PC cell line is that represents an AR targeted treatment refractory disease (late stage), the sensitivity to DNA-PKcsi is still seen. As such, potentially through looking into alternate PC cell lines, there may be more drastic sensitivities seen. Investigation into alternate cell line models would provide greater insights into the contributions of AR signalling on the underlying mechanisms behind the sensitivities observed. Polkinghorn et al., have described that AR activity contributes towards the expression of key DDR genes. As such, in the CWR22Rv1-AR-EK background, it would be expected that there is constitutive activation of DDR genes (Polkinghorn et al., 2013, Kounatidou et al., 2019). If alternate PC cell lines were to be used, such as LNCaP, where AR activity could be modulated through anti-androgens, or PC3, which are AR negative, then there may be increased sensitivity to DNA-PKcsi in an SPOP mutant background due to increase amounts of un-repaired DNA damage. With the AR being a known SPOP substrate, the use of these PC cell lines will also provide the opportunity to investigate the interplay between SPOP and AR, alongside how introducing DNA-PKcsi treatment and IR exposure affect this relationship.

Another caveat to the data presented is that the representation of SPOP mutations used was depletion using siRNA transfection and CRISPR/Cas9 mediated knock-out. These do look to phenocopy the repression of HR genes described in the literature, however it is unknown what the effects of loss of the SPOP protein compared to the protein being present but functionally inactive, as in the case of the MATH domain mutations seen in patients. Currently, there is nothing described in the literature addressing this fact, however through the use of knock-out models, where SPOP wild-type or mutant expression has been reintroduced through plasmid-based expression (as described in Section 4.4.4), the presence or absence of any such function of the SPOP protein can be confirmed. This may be the basis for any differences seen between overexpression models and the knock-out or siRNA mediated depletion models used here.

A key issue with the data presented is the use of NU7441, which is not a current clinical candidate. The data here is still relevant although if this work was to be taken further, the relationships observed here would have to be recapitulated with current clinical candidate DNA-PKcs inhibitors. Re-runs of the experiments using AZD7648 or peposertib, which is in early phase clinical trials. are necessary to confirm the if the effects seen are consistent with new generations of drugs which would be of great importance. Further, although initial indications have hypothesised unrepaired DNA damage being the cause of the proliferation and cell viability reductions seen, more in-depth understanding to how this causes either cell death or senescence is necessary. Use of a caspase cleavage assay to assess the induction of apoptosis or an Annexin V stain and assessment of cell death via flow cytometry would be of great interest to confirm the suspected apoptosis. If the results of this looked to be directing towards cell senescence, in depth analysis of cell senescence markers would be useful to confirm this as the cause of reduced cell proliferation.

One of the most promising results presented within this chapter is the effects of DNA-PKcsi treated SPOP knock-out cells in response to an IR dose range. These findings are consistent with data presented by Bezaway et al., which describes DU-145 and PC3 PC cell lines showing significantly enhanced sensitivity to 2 Gy IR exposure when SPOP is depleted using siRNA transfections (El Bezawy et al., 2020). In data presented in this chapter, even down to the smallest IR dose used, significant losses in cell viability and enhanced sensitivity to NU7441 was displayed in both CWR22Rv1-AR-EK-SPOP$^{-/-}$ clones when compared to parental controls. Ultimately this data indicates that the dose of IR or NU7441 can be lowered when applied in combination which may potentially alleviate some of the clinical toxicity seen from the drug and the negative side-effects of IR exposure.

With most of the experiments in this chapter using IR as an exogenous source of DNA damage, there may be questions regarding to how cells respond to alternate DNA damage inducing agents. Data generated using Doxorubicin to induce DNA damage in combination with DNA-PKcsi in the presence and absence of SPOP knock-down looks to have displayed the same trends as seen with IR exposure. This data is very useful as it gives early indications that the synthetic-lethality relationship can be enhanced using different forms of DNA damaging agent although this would need more investigation using other DNA damaging agents to fully demonstrate that the source of DNA damage is irrelevant in terms of enhancing sensitivity to DNA-PKcsi. Another experiment of interest would be assessing if the results of the clinical trial

combining AZD7648 and Doxorubicin (analogues) could be recapitulated in the SPOP knock-out models, as this could provide further confidence how representative these models are of patient *SPOP* mutant tumours. Unfortunately, due to the known issues of with using Doxorubicin clinically in a PC setting, this work was not continued although using alternate DSB inducers like focal radiation are of high interest and would be interesting future work for the project.

Experiments looking at the resolution of γH2AX foci following IR exposure showed delayed DSB resolution in response to DNA-PKcsi treatment in the background of SPOP knockdown or knockout. When comparing to the control experimental arms in the SPOP knock-down experiments there was clear resolution of γH2AX foci across the initial 8-hour time-course in both the CWR22Rv1 and CWR22Rv1-AR-EK derivative PC cell lines, although no clear resolution in the PC3 cell line. This may be due to the previously mentioned AR contributions to DNA repair, with the PC3s lacking this activity and hence DNA damage could take longer to resolve, which is supported by the experiments set up with the 24-hour endpoint beginning to show significant differences between the control experimental arm. Looking to the literature, experiments using PC3 cells have used higher initial doses of IR to assess DNA damage resolution, so experimental re-runs with IR doses greater than 2 Gy may provide clearer results (Shaheen et al., 2011).

Another issue with this initial experimental set-up lies with the lack of negative control looking into endogenous γH2AX levels, as this would have allowed to track if by 8 or 24-hours post IR exposure foci levels had returned to normal levels, whilst also giving insights into if SPOP depletion caused increased steady-state levels of DNA damage. This was taken onboard when re-running this experiment with the knock-out models, where increased γH2AX levels can be seen in response to SPOP loss in the absence of IR, alongside the previously observed lack of ability to resolve DNA damage across the 24-hour time course. Of interest here, although it looked to enhance γH2AX levels at 24-hours, DNA-PKcsi did not have a substantial effect over the complete loss of functional SPOP within the cells, indicating that SPOP may play wider roles in DDR than previously described in the literature. Ultimately this gives strong indications that in response to SPOP depletion and DNA-PKcsi, unresolved DNA damage is causing the loss of cell viability and subsequent synthetic lethality.

Another experiment that shows very promising results was the ability to phenotypically rescue the lack of γH2AX foci resolution in response to SPOP depletion and DNA-PKcsi through the overexpression of SPOP via inducible plasmid-based expression systems within the PC3 cell lines. Similar experiments were set up using the PC3-iSPOP^{W131G} cell line derivatives although complete data was not able to be generated within the timeframe of the project. It is speculated that there would be no difference to the rate of foci resolution in the presence of overexpressed mutant SPOP, which would confirm that the phenotypes observed in other similar experiments were truly SPOP dependent changes. There were some differences when comparing this data to the previous knock-down experiments in the PC3 cell line, namely that no clear foci resolution was seen at 8-hours. However, in this rescue set-up, a marked increase in foci was observed at 8-hours. This could potentially be due to the variability seen when using the SPOP siRNA pool and the differences in siRNA efficiency when compared to the SPOP-UTR targeting siRNA. Also, the number of cells per image used in the downstream analysis was less than that of the previous experiments, as only cells that were looked to be expressing SPOP (via GFP fluorescence) were selected for analysis, as expression was not homogenous across the timepoints when cells were harvested. This was potentially due to cells being treated 48 hours after Dox induction, so in the 24-hours following this expression might be lost. If these lines were to be used further wash out experiments, it may be necessary to assess how long SPOP expression is sustained following Dox induction, either using GFP via immunofluorescence or SPOP protein levels via western blot across an extended time period. As promising as these initial results are, in future it would be best to use inducible expression systems in the knock-out background to remove siRNA variability and any residual endogenous SPOP expression from the system.

These findings align with data presented by Boysen et al., and Hjorth-Jensen et al., suggesting impaired ability to resolve DNA damage in response to loss of SPOP function. In these studies, overexpression of SPOP-WT or SPOP mutants resulted in increased COMET tail movement, delayed clearance of 53BP1 foci (NHEJ positive regulator) and decreased RAD51 foci (HR engagement marker) formation in RWPE cells (Boysen et al., 2015). Here, no clear differences in clearance of γH2AX foci was observed across the 3-hour time course in all experimental arms. Consistent with this, SPOP siRNA mediated knock-down showed delayed RAD51 foci formation, which could be phenotypically rescued through overexpression of wild type SPOP but not mutant SPOP (Hjorth-Jensen et al., 2018). With respect to these experiments, factoring

in these alternate markers of DDR could greatly boost the impact of future investigation into SPOP loss and DNA damage repair and how DNA-PKcsi adds to this.

To further understand the mechanistic effects of SPOP knock-down and DNA-PKcsi, cell cycle analysis via flow cytometry was used. From the experiments conducted, no major trends were able to be interpreted which was made more challenging by only 2 experimental runs performed from each experiment. Lack of replicates aside, a key issue with the inter-experimental variability seen could be due to the variability of siRNA knock-down efficiency which was also seen across other experiments using this method of SPOP depletion. As with the investigation into the sensitivity to NU7441 treatment, high degrees of variability between experimental replicates were observed from early cell count and clonogenic assay experiments using SPOP knock-down. To alleviate the impact of knock-down variability, re-running similar experiments using the CWR22Rv1-AR-EK-SPOP$^{-/-}$ cell line derivatives were attempted. Unfortunately, only one experimental replicate was able to be completed within the project thus was not included. To generate more useful data from this experiment, as mentioned above, factoring in an Annexin V stain into the protocol could reveal more useful insights into the underlying cellular events causing the lack of proliferation seen, as well as potentially stalling cell cells within a given cell cycle stage then releasing and gathering data after a set time course could also remove some of the variability seen in this initial set-up.

Together data presented here strongly suggests that SPOP mutations can be exploited through inhibition of DNA-PKcs to drive accumulation of DNA damage resulting in synthetic-lethality. Further evidence suggests that exogenous DNA damage induction will ultimately enhance this relationship, allowing for lower doses of inhibitors to be used. With all current data being cell line based, further investigation into alternate cell line backgrounds as well as more representative *in vivo* models are vital to confirm if this work could be used to bring about patient benefit clinically. Also of note is the now commercially available SPOP inhibitor which could, if validated, see use in expanding the benefits of repurposing DNA-PKcs inhibitors from the subset of patients harbouring *SPOP* mutations to PC patients as a whole through combination treatments. Seeing and validating the presence of the sensitivity to DNA-PKcs inhibition, alongside the range of models able to replicate and align with published data, there is scope to take this work further and look to other DDR inhibitors, as well as other key SPOP regulated pathways, and assess if these see similar vulnerabilities to that presented within this chapter.

Chapter 6 – Transcriptomic effects of loss of SPOP function in PC cell line models

6.1 Introduction

A range of SPOP substrates have been described which can aggregate and have drastic effects on normal cell homeostasis in the context of PC-associated *SPOP* loss of function mutations. Many publications have investigated specific SPOP-substrate interactions and the impact on this interplay by mutation, although few have attempted to look at the broader impacts of loss of SPOP function in the terms of transcriptomic or proteomic changes.

Using the Gene expression omnibus (GEO), 4 datasets investigating the effects of SPOP manipulation have been published. Of these, 3 investigate overexpression of either wild-type or mutant forms of SPOP, with the other looking into SPOP depletion via siRNA knock-down. The first publication conducted experiments assessing general transcriptomic changes in LNCaP-Abl prostate cancer lines, overexpressing a control vector, $SPOP^{WT}$, $SPOP^{F102C}$, $SPOP^{F133V}$ or $SPOP^{F133L}$ (Geng et al., 2014). The next focussed more on the effects of SPOP and its mutants on the AR signalling axis, where, amongst other samples, either $SPOP^{WT}$ or $SPOP^{F102C}$ was overexpressed and grown in either androgen replete or deprived media (Burleson et al., 2022). The final overexpression study conducted investigated global expression changes, with a focus on DDR effects in DU145 cells overexpressing either $SPOP^{WT}$ or $SPOP^{F133V}$ (El Bezawy et al., 2020). Finally, looking at comparisons in differentially expressed genes between SPOP knock-down and knockdown of its upstream regulator G3BP1, data is available for the effects of SPOP siRNA mediated knock-down in CWR22Rv1 cells (Mukhopadhyay et al., 2021). Two publications have been identified assessing cell proteome changes caused by the overexpression of wild-type or mutant SPOP, with one in C4-2 cells and the other in mouse prostate cells (Blattner et al., 2017, Geng et al., 2023).

Data presented in Chapter 5 indicated loss of SPOP function confers sensitivity to DNA-PKcs inhibition. Alongside the novel SPOP knock-out models generated in the CWR22Rv1-AR-EK cell background in Chapter 4, investigation into the transcriptomic changes may provide

indications of alternate cellular pathways sensitive to inhibition in a SPOP mutant background. SPOP depletion has been noted to cause increased radiosensitivity, indicating a shift in cellular response to IR (El Bezawy et al., 2020). Therefore, investigations into the transcriptional shifts caused by SPOP loss combined with IR exposure may provide insights into target pathways for inhibition to potentiate IR effects.

6.2 Aims

Data presented in Chapter 5, using CWR22Rv1-AR-EK-SPOP$^{-/-}$ PC cell line derivatives, has indicated the presence of a synthetically-lethal relationship between *SPOP* loss of function mutations and DNA-PKcs inhibitors. With many indications across the literature showing SPOPs involvement in a range of cellular processes, using next-generation sequencing approaches assessing the effects of complete SPOP loss on cell transcriptomes may reveal more pathways impacted by SPOP loss and further potential exploitable drug targets. Combining this information with the available public data may further enhance the understanding of the diverse roles SPOP plays in normal cell homeostasis.

In attempts to identify new potential druggable pathways using *in silico* next-generation sequencing analysis, work conducted within this chapter will look to satisfy the following aims.

i. Send isolated whole cell RNA from the CWR22Rv1-AR-EK-SPOP$^{-/-}$-C1 and CWR22Rv1-AR-EK-SPOP$^{-/-}$-C2 derivatives and parental CWR22Rv1-AR-EK-iCas9 cell lines for RNA sequencing analysis.
ii. Compare the results of the above analysis to relevant publicly available data whilst experimentally validating some of the results output from the analysis pipelines.

6.3 Materials and methods

6.3.1 RNA-Sequencing sample preparation

Sample preparation for RNAseq analysis was performed as outlined in Figure 6.1. In brief, CWR22Rv1-AR-EK-iCas9, CWR22Rv1-AR-EK-SPOP$^{-/-}$-C1 and CWR22Rv1-AR-EK-SPOP$^{-/-}$-C2 cells were seeded in 6-well cell culture plates (Corning) and left for 48 hours before RNA harvest (in addition to production of parallel protein lysates to validate biomarker expression). Additional plates were seeded for the CWR22Rv1-AR-EK-iCas9 and CWR22Rv1-AR-EK-SPOP$^{-/-}$-C1 lines which were exposed to 2 Gy IR at 42 hours, before being returned to incubate for a further 6 hours prior to RNA harvest. This was repeated twice again with separate passages of cells for a total of 3 biological replicates per treatment arm which were sent for RNAseq analysis.

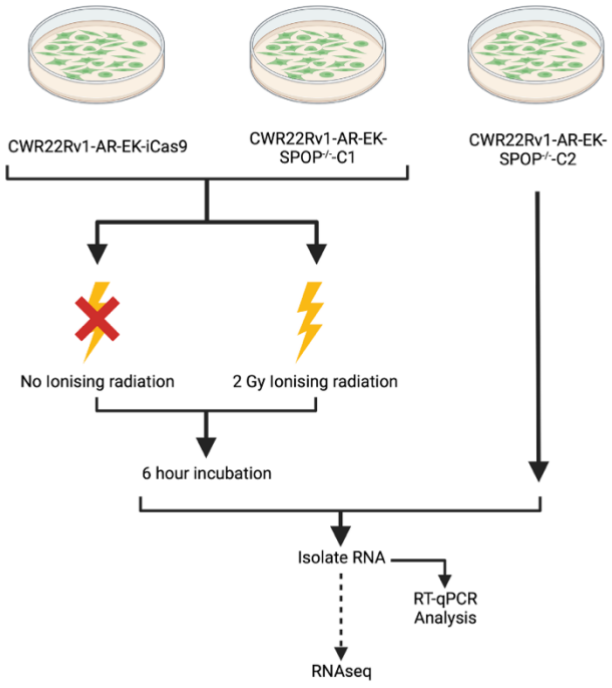


Figure 6.1 – Schematic for sample preparation before RNA sequencing analysis. *Whole cell RNA was extracted from CWR22Rv1-AR-EK-iCas9, CWR22Rv1-AR-EK-SPOP-/--C1 and CWR22Rv1-AR-EK-SPOP-/--C2 cells and sent for RNA sequencing. Populations of both CWR22Rv1-AR-EK-iCas9 and CWR22Rv1-AR-EK-SPOP-/--C1 cells were also exposed to 2 Gy IR, returned to incubate for 6 hours before RNA was extracted and sent for analysis.*

6.3.2 RNA-sequencing analysis

Samples sent for RNA-sequencing were generated as described above and sequenced using Illumina NovaSeq, 2x150bp configuration, with a data output of ~20 million raw paired-end reads per sample. Quality checks were performed using FastQC and MultiQC to ensure raw sequencing reads were suitable before use in downstream analysis (Andrews, 2010). Alignment to the human genome was performed using Salmon to the RefGenie Salmon index. Following alignment, principal component analysis (PCA), to display clustering of biological repeats, and DEseq2 to determine fold change (FC) and statistical p-values of differential gene expression between sample groups was performed in R. From the resultant total lists of differentially expressed genes (DEGs), lists of significantly differentially expressed genes (SDEGs) were extracted by filtering for only genes with a logFC of ±0.58 and an adjusted p-value of <0.05.

6.3.3 Gene-set enrichment analysis

DEG lists were ranked by logFC and input into the Broad institute Gene-set enrichment analysis (GSEA) tool (Version 4.3.2). Ranked lists were ran looking into enrichment in the Hallmark pathways and Kyoto Encyclopedia of Genes and Genome (KEGG) gene sets. Example of GSEA output plots can be seen in Figure 6.2

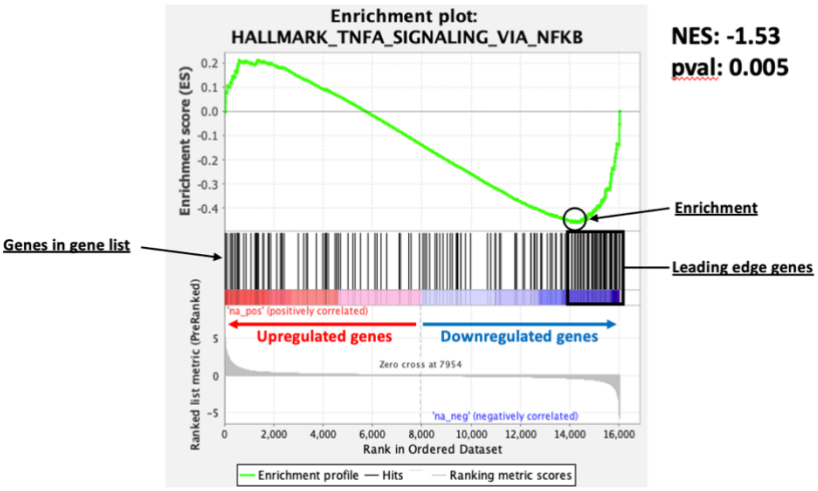


Figure 6.2 – Representation of GSEA output. *Output from GSEA for selected TNFA_SIGNALLING_VIA_NFKB hallmark displaying negative enrichment with leading edge highlighted.*

6.4. Results

6.4.1 Quality control

The FastQC and MultiQC tools were used to assess the quality of the RNA-sequencing data returned and to determine if any further processing of the data was necessary before downstream use [Figure 6.3] Initial FastQC was performed on the files returned for the untreated control CWR22Rv1-AR-EK-iCas9 samples, with samples passing all expected parameters tested apart from the adapter content [Figure 6.3A]. Having seen this across multiple samples, reads from all samples were trimmed using the cutAdapt tool, and then samples were run again through FastQC to check that all adaptor sequences had been removed from each sample [Figure 6.3B]. Following this, MultiQC was ran on all total samples, with samples passing (green) most of the parameters measured whilst failing (red) checks for sequence duplication and per base sequence content which is expected due to the presence of highly expressed genes from the cell lines sequenced [Figure 6.3C]. Also assessed were the mean quality scores for each sample and its individual replicates [Figure 6.3D]. Satisfied with samples passing the quality control checks performed, samples proceeded to downstream analysis.

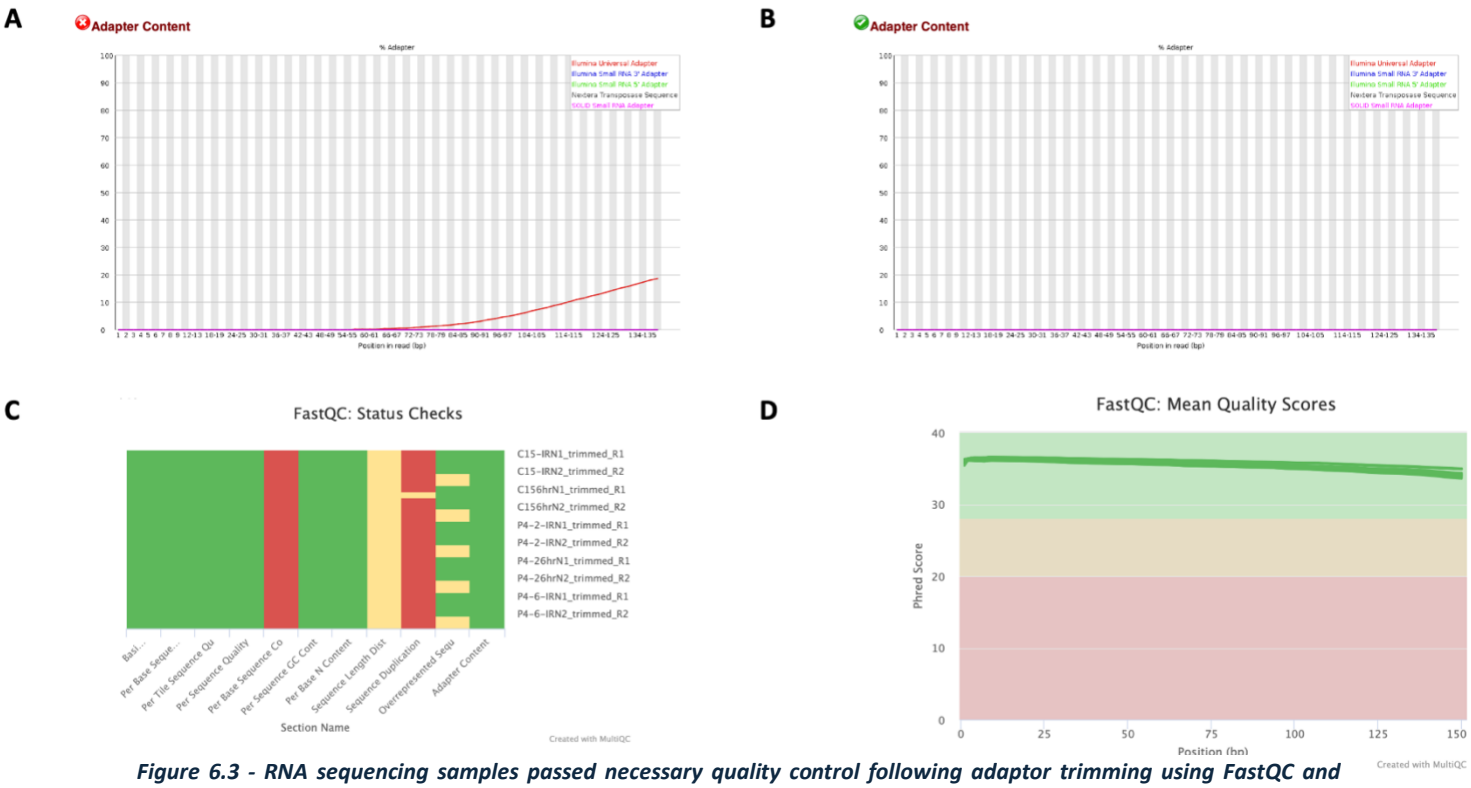


Figure 6.3 - RNA sequencing samples passed necessary quality control following adaptor trimming using FastQC and MultiQC tools. A. *Example of failed adaptor content quality check when putting samples through FastQC.* ***B.*** *Example adaptor content quality check following removal of adaptor sequences using cutAdapt tool.* ***C.*** *MultiQC summary for all samples submitted for all parameters assessed by FastQC.* ***D.*** *Fast QC Mean quality scores for all samples submitted for RNA sequencing analysis.*

6.4.2 Dispersion plots and Principal Component Analysis

Before running samples through DESeq2, checks were performed to assess the variation within the total data and if any replicates within a given sample were outliers. To assess the variation within the total data, a dispersion plot was generated for the total data [Figure 6.4A]. Here the data follows the expected relationship from the data with variance (dispersion) decreasing with increased read counts. Following this, the clustering of individual replicates within samples was assessed using principal component analysis [Figure 6.4B]. The plot shows clear clustering can be seen within the replicate samples of each of the experimental arms: CWR22Rv1-AR-EK-iCas9, CWR22Rv1-AR-EK-SPOP$^{-/-}$-C1 and CWR22Rv1-AR-EK-SPOP$^{-/-}$-C2; with the +IR samples for CWR22Rv1-AR-EK-iCas9 and CWR22Rv1-AR-EK-SPOP$^{-/-}$-C1 lines clustering together with the -IR samples for each. This indicates the individual repeats for each sample are consistent with one-another and distinct expression profiles between all the cell lines tested. Together, these provide great confidence in the results output by DESeq2 and subsequent lists of DEGs and SDEGs generated.

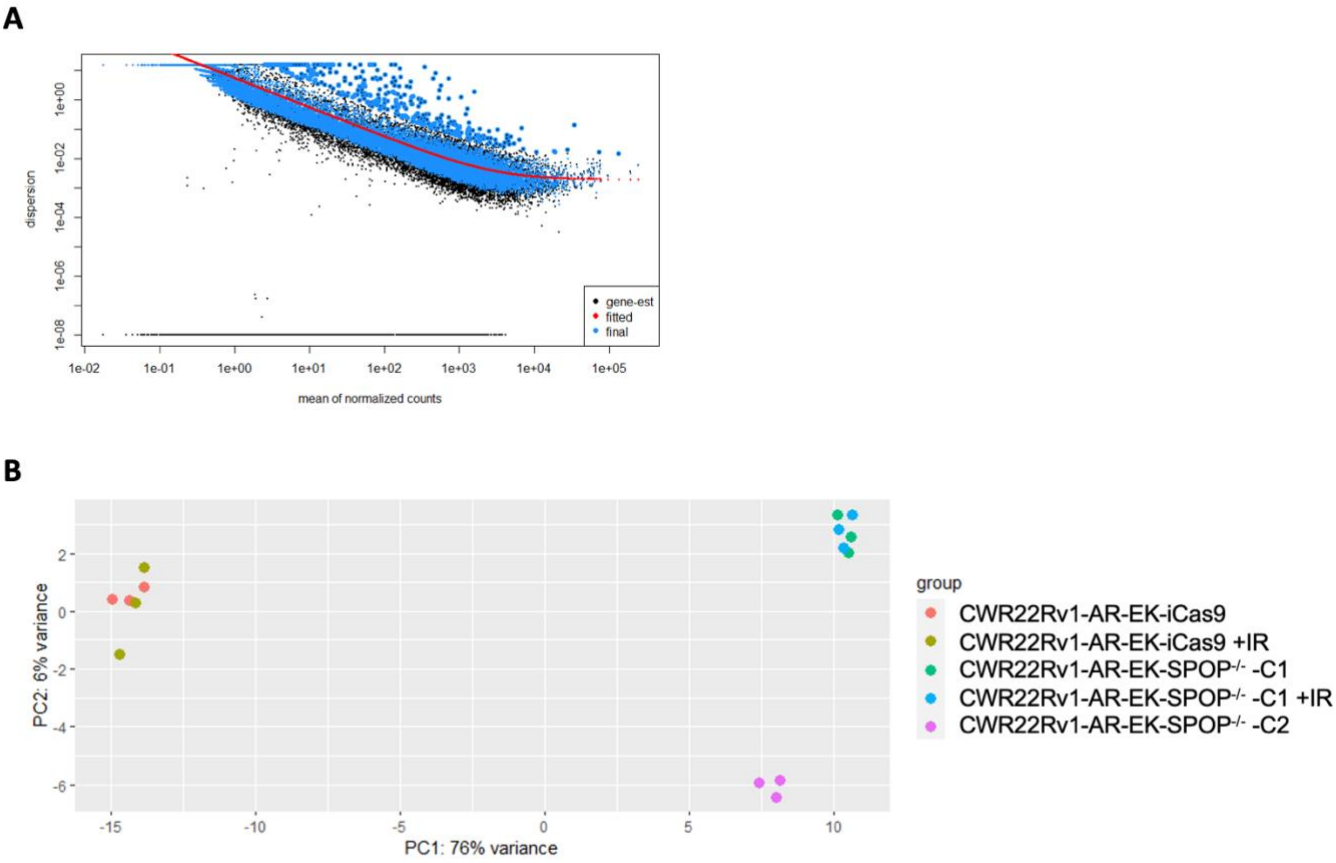


Figure 6.4 – Dispersion plots and principal component analysis for all submitted samples. A. *Dispersion plot for all samples submitted for RNA sequencing analysis showing final gene set conforming the predicated model.* ***B.*** *Following alignment to genome and quantification, PCA plots generated in R., with clustering between experimental repeats was generated using the gene count matrix.*

6.4.3 Differential gene expression using DESeq2

The R package DESeq2 was used to compare different samples to produce lists of DEGs. The total DEG lists were filtered to include genes with a $\log_2$FC of $\pm$ 0.5 and an adjusted p-value of <0.05 to produce the lists of SDEGs. Initially, both the CWR22Rv1-AR-EK-SPOP$^{-/-}$-C1 and the CWR22Rv1-AR-EK-SPOP$^{-/-}$-C2 clones were compared to the CWR22Rv1-AR-EK-iCas9 parental control (Table 6.1). Here, 2101 and 1901 SDEGs were returned for each SPOP knock-out clone respectively, representing those genes altered as a result of the loss of SPOP within this cellular background. A list of total SDEGs were used in creating a representative heatmap for all individual replicates using the normalised raw counts from the RNAseq runs [Figure 6.5]. Clear clustering can be observed within the total 2551 SDEGs (total SDEGs from Figure 6.7A), representing the effects of loss of SPOP in both SPOP$^{-/-}$ clones which demonstrate altered gene expression when compared to the parental control samples. Importantly, both SPOP$^{-/-}$-C1 and -C2 clones showed similar gene expression profiles when compared to the parental controls

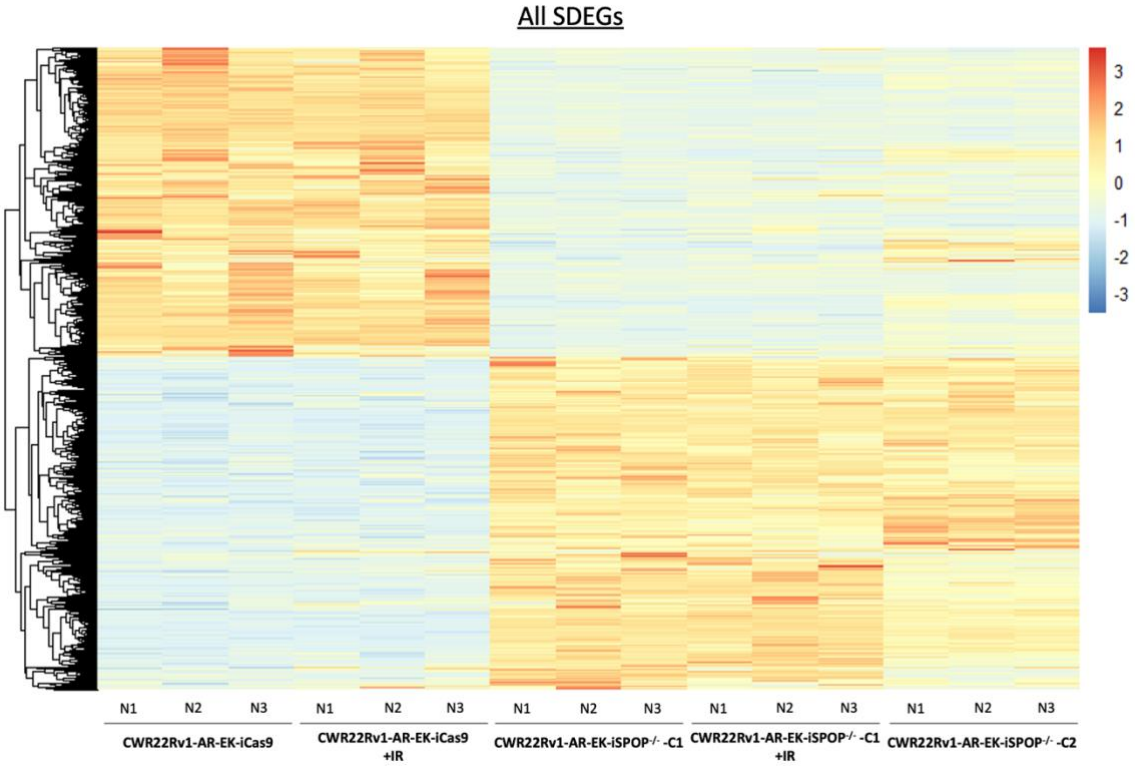


Figure 6.5 – CWR22Rv1-AR-EK-SPOP-/- clones show distinct transcriptomes when compared to parental CWR22Rv1-AR-EK-iCas9 cells. *Heatmap of normalised raw counts (counts-per-million) for a list of all significantly differentially expressed genes generated within R for all samples submitted.*

suggesting commonality in transcriptomic impact of SPOP knockout in both clonal derivatives whilst also giving insight into the distinct transcriptional profiles between the two knock-out clones [Figure 6.7A]. Further visualisations of differential gene expression were made between the experimental arms using MA plots in Figure 6.6, representing the total DEG and SDEGs returned from each comparison.

Table 6.1 – *Numbers of DEG and SDEG from SPOP$^{-/-}$ clones compared to parental controls*.

Sample	Number of DEG	Number of SDEG
CWR22Rv1-AR-EK-Cas9-SPOP$^{-/-}$-C1	17515	2101
CWR22Rv1-AR-EK-Cas9-SPOP$^{-/-}$-C2	17036	1950

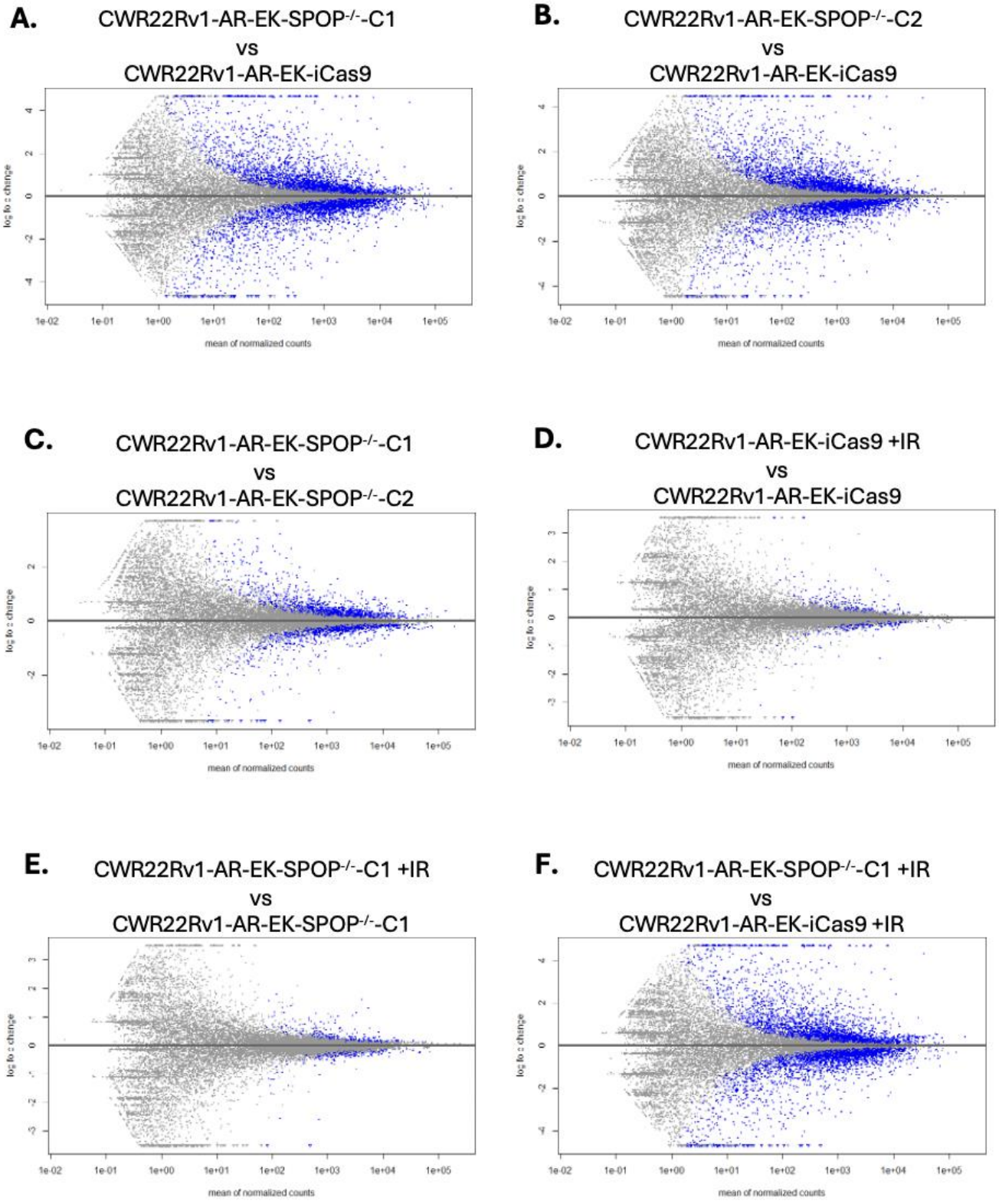


Figure 6.6 - MA plots to show proportion of significantly up and down regulated genes for all caparisons made between samples. *MA plots from DEGs and SDEGs were made for each comparison made between samples. Gray points indicate total DEGs up/downregulated. Bule points indicate up/downregulated SDEGs that have Log2FC of ± 1.5 and p value <0.05.*

6.4.4 Assessing transcriptomic variation between CWR22Rv1-AR-EK- SPOP$^{-/-}$-C1 and CWR22Rv1-AR-EK- SPOP$^{-/-}$-C2 derivatives

Although both clonal populations harbour *SPOP* knock-out, it can be seen from Figure 6.5 that the effects on gene expression differ slightly. To further interrogate the differences between the two clones, the SDEG lists from Table 6.1 were compared to determine overlap between the two lists [Figure 6.7A]. Interestingly, 660 (25.8%) and 509 (20%) SDEG were selectively altered in the CWR22Rv1-AR-EK-SPOP$^{-/-}$-C1 and -C2 derivatives, respectively; with 1382 (54.2%) SDEGs shared between the two. To further investigate the differences between these two clones, the full DEG lists from the comparison between CWR22Rv1-AR-EK-SPOP$^{-/-}$-C2 VS CWR22Rv1-AR-EK-SPOP$^{-/-}$-C1 were taken and subject to GSEA to identify alterations in *Hallmark* pathways [Figure 6.7B]. Only positively enriched pathways were returned from GSEA, with the highest enriched pathway being *pancreas beta cells*, *angiogenesis* and *hypoxia* with a normalised enrichment score (NES) of 0.8, 0.7 and 0.47, respectively.

6.4.5 Transcriptome effects of SPOP knock-out within CWR22Rv1-AR-EK cells.

To assess the effects of loss of SPOP function within the CWR22Rv1-AR-EK cell derivatives, the complete DEG lists from each clone compared to the parental line was subject to GSEA using the *Hallmark* lists for significantly enriched gene sets with the overlaps between the conditions and NES of each pathway presented in Figure 6.8. Interestingly a large disparity was seen in the number of positively enriched hallmarks between the two clones, although both shared the Hedgehog signalling hallmark with an NES of 1.54 and 1.6 for CWR22Rv1-AR-EK-SPOP$^{-/-}$-

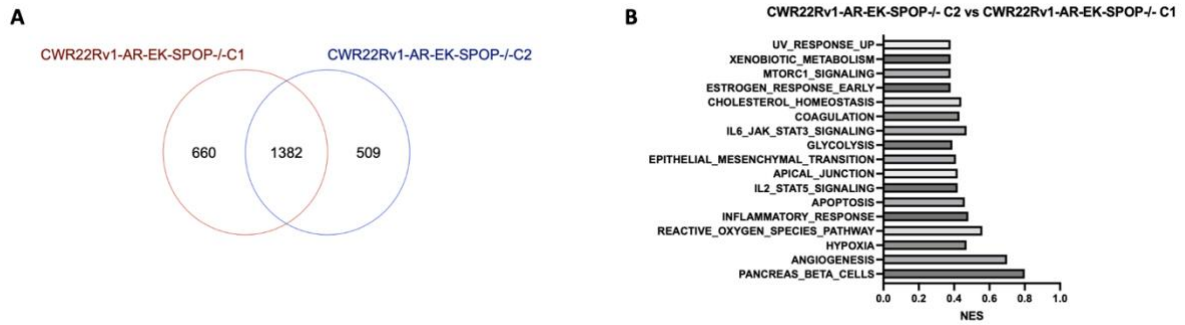


Figure 6.7 – Comparisons between CWR22Rv1-AR-EK-SPOP-/--C1 and CWR22Rv1-AR-EK-SPOP-/--C2. A. *Total SDEGs lists from both CWR22Rv1-AR-EK-SPOP-/- clones compared to the CWR22Rv1-AR-EK-iCas9 parental line taken and compared against each other to assess transcriptome differences between the two clones.* ***B.*** *GSEA analysis on ranked DEG lists from comparisons between CWR22Rv1-AR-EK-SPOP-/--C1 and CWR22Rv1-AR-EK-SPOP-/--C2 lines ran for Hallmark pathways with p value <0.25.*

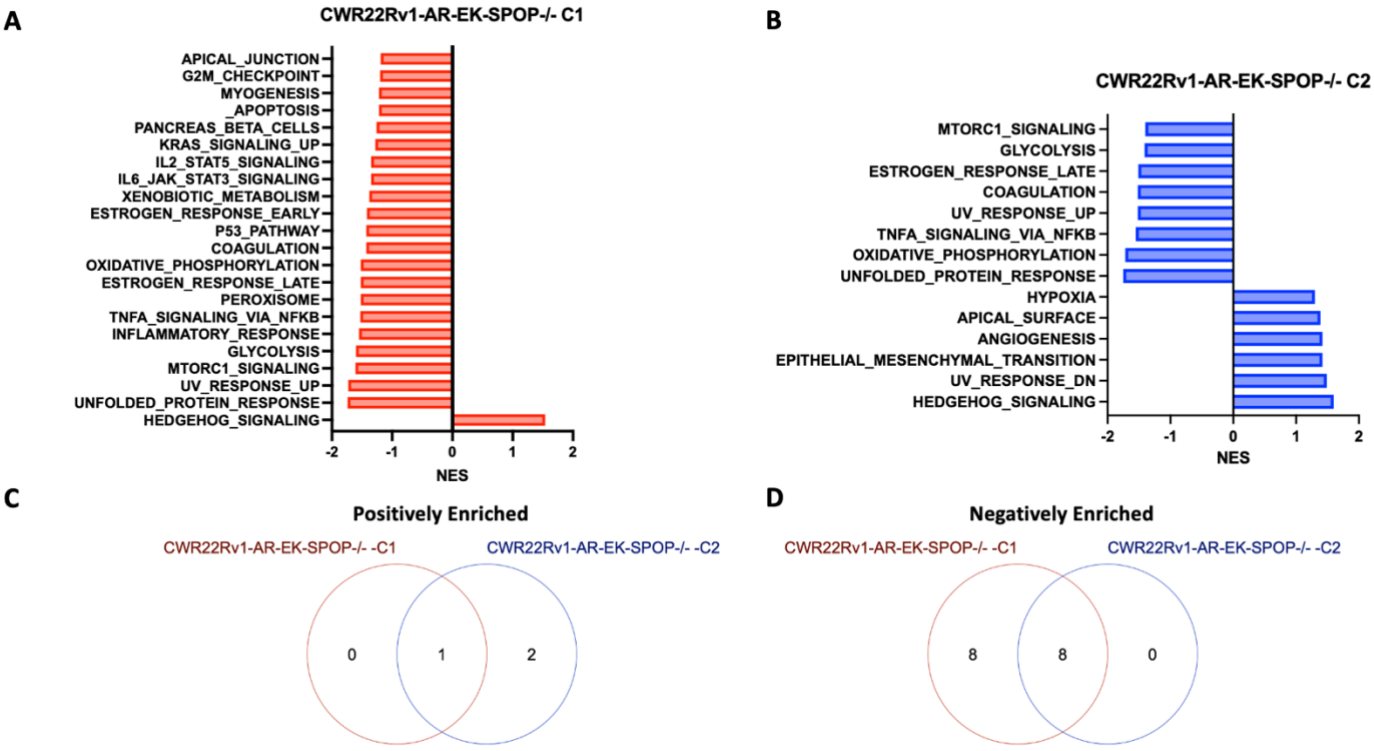


Figure 6.8 - CWR22Rv1-AR-EK-SPOP-/--C1 and CWR22Rv1-AR-EK-SPOP-/--C2 share positively and negatively enriched Hallmark pathways. *GSEA analysis on ranked DEG lists from comparisons between CWR22Rv1-AR-EK-SPOP-/--C1 and CWR22Rv1-AR-EK-iCas9 (**A**) CWR22Rv1-AR-EK-SPOP-/--C2 and CWR22Rv1-AR-EK-iCas9 (**B**) lines ran for Hallmark pathways with p value <0.25. Positively (**C**) and negatively (**D**) enriched Hallmark pathway lists from CWR22Rv1-AR-EK-SPOP-/--C1 and CWR22Rv1-AR-EK-SPOP-/--C1 taken and compared against each other for overlaps in Hallmark pathways between the two clones when compared to the parental line.*

C1 and CWR22Rv1-AR-EK-SPOP$^{-/-}$-C2, respectively. Furthermore, CWR22Rv1-AR-EK-SPOP$^{-/-}$-C2 displayed 5 more positively enriched hallmarks when compared to CWR22Rv1-AR-EK-SPOP$^{-/-}$-C1. With regards to the negatively enriched hallmarks, CWR22Rv1-AR-EK-SPOP$^{-/-}$-C1 and CWR22Rv1-AR-EK-SPOP$^{-/-}$-C2 returned 21 and 8 hallmarks, respectively; the latter demonstrating a total overlap with CWR22Rv1-AR-EK-SPOP$^{-/-}$-C1.

For more in-depth pathway analysis, the same lists were subject to GSEA using the *KEGG* pathways [Figure 6.9]. As observed previously, only a small number of pathways were seen to be positively enriched, although both shared a significant upregulation of nitrogen metabolism, with NES values of 1.83 and 1.9 for CWR22Rv1-AR-EK-SPOP$^{-/-}$-C1 and CWR22Rv1-AR-EK-SPOP$^{-/-}$-C2, respectively when compared to the parental samples. Comparing the negatively enriched pathways between the two clones, there is considerable disparity between the pathways being downregulated in each clone, contrasting with that seen in the *Hallmark* analysis. Only 6 significantly negatively enriched KEGG pathways were shared

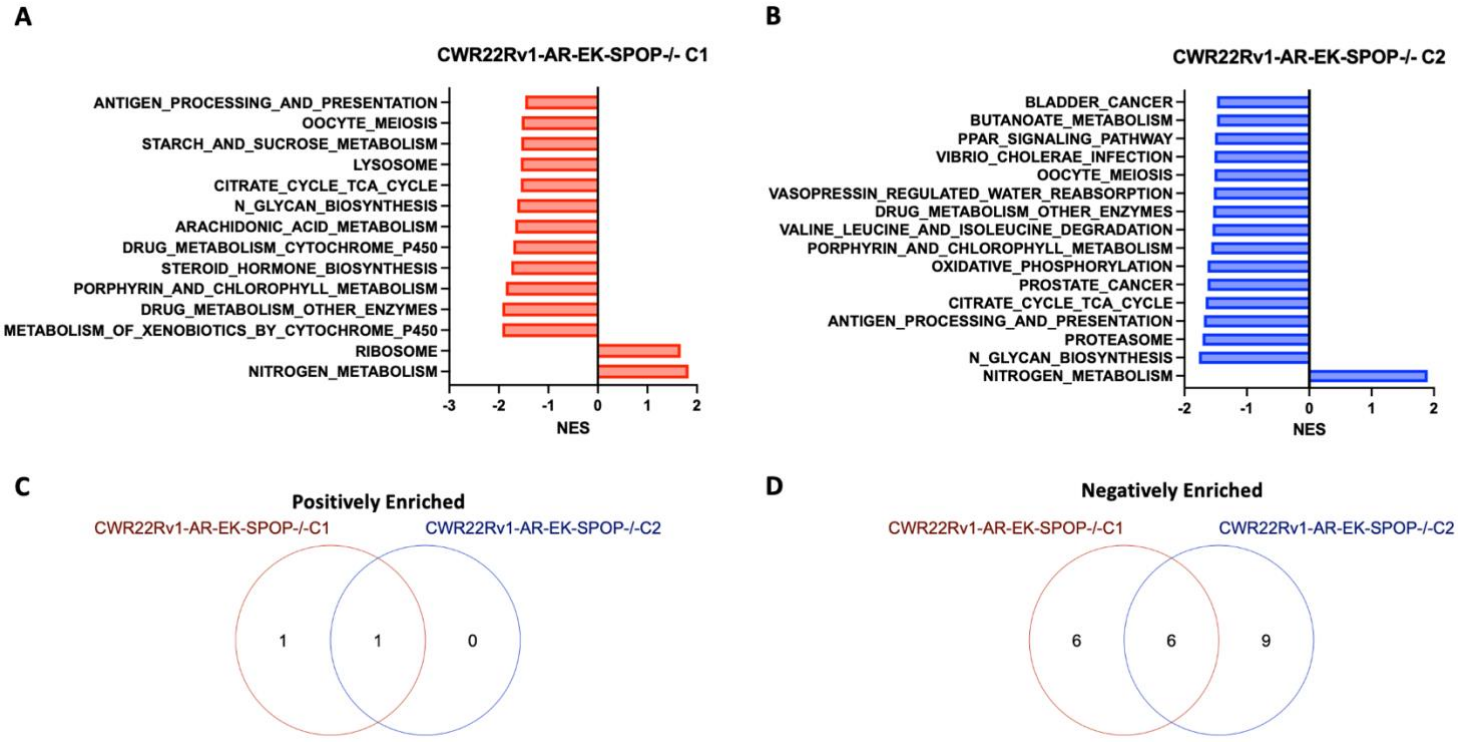


Figure 6.9 - CWR22Rv1-AR-EK-SPOP-/--C1 and CWR22Rv1-AR-EK-SPOP-/--C2 share positively and negatively enriched KEGG pathways. *GSEA analysis on ranked DEG lists from comparisons between CWR22Rv1-AR-EK-SPOP-/--C1 and CWR22Rv1-AR-EK-iCas9 (**A**) CWR22Rv1-AR-EK-SPOP-/--C2 and CWR22Rv1-AR-EK-iCas9 (**B**) lines ran for KEGG pathways with p value <0.25. Positively (**C**) and negatively (**D**) enriched KEGG pathway lists from CWR22Rv1-AR-EK-SPOP-/--C1 and CWR22Rv1-AR-EK-SPOP-/--C1 taken and compared against each other for overlaps in KEGG pathways between the two clones when compared to the parental line.*

between the two clones, with C1 having 6 distinctive pathways and C2 having 9. Taken together, these results show that although both clones harbour *SPOP* knock-out, the resultant transcriptomic changes within the lines vary; the reason/s for this remaining undefined. That said, having both clones provides a useful tool for downstream experimental set-up; where hypotheses can be tested in one clone and further validated in the second to add further weight to the results observed.

In attempts to see if *in silico* analysis of RNAseq data and resultant GSEA hallmarks were representative of phenotypic effects in cells, scratch wound assays were conducted to validate the results from previous RNAseq analysis [Figure 6.10]. Looking across the total GSEA outputs for comparisons made with each of the SPOP$^{-/-}$ clones vs the parental controls, both SPOP$^{-/-}$ clones displayed positive enrichment in the epithelial mesenchymal transition (EMT)

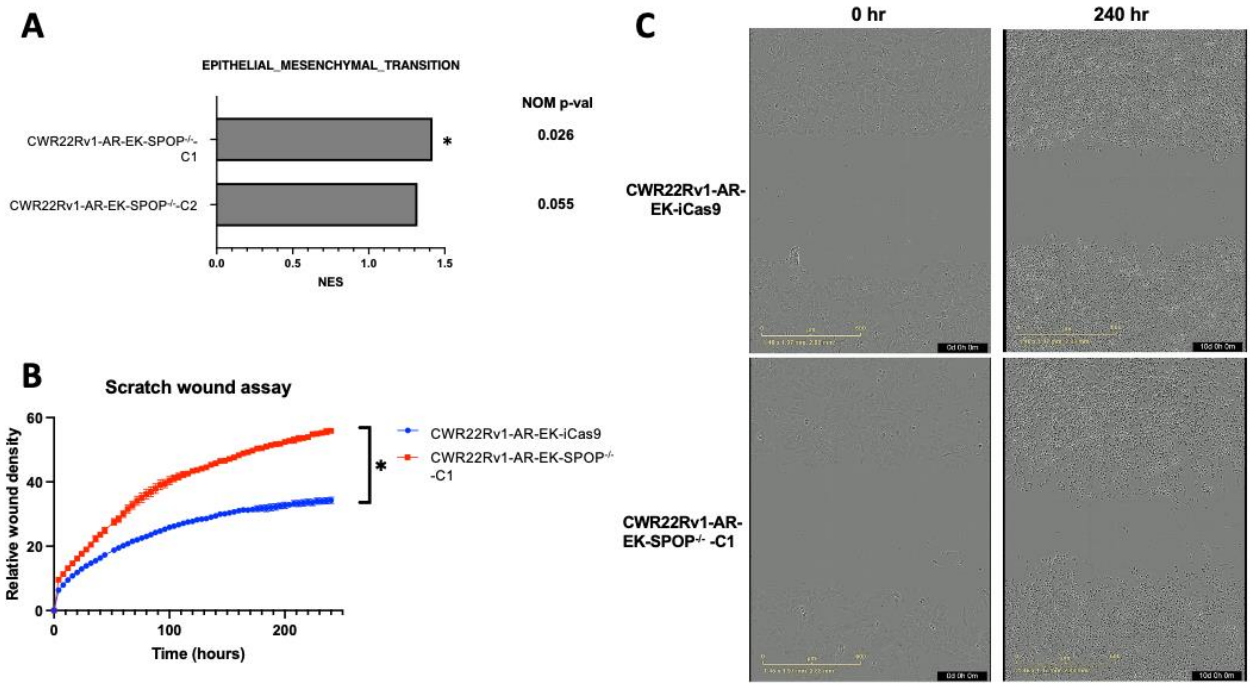


Figure 6.10 - Positive enrichment of EMT indicated form CWR22Rv1-AR-EK-SPOP$^{-/-}$ clone RNAseq data can be experimentally recapitulated. A. *GSEA hallmark analysis results of EMT hallmark from CWR22Rv1-AR-EK-SPOP$^{-/-}$-C1 and –C2 vs CWR22Rv1-AR-EK-iCas9 comparisons. * denotes NOM p-val <0.05.* ***B.*** *Relative wound density from 250-hour time course scratch wound assay between CWR22Rv1-AR-EK-iCas9 and CWR22Rv1-AR-EK-SPOP$^{-/-}$-C1 cells. Data normalised to 0-hour timepoint for each cell line and is representative of 3 individual experiments ± SEM. Statistical significance was determined by a One-way ANOVA with Dunnett's multiple comparison test (*, **, *** and **** denote P-values <0.0332, <0.0021, <0.0002 and <0.0001 respectively, ns denotes not significant).* ***C.*** *Representative images of scratch wound assay at 0 hours and 250 hours for CWR22Rv1-AR-EK-iCas9 and CWR22Rv1-AR-EK-SPOP$^{-/-}$-C1 cells. Data is representative of 3 individual experiments.*

phenotype with SPOP$^{-/-}$-C1 being statistically significant [Figure 6.10A]. As a result, a scratch wound assay was conducted both in the CWR22Rv1-AR-EK-iCas9 cell line and CWR22Rv1-AR-EK-SPOP$^{-/-}$-C1 cell line derivative [Figure 6.10B&C]. Across the total time course, the CWR22Rv1-AR-EK-SPOP$^{-/-}$-C1 showed significantly increased relative wound density, closing the initial scratch faster than the CWR22Rv1-AR-EK-iCas9 cell line. With this hallmark pathway experimentally validated, greater confidence can be had in the findings from the RNAseq analysis and future experimental results to mirror that from the *in silico* analysis.

6.4.6. Changes in response to IR as a result of SPOP knock-out within CWR22Rv1-AR-EK cells

Samples from the CWR22Rv1-AR-EK-iCas9 and CWR22Rv1-AR-EK-SPOP$^{-/-}$-C1 lines were exposed to 2 Gy IR, left to incubate for 6 hours before being harvested. These samples were run through DESeq2 and differential gene expression assessed for each arm and compared to their own representative unirradiated samples. A further comparison was made between the

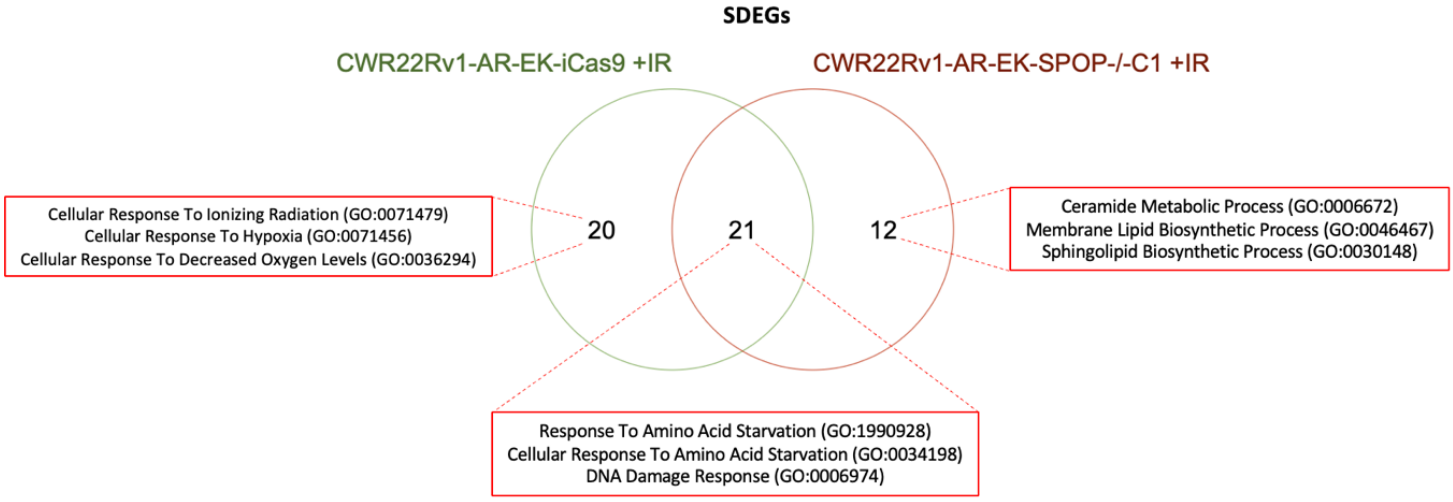


Figure 6.11 – Gene ontology analysis from SDEG lists for IR response in CWR22Rv1-AR-EK-iCas9 and CWR22Rv1-AR-EK-SPOP-/--C1 lines. *SDEG lists from the CWR22Rv1-AR-EK-iCas9+IR vs CWR22Rv1-AR-EK-iCas9 and CWR22Rv1-AR-EK-SPOP-/--C1+IR vs CWR22Rv1-AR-EK-SPOP-/--C1 comparisons underwent gene ontology analysis before GO terms for each comparison were taken, with individual and shared GO terms assessed. Top 3 GO terms from each list are highlighted.*

IR-exposed CWR22Rv1-AR-EK-SPOP$^{-/-}$-C1 and IR-exposed parental control (Table 6.2). The total DEGs returned from the IR-treated samples compared against their untreated counterparts were considerably lower than the number returned from all of the other comparisons made, including comparisons between the irradiated SPOP$^{-/-}$ and the irradiated parental. Further, the number of SDEGs returned from these irradiated comparisons were drastically lower than that observed from other comparisons, with the parental CWR22Rv1-AR-EK-iCas9 +IR vs -IR returning only 40 SDEGs and CWR22Rv1-AR-EK-SPOP$^{-/-}$-C1 +IR vs CWR22Rv1-AR-EK-SPOP$^{-/-}$-C1 -IR returning just 33 SDEGs. Although these lists are limited, there was still SDEGs returned; both shared between and exclusive to each list which were subsequently input to the ENRICHR tool to assess the associated gene ontology (GO) terms for biological processes. The top 3 terms are listed in Figure 6.11. The shared genes from both – IR and + IR treatment arms when ran through GO for biological processes returned *response to amino acid starvation* alongside the *DNA damage response*, both expected to be seen in response to external stress such as IR exposure. GO terms returned from the parental control +IR treatment gene list included

cellular responses to IR and *hypoxia*. Those genes exclusive to the CWR22Rv1-AR-EK-SPOP$^{-/-}$-C1 +IR treatment returned predominantly metabolic based GO terms.

As before, the full DEG lists for each comparison were compiled and ran through GSEA looking specifically into *Hallmark* pathways [Figure 6.12]. In response to the IR treatment across the two cell lines, both resulted in a higher number of positively enriched pathways than those negatively enriched. Both the CWR22Rv1-AR-EK-iCas9 +IR and CWR22Rv1-AR-EK-SPOP$^{-/-}$-C1 +IR shared a majority of positively enriched *Hallmarks*, including the *p53 pathway,* which saw a NES of 2.76 and 3.06, respectively. *Hallmarks* exclusive to the CWR22Rv1-AR-EK-iCas9 +IR treatment arm included *epithelial-mesenchymal transition* with an NES of 1.54. The CWR22Rv1-AR-EK-SPOP$^{-/-}$-C1 +IR yielded individual *Hallmarks* in *oxidative phosphorylation*, *PI3K-AKT-MTOR signalling* and *myogenesis*, returning NES of 2.29, 1.56 and 1.48, respectively. Both treatments returned no shared overlaps in negatively enriched hallmarks: with the CWR22Rv1-AR-EK-iCas9 +IR treatment returning negative enrichment in the *unfolded protein response* and *Myc targets Hallmarks* whilst the CWR22Rv1-AR-EK-SPOP$^{-/-}$-C1 +IR returned only *Hedgehog signalling* as negatively enriched.

To further understand the differences between the two irradiated cell line derivatives, a final comparison was made using DESeq2 between the CWR22Rv1-AR-EK-SPOP$^{-/-}$-C1 +IR and CWR22Rv1-AR-EK-iCas9 +IR samples. The resultant DEG list was taken and ran through GSEA to reveal the differences in IR response between the CWR22Rv1-AR-EK-iCas9 and CWR22Rv1-AR-EK-SPOP$^{-/-}$-C1 lines [Figure 6.13]. Here, only *hedgehog signalling* was found to be positively enriched with an NES of 1.72; whereas multiple *Hallmarks* were found to be negatively enriched, including the *unfolded protein response*, *p53 pathway* and *G2M checkpoint* which saw NES scores of -1.51, -1.38 and -1.24, respectively.

Table 6.2 – *Number of DEGs and SDEGs from IR exposed samples*

Sample	Number of DEG	Number of SDEG (p val< 0.05, logFC >±0.5)
CWR22Rv1-AR-EK-iCas9 +IR **vs** CWR22Rv1-AR-EK-iCas9	13593	40
CWR22Rv1-AR-EK-Cas9-SPOP$^{-/-}$-C1 +IR **vs** CWR22Rv1-AR-EK-Cas9-SPOP$^{-/-}$-C1	12123	33
CWR22Rv1-AR-EK-Cas9-SPOP$^{-/-}$-C1 +IR **vs** CWR22Rv1-AR-EK-iCas9 +IR	17533	2107

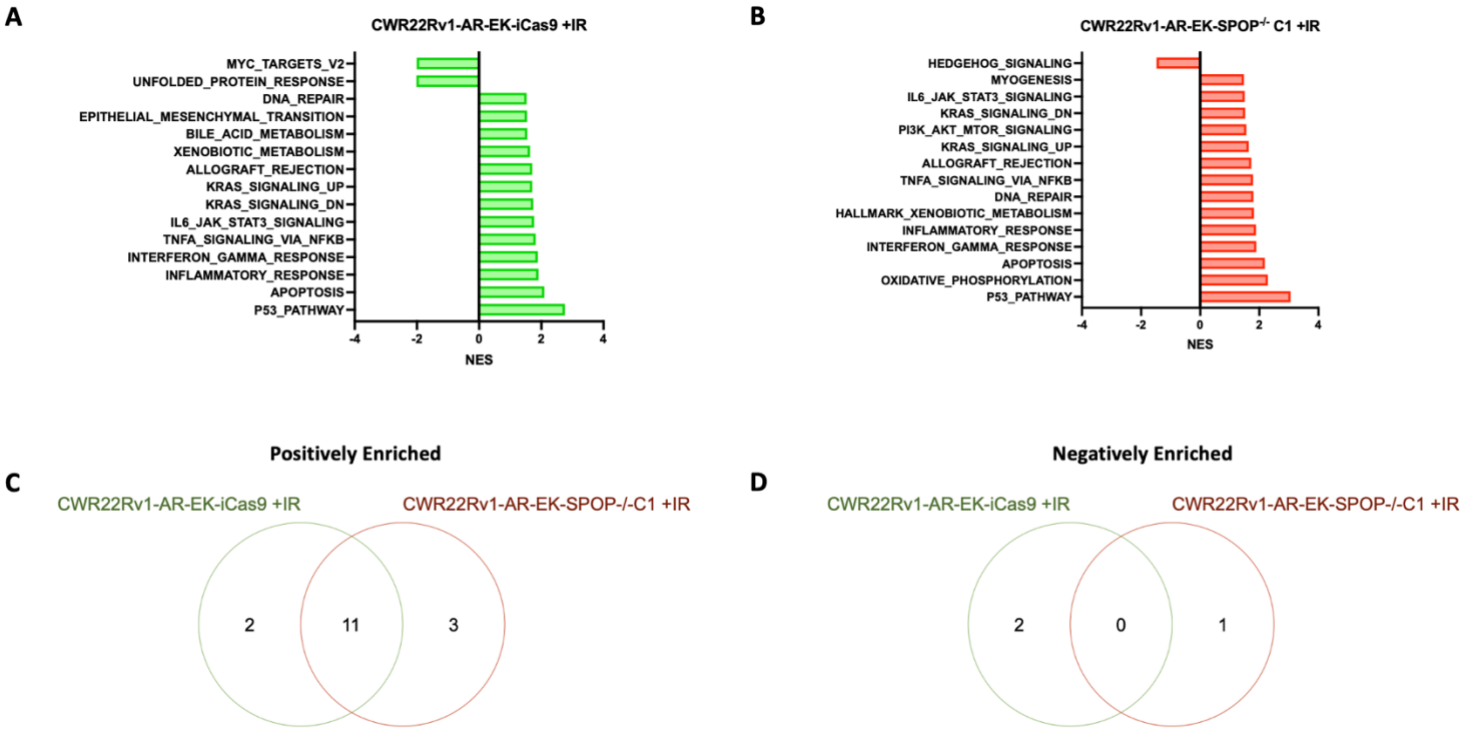


Figure 6.12 – Comparisons between CWR22Rv1-AR-EK-iCas9 and CWR22Rv1-AR-EK-SPOP-/--C1 positively and negatively enriched Hallmark pathways. *GSEA analysis on ranked DEG lists from comparisons between CWR22Rv1-AR-EK-iCas9+IR and CWR22Rv1-AR-EK-iCas9 (**A**) CWR22Rv1-AR-EK-SPOP-/--C1+IR and CWR22Rv1-AR-EK-SPOP-/--C1 (**B**) treatment arms for Hallmark pathways with p value <0.25. Positively (**C**) and negatively (**D**) enriched Hallmark pathway lists from above comparisons taken and compared against each other for overlaps in Hallmark pathways between the two +IR treatment arms when compared to the un-irradiated cells.*

Finally, to assess which of these genes and *Hallmark* pathways were exclusive to the +IR treatments in the CWR22Rv1-AR-EK-SPOP$^{-/-}$-C1 cell line, SDEG lists and *Hallmarks* returned from GSEA were investigated from DESeq2 comparisons between CWR22Rv1-AR-EK-SPOP$^{-/-}$-C1 +IR vs CWR22Rv1-AR-EK-iCas9 +IR and CWR22Rv1-AR-EK-SPOP$^{-/-}$-C1 -IR vs CWR22Rv1-AR-EK-iCas9 -IR samples [Figure 6.13]. When comparing the SDEG lists from the two comparisons, the effects of the IR treatments on the transcriptome of the SPOP$^{-/-}$ cells can be seen, with 494 SDEGs being exclusive to the +IR treated CWR22Rv1-AR-EK-SPOP$^{-/-}$-C1 populations. Both comparisons shared positive enrichment in the *hedgehog signalling Hallmark*, while all the negatively enriched *Hallmarks* observed in the -IR treatments were also observed in the +IR treated comparisons. 5 *Hallmarks* were seen exclusively negatively enriched in the +IR treated comparison, including *oxidative phosphorylation* with a NES of -1.26 and IL2-STAT5 signalling with NES of -1.29.

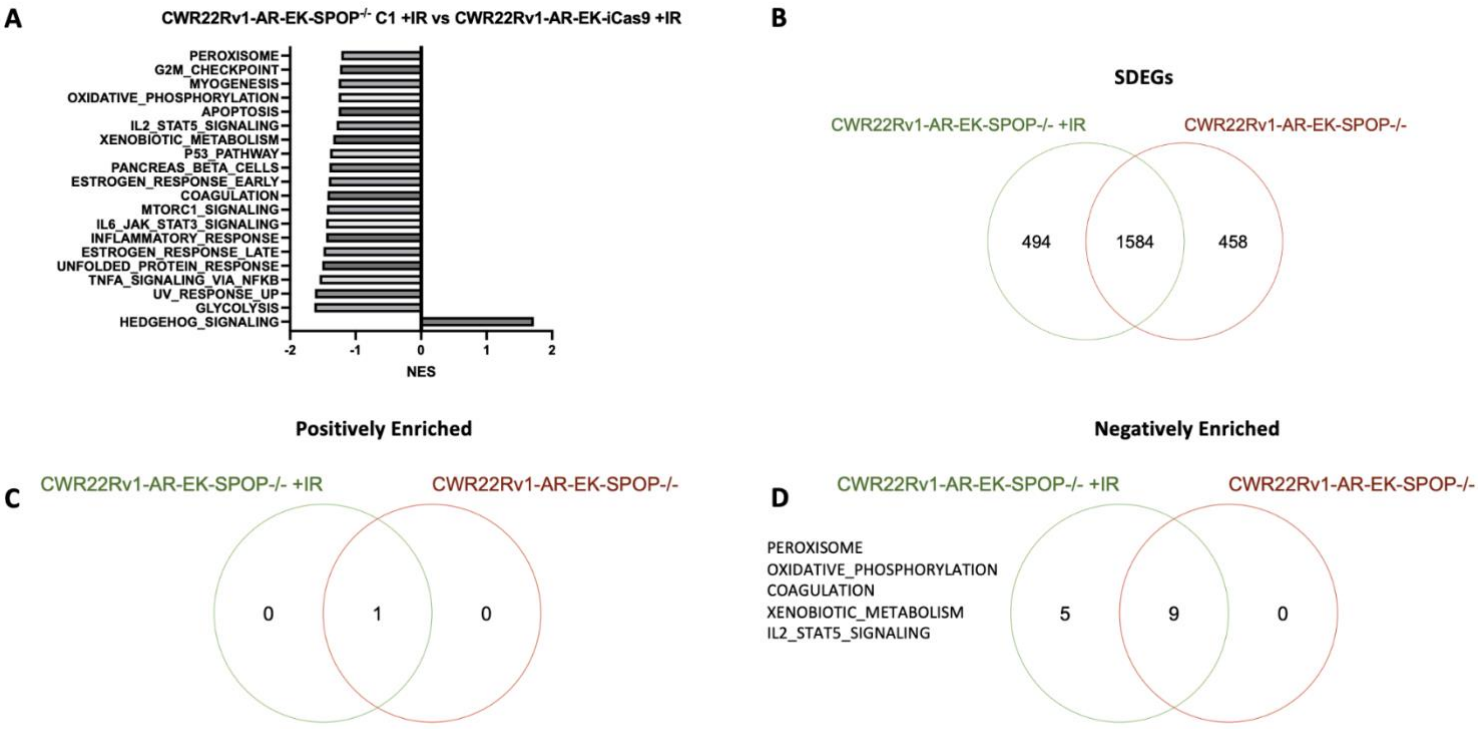


Figure 6.13 - Comparisons between the irradiation response in CWR22Rv1-AR-EK-iCas9 and CWR22Rv1-AR-EK-SPOP-/--C1 cells. A. *GSEA analysis on ranked DEG lists from comparisons between CWR22Rv1-AE-EK-SPOP-/--C1+IR and CWR22Rv1-AR-EK-iCas9+IR for Hallmark pathways with p value <0.25.* ***B.*** *Total SDEGs lists from CWR22Rv1-AE-EK-SPOP-/--C1+IR vs CWR22Rv1-AE-EK-SPOP-/--C1 comparison compared to the CWR22Rv1-AE-EK-SPOP-/--C1 vs CWR22Rv1-AR-EK-iCas9 comparison taken and overlaps in transcriptome assessed. Positively (**C**) and negatively (**D**) enriched Hallmark pathway lists from above comparisons taken and compared against each other for overlaps in Hallmark pathways from comparisons between CWR22Rv1-AE-EK-SPOP-/--C1+IR vs CWR22Rv1-AE-EK-SPOP-/--C1 and CWR22Rv1-AE-EK-SPOP-/--C1 vs CWR22Rv1-AR-EK-iCas9.*

6.4.7 Transcriptome analysis of TCGA patient data harbouring SPOP mutations conforms with trends seen in in-house SPOP$^{-/-}$ transcriptome data.

Using the TCGAbiolinks and EdgeR tools, DEG lists were extracted from the publicly available TCGA patient datasets comparing patients harbouring a *SPOP* mutation against those with wild-type *SPOP*. As before, the resultant lists were then run through GSEA to assess effects of *SPOP* mutations on the *Hallmark* pathways from patient data [Figure 6.14A]. *Hallmarks* found to be positively enriched included *myogenesis* (NES 2.67) and the *androgen response* (NES 1.72), whilst the *epithelial-mesenchymal transition* (NES -1.69) and *apoptosis* (NES -1.37) *Hallmarks* were found to be negatively enriched.

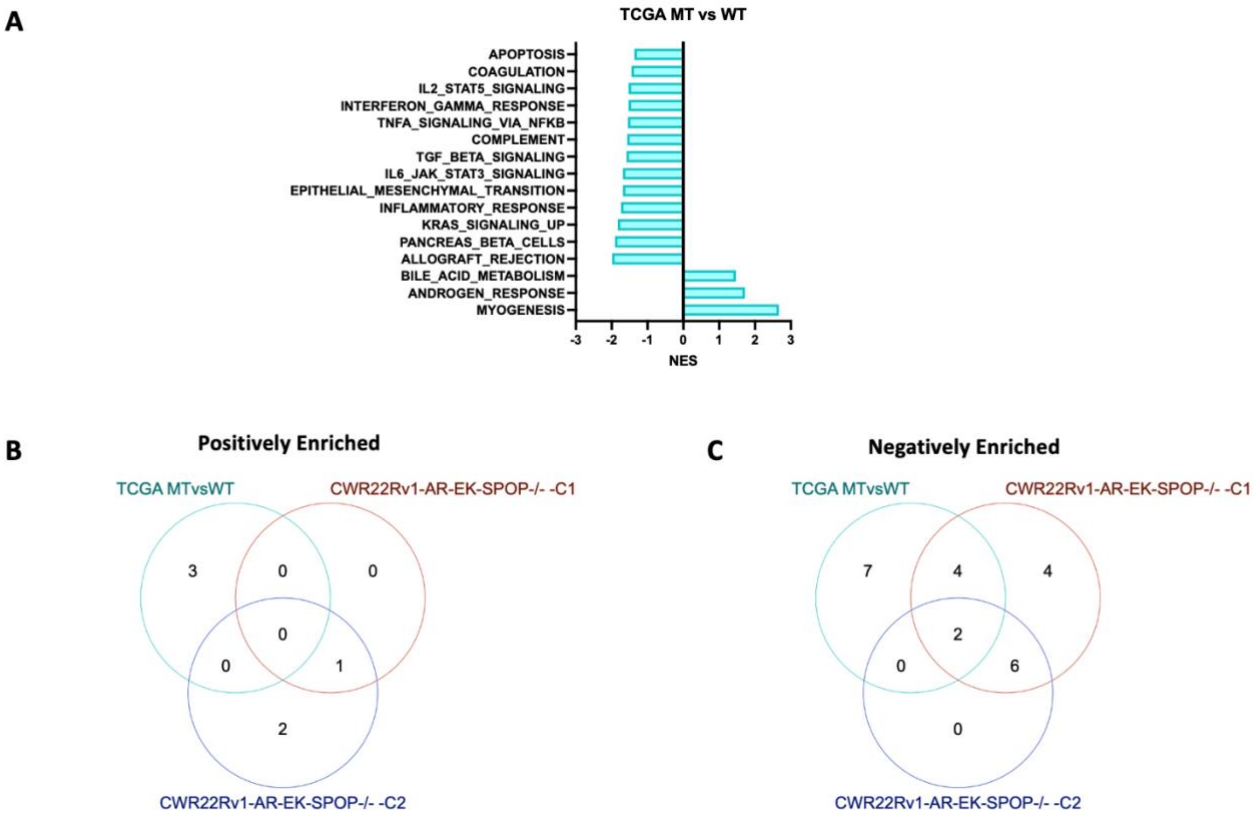


Figure 6.14 – Hallmark analysis from TCGA SPOP mutant vs wild-type patients and comparisons to RNA sequencing analysis from SPOP-/- clones. A*. GSEA analysis on ranked DEG lists from comparisons between CWR22Rv1-AE-EK-SPOP-/--C1+IR and CWR22Rv1-AR-EK-iCas9+IR for Hallmark pathways with p value <0.25. Positively (**B**) and negatively (**C**) enriched Hallmark pathway lists from TCGA MTvsWT compared against CWR22Rv1-AR-EK-SPOP-/--C1 vs CWR22Rv1-AR-EK-iCas9 AND CWR22Rv1-AR-EK-SPOP-/--C2 vsCWR22Rv1-AR-EK-iCas9 Hallmarks from RNAseq analysis.*

Finally, to investigate how the transcriptome effects of *SPOP* knock-out within the CWR22Rv1-AR-EK-iCas9 models generated in this study align with the transcriptome effects of *SPOP* mutations within patient samples, the retuned hallmarks for each comparison were compared against each-other [Figure 6.14B&C]. Here, none of the hallmarks found to be positively enriched in the TCGA dataset were found within the CWR22Rv1-AR-EK-SPOP$^{-/-}$-C1 and CWR22Rv1-AR-EK-SPOP$^{-/-}$-C2 SPOP$^{-/-}$ models. Interestingly, further examination revealed the negatively enriched *Hallmarks coagulation* and *TNFα signalling via NFκB* were shared between all three comparisons, with a further 4 *Hallmarks* shared between CWR22Rv1-AR-EK-SPOP$^{-/-}$-C1 and the TCGA dataset. Also of note, the TCGA MT vs WT dataset returned *epithelial-mesenchymal transition* as positively enriched, which was found to be negatively enriched in both SPOP$^{-/-}$ clones.

6.4.8. Transcriptome analysis of published RNAseq data aligns with trends seen in in-house SPOP$^{-/-}$ and TCGA transcriptome and data.

Of the publicly available datasets, data published in Geng et al. (2014) demonstrates overexpression of a range of SPOP mutants in the LNCaP-Abl PC cell background alongside expression of a control vector and SPOP^{WT} (Geng et al., 2014) [Figure 6.15]. Using UMAP plots for all the samples available from this dataset, clear clustering of experimental arms can be seen [Figure 6.15A]. Interestingly here samples expressing any of the SPOP mutants cluster together and overlap with one-another, giving indications that there will be similar transcriptomic changes seen within cells regardless of the SPOP mutation present. With the experimental arms present in this experimental set up, deciding what to use as a control may be a point of issue as overexpression of SPOP^{WT} will cause transcriptional changes alone. As such, overlaps in SDEGs from all SPOP mutants were compared against the vector control and SPOP^{WT} overexpression experimental arms [Figure 6.15B&C]. In both comparisons, the proportion of shared SDEGs between each mutant looked to remain relatively similar

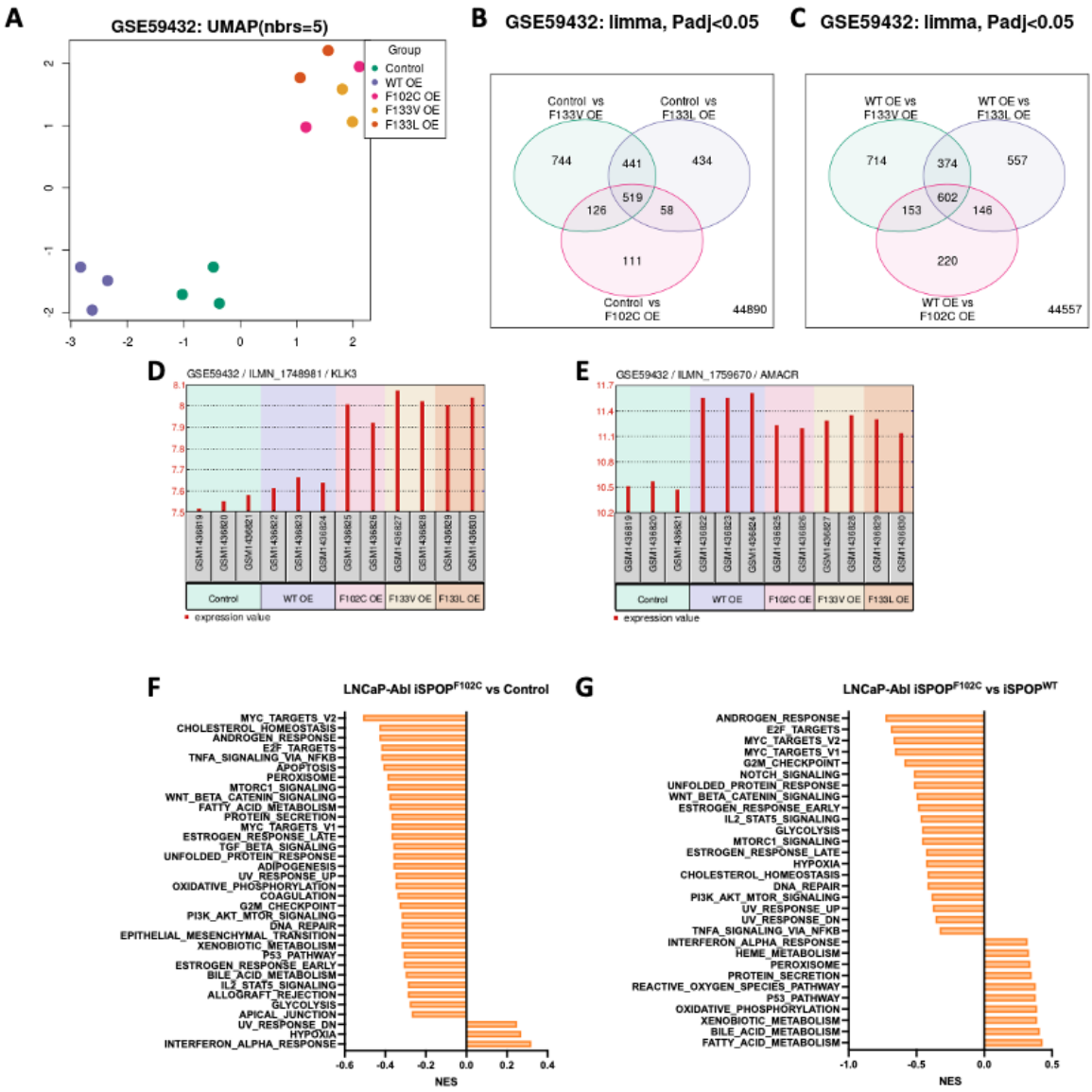


Figure 6.15 - Analysis of publicly available SPOP overexpression datasets reveals minimal difference across SPOP mutants. ***A.*** *UMAP plots for individual samples from control vector, SPOP^{WT}, SPOP^{F102C}, SPOP^{F133V} and SPOP^{F133L} overexpression in LNCaP-Abl cells. Vector control and SPOP^{WT} represent 3 individual replicates, SPOP$^{F102C/F133V/F133L}$ data represents 2 individual replicates. SDEG comparisons comparing overexpression of each individual SPOP mutant against vector control (**B**) and SPOP^{WT} overexpression (**C**). Plots of raw counts across all experimental repeats for KLK3 (PSA) (**D**) and AMACR (**E**). Significantly enriched hallmarks from GSEA analysis for SPOP^{F102C} overexpression from ranked DEG lists compared against vector control expression (**F**) and SPOP^{WT} expression (**G**).*

regardless of mutant expressed. Finally, using the raw counts for select genes, differences across the experimental arms was assessed. *KLK3* (PSA) presented itself as one of the genes demonstrating the highest alterations in expression across all datasets, and with it being an AR responsive gene and AR being an SPOP substrate targetable in the LNCaP-Abl cell line; where levels increase in response to overexpression of all SPOP mutants [Figure 6.15D]. This trend of altered expression in response to SPOP mutant overexpression was continued for other gene targets. However in some cases, such as *AMACR*, expression was elevated in response to overexpression of SPOP regardless of mutation status [Figure 6.15E]. With similar responses

seen across the range of SPOP mutants expressed within this study, alongside a similar overexpression RNAseq study performed in the LNCaP cell background only using this mutation, $SPOP^{F102C}$ was selected as a candidate mutant to be taken for further GSEA analysis [Figure 6.15F&G]. Comparing the upregulated pathways between both sets of hallmark analysis only the Interferon alpha response was shared, with a majority of negatively enriched pathways being shared including Androgen response, DNA repair and TNFa signalling via NFKB.

Data included in Burleson et al. (2022) looks at the transcriptional effects of overexpressing $SPOP^{F102C}$ and $SPOP^{WT}$ in LNCaP PC cells in the presence and absence of DHT (Burleson et al., 2022). To further understand the effects of SPOP mutant and wild type overexpression in PC cells, as before, GSEA hallmark analysis was performed using the cells stimulated with DHT comparing the effects of overexpression of $SPOP^{F102C}$ vs $SPOP^{WT}$ [Figure 6.16A]. Here, the androgen response saw negative enrichment, as expected, with an NES of -0.52, although a positive NES of 0.47 for the DNA repair hallmark. SPOP overexpressing cells stimulated with DHT allows for more representative comparisons to the previous analysis within the LNCaP-Abl cells in which the AR is active independent of its ligand. As such, positively and negatively enriched hallmarks from GSEA hallmark analysis were compared to show shared enriched pathways from both datasets [Figure 6.16B]. Across both the datasets, positive enrichment was seen in 5 hallmarks such as the p53 pathway while only 3 negatively enriched pathways, including the androgen response and TNFa signalling by NFKB were observed.

As before, RNAseq data was taken from work published in Mukhopadhyay et al. (2021) where CWR22Rv1 cells were transfected with SPOP-targeting or nonspecific siRNA (Mukhopadhyay et al., 2021). Data from this comparison would be of great interest as out of all the available published data, it most represents the SPOP knock-out RNAseq data analysed previously. To that end, GSEA hallmark analysis was conducted on this dataset [Figure 6.16C]. Of note, negative enrichment of the DNA repair hallmark and positive enrichment in the p53 pathway, epithelial mesenchymal transition and TNFa signalling by NFKB hallmarks was observed, with NES of 0.31, 0.35 and 0.44, respectively. In attempts to draw comparisons with the other publicly available data, utilising the dominant negative loss of function effects exerted on cells

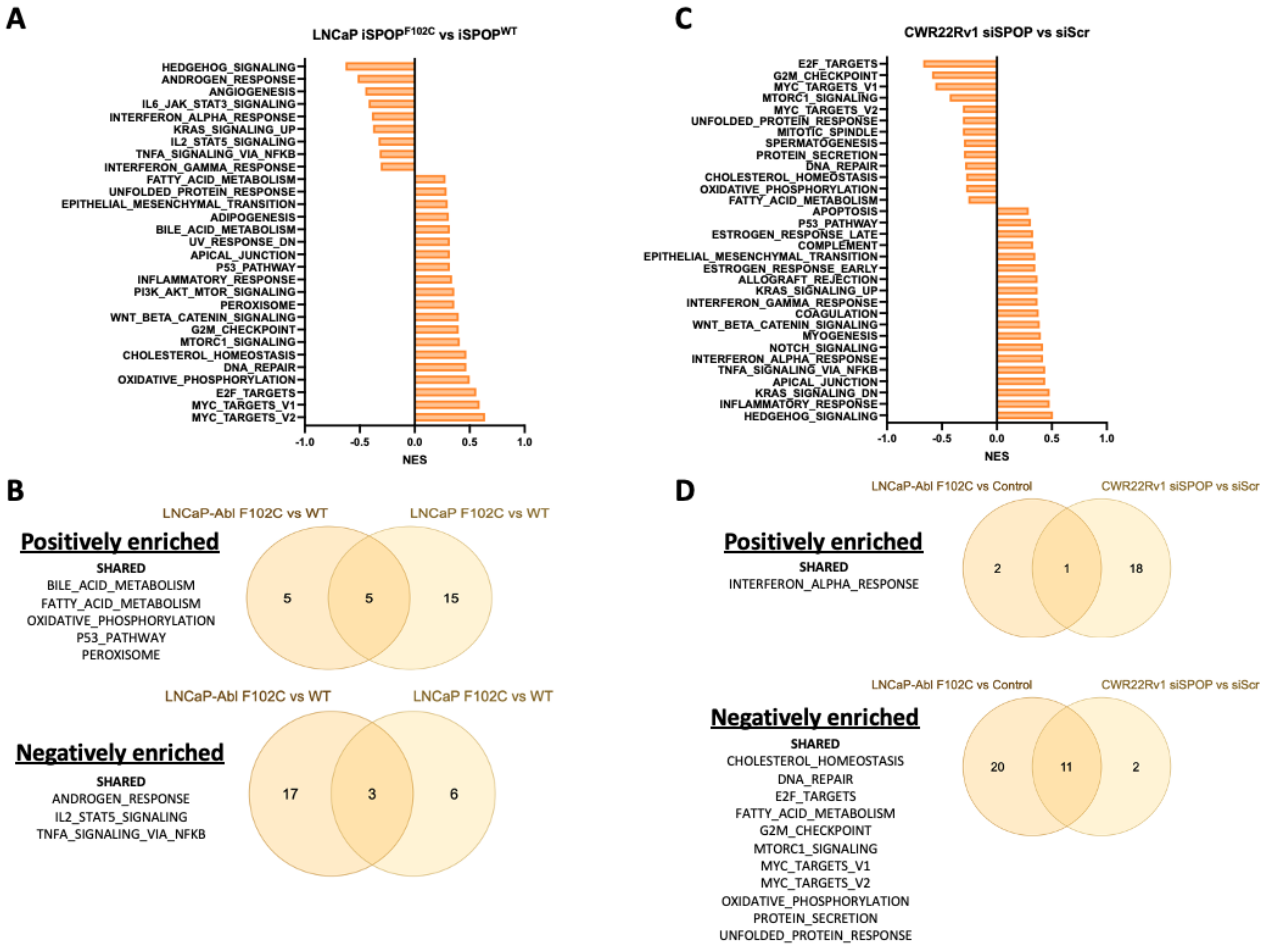


Figure 6.16 – GSEA hallmark analysis across published datasets show shared responses to SPOP manipulation. A. *Significantly enriched GSEA hallmarks for LNCaP cells overexpressing SPOP^{F102C} vs SPOP^{WT} from Burleson et al.2022* ***B.*** *Comparisons of hallmarks from GSEA analysis between SPOP^{F102C} vs SPOP^{WT} overexpression in LNCaP-Abl cells and LNCaP cells with shared hallmarks annotated.* ***C.*** *Significantly enriched GSEA hallmarks for CWR22Rv1 cells transfected with SPOP targeting siRNA vs scrambled siRNA from Mukhopadhyay et al. 2021* ***D.*** *Comparisons of hallmarks from GSEA analysis between SPOP^{F102C} vs control vector overexpression in LNCaP-Abl cells and CWR22Rv1 cells transfected with siSPOP vs siScr siRNA with shared hallmarks annotated.*

by SPOP mutant expression, GSEA hallmarks from the knockdown RNAseq was compared to hallmarks from the LNCaP-Abl SPOP^{F102C} vs control vector dataset [Figure 6.16D]. Minimal similarities were seen in from the positively enriched hallmarks, however negative enrichment in the DNA repair hallmark was shared across both datasets.

Finally, to potentially provide indications of any shared cellular pathways dysregulated across all datasets analysed within this chapter, comparisons were made using GSEA hallmarks from both the SPOP$^{-/-}$ clones and TGCA SPOP^{MT} and ran against each of the published datasets [Figure 6.17A-D]. Unfortunately, only minimal shared pathways between the published data were seen when looking at those positively enriched. Looking only at those negatively

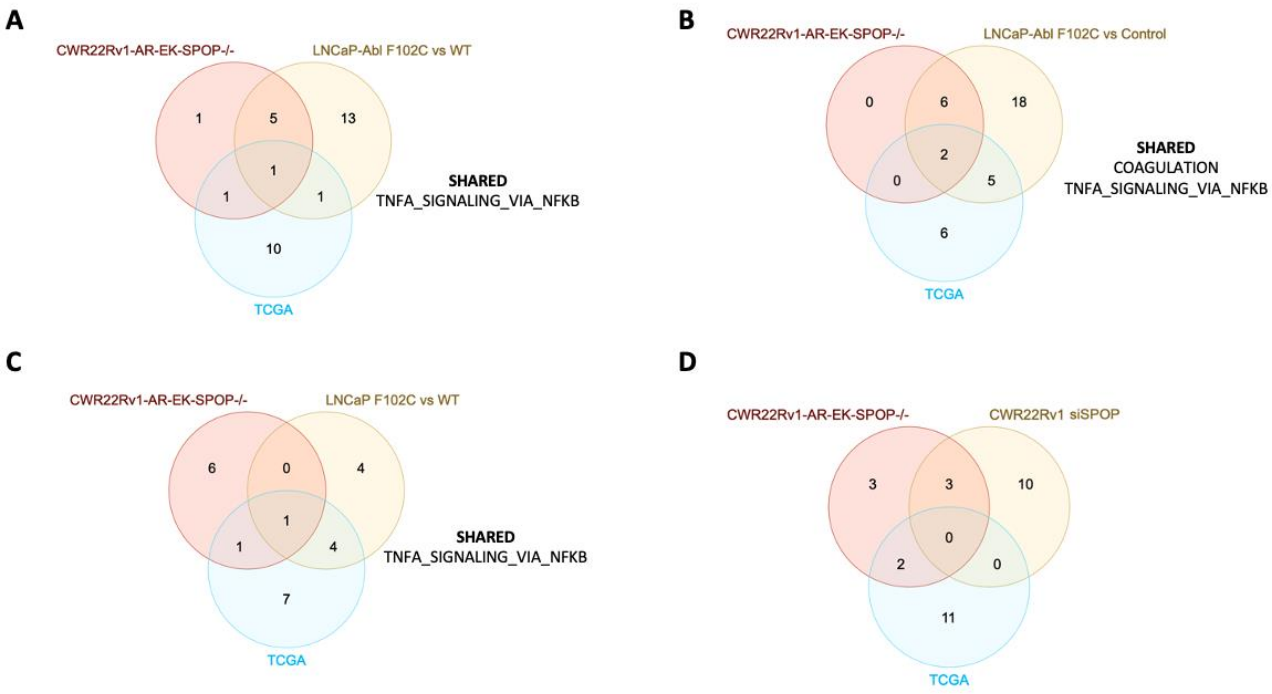


Figure 6.17 - Comparisons between publicly available data to CWR22Rv1-AR-EK-SPOP$^{-/-}$ and TCGA negatively enriched hallmarks. *GSEA hallmark comparisons between CWR22Rv1-AR-EK-SPOP$^{-/-}$ clones vs CWR22Rv1-AR-EK-iCas9, TCGA SPOP^{MT} vs SPOP^{WT} and LNCaP-Abl F102C vs WT (**A**), LNCaP-Abl F102C vs control vector (**B**), LNCaP F102C vs WT (**C**) or CWR22Rv1 siSPOP vs siScr (**D**) with hallmarks shared across all datasets annotated.*

enriched, the TNFa signalling by NFKB hallmark was shared across 3 of the 4 datasets, with the SPOP knockdown in the CWR22Rv1 cell line being the only dataset where this negative enrichment was not seen. Also of note, this was the only dataset to see no shared hallmarks with the TCGA dataset, where the other published data-sets demonstrated at least one shared hallmark [Figure 6.17D]. Encouragingly, negative enrichment was seen in the *mTORC1 signalling*, *oxidative phosphorylation* and *unfolded protein response* hallmarks between the CWR22Rv1-AR-EK-SPOP$^{-/-}$ clones and the siSPOP-treated CWR22Rv1 cells. Taken together, through expanding the RNAseq analysis to incorporate published datasets, more *in silico* evidence has been provided, supporting the use of these models in the investigation of SPOP controlled pathways. Importantly, TNFa signalling by NFKB is of key interest due to it being shared across SPOP manipulated cell line-based experiments as well as data from SPOP mutant patients.

6.5 Discussion

With a wealth of data now focussing on the diverse roles SPOP plays within normal cell homeostasis, understanding the wider effects of SPOP functional loss on cancer cells is of great interest. Using the available data for SPOP mutant patients, published datasets utilising different methods of mimicking SPOP loss of function, alongside the novel CWR22Rv1-AR-EK-SPOP$^{-/-}$ cell line derivatives developed previously, allows for more in-depth investigation, and combining insights from the three allows for greater confidence in the results from *in silico* analyses. To that end, work described in this chapter looked to run *in silico* RNAseq analysis on the new CWR22Rv1-AR-EK-SPOP$^{-/-}$ clones and compare the results from these to relevant publicly available data to assess how representative these new models were of patient disease. Data from the *SPOP* knock-out clones revealed enrichments in many cellular pathways, including upregulation of genes involved in the EMT pathway, which was experimentally validated through *in vitro* scratch wound assays. With direct *in vitro* cell phenotypes recapitulating *in silico* analysis results, data generated from this analysis revealed many potential pathways worth further investigation which could direct to the discovery of further drug susceptibilities as a result of loss of function *SPOP* mutations in PC.

Analysis from the SPOP knock-out clones showed similar expression profiles in relation to total SDEGs, but comparisons between SDEG lists from each clone revealed ~25% SDEGs were not shared. As a result, the differences between the two clones' expression profiles resulted in alternate hallmark pathways enriched between the two. These distinct transcriptional profiles demonstrated by each clone could be due to off-target effects of the CRISPR guides used. However, the guide designs used to develop the clones had the lowest off-target stores across the range of design tools used. Using more in-depth analysis tools on the data may reveal if there have been any off target indels introduced. The benefit of this heterogeneity between clones means that even with transcriptional differences, both clones showed similar results in experiments conducted in Chapter 5, indicating that DNA-PKi treatment resulted in similar sensitivity regardless of the expression changes between the clones giving more confidence in the results produced.

In silico data analysis from the SPOP knock-out clones revealed EMT as a positively enriched pathway. SPOP playing roles in EMT has been described in the literature in both prostate and breast cancer (Ma et al., 2021, Wei et al., 2022). In prostate cancer ITCH, a key regulator of both the immune response and EMT, was described as an SPOP substrate, with SPOP knockdown shown to increase cell migration (Ma et al., 2021). In breast cancer TWIST1, a key regulator of EMT, was described as an SPOP substrate where SPOP depletion resulted in enhanced TWIST1 levels (Wei et al., 2022). Encouragingly, the positive enrichment observed in the *EMT* hallmark translated to increased migration of CWR22Rv1-AR-EK-SPOP$^{-/-}$ cells compared to the parental control, mirroring that described from both *in silico* analysis and the published literature. Ultimately, thus indicates the data generated from the RNAseq analysis aligns to phenotypic impact and provides further confidence for investigating those alternate pathways that have been significantly altered in response to SPOP loss.

With the aim of elucidating how representative these new *SPOP* knock-out models were of patient disease, results from the RNAseq analysis were compared against RNAseq analysis from SPOP^{MT} vs SPOP^{WT} patient data available from the TCGA. As expected, from the GSEA hallmark analysis, *androgen response* was seen amongst the positively enriched hallmarks in the mutant *SPOP* patients. This elevation in AR signalling was not mirrored in the knock-out clones, likely due to the previously mentioned lack of SPOP binding motif in the CWR22Rv1-AR-EK cell background due to the line only expressing AR-Vs. Promisingly, there were negatively-enriched hallmarks shared between this TCGA dataset and either both or individual clones, indicating that the underlying cellular processes within these cells are somewhat representative of that in patient cancers. Unfortunately, there was no common positively-enriched hallmarks shared between either clone or the TCGA dataset. Seeing common negatively-enriched hallmarks, although few in number, across the three comparisons is encouraging and highlights key pathways of interest as the results from *in vitro* cell line work is more likely to translate *in vivo*. Another benefit of this TCGA dataset is that patients have only been sorted based of *SPOP* mutation status, meaning that GSEA hallmark outputs are present regardless of other genetic factors at play, such as AR-V positivity or other oncogenic drivers.

To expand this analysis further, RNAseq data from published articles was investigated and subsequently compared against both in-house SPOP knock-out analysis and TCGA $SPOP^{MT}$ vs $SPOP^{WT}$ datasets. Of the datasets available, the work published in Bezaway *et al*., was not included as the overexpression study was conducted in the AR negative DU-145 line. As the effects of SPOP mutations on AR signalling were not the main focus of this project, this dataset was not included and instead, those studies utilising cell lines with active AR signalling were included to mirror the constitutive AR activity in the CWR22Rv1-AR-EK cell background. This DU-145 dataset would provide greater insights as the repertoire of SPOP knock-out cell lines is expanded to alternate cellular backgrounds such as the planned AR negative PC3 line. Of note from the initial analysis performed on the data from the LNCaP-Abl cells overexpressing $SPOP^{WT}$ and a range of mutant forms of the protein, was that a subset of genes saw increased expression in all experimental arms where any form of SPOP was overexpressed. As discussed previously in Section 5.5, there may be effects of SPOP on cell homeostasis out with its substrate binding region in the MATH domain. This data would indicate that increased SPOP protein abundance within cells will cause transcriptional changes regardless of mutations status, indicating alternate roles for SPOP. The issue here is that in the case of knock-out cell models, these alternate roles will also be lost, thus endeavours into generating knock-in SPOP models or re-introducing mutant SPOP expression in SPOP knock-out backgrounds would be of great interest and result in the most impactful models.

With the *TNFa signalling via NFKB* hallmark being the most common negatively-enriched across the datasets analysed, deeper investigation into this pathway may reveal highly translatable results. When looking to published literature, PC-associated *SPOP* mutations have been implicated in upregulation of non-canonical STING-NFKB signalling (Geng et al., 2023). Within this recent study, GSEA of CRPC patient tumours displayed positive enrichment of the *TNFa signalling via NFKB* hallmark. This directly contrasts data presented from the CWR22Rv1-AR-EK-$SPOP^{-/-}$ clones, TCGA analysis and the majority of the publicly available data, with positive enrichment only seen in the CWR22Rv1 siRNA knockdown dataset. Of note here, NES from the public data were much lower than the NES described in the in-house analysis, where public data saw hallmark NES of ~$\pm$0.5 whereas the *SPOP* knock-out and TCGA analysis saw NES of ~$\pm$2-3 meaning that these enrichments may not be as evident from the public data. Also, the patient $SPOP^{MT}$ vs $SPOP^{WT}$ dataset used by Geng *et al*., was from a cohort of

neuroendocrine PC tumours, hence the described large transcriptomic changes seen when tumours shift to neuroendocrine phenotypes could be the cause of the shift from negative to positive enrichment of the *TNFa signalling via NFKB* hallmark (Feng et al., 2022, Zhang et al., 2024). Together, this data suggests deeper investigation into this pathway is necessary, however this could reveal potentially exploitable cellular phenotypes with high translatable potential.

With aberrant gene expression leading to positive and negative enrichment in a range of pathways across the datasets investigated in this chapter, there is many avenues for future lines of experimental investigation. One of the main caveats to the data presented is that this is transcriptomic changes in response to SPOP manipulation, which means that the effects described here are indirect effects of SPOP loss of function facilitated by changes in post-translational regulation of SPOP substrates. To enhance the data presented, proteomic assessment of the effects of complete SPOP loss using the knock-out models would greatly increase the impact of the work. Alongside looking into the general proteomic changes in cells as a whole, doing targeted ubiquitin pulldowns in both the parental and SPOP knock out cell lines may reveal a definitive list of SPOP substrate proteins.

Ultimately, data presented in this chapter has unveiled a range of potential experimental routes of investigation which may reveal therapeutic susceptibilities caused by a loss of SPOP activity within cells, which were importantly backed up by patient data. Expanding these large-scale transcriptome or potential proteome assessments to alternate PC cell line models could not only support the current analysis but also indicate alternate targets of interest for future investigation.

Chapter 7 – Utilisation of DNA based nanotechnology for PC associated mRNA biomarker detection.

7.1 Introduction

7.1.1 AR-Vs as a form of PC therapy resistance

One of the key mechanisms at play in the progression of PC to the more aggressive disease state is an increase in alternatively spliced AR isoforms, termed AR-Vs (Dehm and Tindall, 2011). As a result, AR-Vs present one of the main resistance mechanisms that allow PC to develop ADT resistance and cause PC relapse to more aggressive CRPC. This is potentially due to ADT acting as a selection pressure that favours PC cells able to bypass androgen deprivation via expressing constitutively active AR-Vs which lack the AR LBD; the intended therapeutic target site (Dehm and Tindall, 2011, Lu and Luo, 2013). The generation of constitutively active AR-Vs is primarily by alternative splicing which incorporates C-terminal short cryptic exons, present within intron 2 and 3 of the *AR* gene, at the expense of exons 4-8, causing loss of the LBD region, but retention of the transcriptionally activate NTD, and DBD. The resultant variants produced, such as AR-V1 and AR-V7, are constitutively active, support the CRPC phenotype through hormone-independent AR signalling, and can be detected in clinical PC samples. Another clinically relevant variant is AR^{v567eS} which shifts the open reading frame of the AR gene and results in a constitutively active variant that contains the hinge region but lacks the LBD. Both AR-V7 and AR^{v567eS} regulate expression of AR target genes, even without the presence of active full-length AR (AR-FL) (Lu and Luo, 2013). Further, following downregulation of AR-FL signalling via knockdown, ADT or enzalutamide treatment there is an increase in AR-V7 expression (Sharp et al., 2019). Also of note is that the expression of these AR-Vs has not been shown in normal prostate tissue, indicating that their expression is now wholly pathogenic (Dehm and Tindall, 2011, Lu and Luo, 2013). These, combined with emerging evidence that AR-Vs look to have minimal activity in the absence of AR-FL indicates that expression of these AR-Vs represents a simple mechanism for PC cells to develop resistance (Dehm and Tindall, 2011, Wüstmann et al., 2023). As continued AR signalling is a main driver of disease progression following ADT, research into regulators of AR-Vs have become a key area of focus (Longo, 2010).

7.1.2 AR-V7 as a PC biomarker

CRPC patients that are treated with either enzalutamide or abiraterone, both aimed at inhibiting AR signalling through competition or depletion of androgens (Attard et al., 2008, Scher et al., 2010), show an increased survival rate (de Bono et al., 2011, Scher et al., 2012). However, in the population of patients who initially have a positive response to these agents, almost all develop secondary resistance to these treatments, with one resistance mechanism observed in patient cohorts is the presence of AR-Vs (Mostaghel et al., 2011, Nadiminty et al., 2013). AR-V7 is one of a sub-set of ligand independent transcription factors which can be detectable reliably in circulating tumour cells in clinical specimens (Guo et al., 2009). Due to the lack of LBD, AR-V7 is refractory to the current repertoire of ligand targeting agents such as enzalutamide and abiraterone. In clinical samples with detectable AR-V7 within circulating tumour cells, there was seen to be an associated resistance to both enzalutamide and abiraterone (Antonarakis et al., 2014). Further, clear increases in AR-V7 expression can be seen when comparing castrate-sensitive disease to CRPC patient samples (Sharp et al., 2019).

Taxanes can be used to treat PC and cause mitotic arrest and anti-tumour effects via disruption of the cytoplasmic-to-nuclear trafficking of AR along the microtubule network (Gan et al., 2009, Thadani-Mulero et al., 2012). Thus, taxane therapy can be used to disrupt AR signalling by inhibiting AR translocation to the nucleus and sequestering it in the cytoplasm. The presence of AR-Vs can either confer resistance or sensitivity to taxane chemotherapy (Thadani-Mulero et al., 2014). Although the presence of AR-V7 can be used to denote resistance to enzalutamide and abiraterone, it seems to have no effect on resistance to taxane therapy (Antonarakis et al., 2015). This provides the opportunity, upon detection of AR-V7 in a patient's PC, to redirect the standard enzalutamide or abiraterone treatment to taxane chemotherapy to yield better results.

These are examples of the presence of AR-V protein and RNA being used as a marker for a more advanced disease. There is also the potential with this information to use the presence or absence of AR-V7 to direct patient treatment. Currently available commercial methods of AR-V7 detection look to detect AR-V7-positive patients from CTCs via liquid biopsy following lack of response to AR targeted therapies, such as those marketed by Epic Sciences and Qiagen.

Specifics on the method utilised for detection remains unclear with the test from Epic Sciences. The AdnaTest ProstateCancerPanel AR-V7 from Qiagen uses a combination of metal-conjugated antibodies to capture CTCs from patient blood, then RNA extraction and subsequent RT-qPCR are performed to quantify AR-V7 mRNA within the sample alongside other PC associated biomarkers PSA, PSMA and AR-FL. As with the test provided by Epic Sciences, the AdnaTest boasts great specificity however comes with high costs per run ($3,516 for 12 samples).

7.1.3 Current methods of SNV discrimination

Reliable detection of single base changes from patient samples holds massive importance in patient stratification and direction of bespoke treatment. The current gold-standard for this used in many patient datasets is next-generation sequencing which provides an accurate readout of a patient's mutational burden although this comes with increased costs and high demands on individuals and computational pipelines for data analysis. Although this may be necessary in some situations, it is unlikely for this to be performed routinely on a patient-to-patient basis. However, in attempts to reduce the associated costs of NGS, targeted NGS panels looking at cancer associated SNPs have been utilised on primary and liquid biopsy samples, with commercial available kits utilising the technology also commercially available, such as the AmpliSeq technology from ThermoFisher scientific (Stover et al., 2022). In attempts to overcome the shortcomings of NGS, simpler methods have been developed to detect point mutations within given samples to better identify key mutations that may confer pro-tumorigenic genotypes of a given patients cancer, such as traditional methods like qPCR-based techniques, or non-conventional methods such as the development of elaborate novel biosensor platforms (Jiang et al., 2021).

Of these, qPCR-based techniques boasts one of the simplest forms of SNV detection which has been implicated through a range of different techniques as a faster more cost-effective method of single-base mutation detection. Work has been conducted by multiple groups in attempts to adapt simple qPCR techniques for use in routine SNV detection, such as the work presented by Alvarez-Garcia *et al.,* (2018) which targets two cancer associated mutations (E545K & H1047R) within *PIK3CA* (Alvarez-Garcia et al., 2018). Here, using SYBR green based reagents and mutant allele-specific primers and a phosphate-modified oligonucleotide targeting the WT

sequence to block amplification of the WT allele, only amplification of the mutant input is permitted. Using this technique, the proposed assay was sensitive enough to detect 5% of mutant sequences amongst a 95% WT sequences (seemingly the common form of measuring sensitivity of SNV detection assays) and was able to positively identify mutant alleles from fresh frozen and embedded patient tumour biopsy material.

Another method of single-base mutant detection has been proposed by Xue *et al.,* (2021) which utilised primers targeting EGFR mutations (Xue et al., 2021). Here primers are designed incorporating the base change present in the T790M mutant, where upon interaction with the WT sequence will leave an open mismatch. From here, gold nanoparticles to bind and inhibit amplification of WT inputs are utilised, leading to detectable changes in the fluorescent output of the Taqman probes. Using EGRF mutations as a candidate marker, the assay was able to detect 0.95% mutant input abundance within spiked samples. A key drawback of these qPCR-based methods of SNV detection is that they are typically developed and optimised for specific mutations and whether these can be translated to other key mutations within different cancer subtypes is currently unknown.

Alongside these more experimental techniques, commercial options have been developed for qPCR-based point mutation detection, with two of the main techniques used being castPCR™ (Applied Biosystems) and ARMS-qPCR (Qiagen, Roche) describing 0.01% and 1% mutant sequence detection, respectively. Although boasting high sensitivity for mutant allele detection, evidence from Vallée *et al.,* (2014) suggested that in the case of specifically detecting $EGFR_{T790M}$ mutations via ARMS-qPCR based techniques, a limit of detection (LOD) of 7.02% MT input is reached in contrast to the reported 1% (Vallée et al., 2014). Assessing detection of the same EGFR mutation using castPCR™, experimentally, Yang *et al*., (2018) found an LOD of 1% rather than the described 0.1% (Yang et al., 2018). Furthermore, the adaptability of these assays for new biomarkers is low, as only a select range of mutations are available for purchase and the relative cost of these assays are much greater than traditional qPCR experiments.

With experimental methods boasting greater cost effectiveness but lack of validation for alternate targets, and commercial kits presenting high specificity but increased costs and limited target choice, novel methods of SNV detection that are cost effective and hold greater sensitivity than the current experimental best of 0.95% MT input are of great interest. The deployment of DNA nanotechnology-based detection here may overcome many of the drawbacks presented and provide a novel detection platform for rapid and cost-effective selective detection of SNVs within a given sample.

7.2 Aims

Currently there is a substantial body of evidence indicating the AR splice variant AR-V7 as a key driver of PC progression. AR-V7 expression has been shown to massively increase in expression as the disease shifts to the hormone therapy-resistant CRPC stage. This high transcript abundance combined with harbouring a unique nucleotide sequence within cryptic exon 3 (CE3), as a consequence of an aberrant splicing event, make it an ideal candidate as a biomarker for aggressive PC. Thus, using AR-V7 as a candidate biomarker, new DNA nanotechnology-based detection platforms will be designed and tested to gain a better understanding of this new technology and assess its suitability for use in biomarker detection.

The current repertoire of SNV detection platforms displaying key drawbacks, predominantly in either selectivity, adaptability to new targets or cost effectiveness. As such, further investigations using DNA-based nanotechnology approaches will be deployed to assess if this is a viable avenue for the development of new SNV detection platforms. Considering interest in SPOP biology within the host laboratory, and that PC-associated mutations are single base changes, *SPOP* mutations will be used as a candidate for validation of this new SNV detection platform.

To achieve this, work will be conducted in attempts to satisfy two main objectives:

I. Design probes targeting the intersection between AR exon 3 (E3) and CE3 in attempts to detect AR-V7 mRNA and assess sensitivity and selectivity in a range of background media with aims of detecting endogenous transcripts within PC cell line lysates.
II. Utilise novel probe designs in attempts to discriminate between wild-type and mutant *SPOP* inputs to a greater selectivity than that presented by the current repertoire of qPCR-based SNV detection platforms.

7.3 Methods

7.3.1 AR-V7 mRNA detection

Probe sequences.

Table 7.1 - *List of sequences used for AR-V7 detection probes and associated 5' and 3' modifications.*

Oligo	**Sequence**	**5' Modifier**	**3' Modifier**
CE3-24 F	ATTGCCAACCCGGAATTTTTCTCC	FAM	
CE3-24 Q	GGAGAAAAATTCCGGG		Dabcyl
CE3-30 F	ATTGCCAACCCGGAATTTTTCTCCCAGAGT	FAM	
CE3-30 Q	ACTCTGGGAGAAAAATTCCGGG		Dabcyl
CE3-40 F	ATTGCCAACCCGGAATTTTTCTCCCAGAGTCATCCCTGCT	FAM	
CE3-40 Q	AGCAGGGATGACTCTGGGAGAAAAATTCCGGG		Dabcyl
E3-30 F	CGGAATTTTTCTCCCAGAGTCATCCCTGCT	FAM	
E3-30 Q	TGACTCTGGGAGAAAAATTCCG		Dabcyl
CE3-24 T	GGAGAAAAATTCCGGGTTGGCAAT		
CE3-30 T	ACTCTGGGAGAAAAATTCCGGGTTGGCAAT		
CE3-40 T	AGCAGGGATGACTCTGGGAGAAAAATTCCGGGTTGGCAAT		
CE3-60 T	CGTCTTCGGAAATGTTATGAAGCAGGGATGACTCTGGGAGAAAAATTCCGGGTTGGCAAT		
E3-30 T	AGCAGGGATGACTCTGGGAGAAAAATTCCG		

7.3.2 SPOP SNV detection

Probe sequences.

Table 7.2 - *List of sequences used for SPOP based SNV discrimination probes and associated 5' and 3' modifications.*

Oligo	**Sequence**	**5' Modifier**	**3' Modifier**
MT-Duplex F	*GCTACGGATGAATTTCTTGAC*	6-FAM	
MT-Duplex Q	*AGAAATTCATCCGTAGC*		*BHQ®-1*
MT-Hairpin	*AGAAATTCATCCGTAGTTTCTCTACGGATGAATTTCTTGACT*	*6-FAM*	*BHQ®-1*
WT-IS	*ATTCAAGAAATTCATCCGTAGAG*		
MT-IS	*AGTCAAGAAATTCATCCGTAGAG*		
WT-IL	*GCATATAGGTTTGTGCAAGGCAAAGACTGGGGATTCAAGAAATTCAT CCGTAGAGATTTTCTTTTGGATGAGGCCAACGGGCTTCT*		
MT-IL	*GCATATAGGTTTGTGCAAGGCAAAGACTGGGGAGTCAAGAAATTCAT CCGTAGAGATTTTCTTTTGGATGAGGCCAACGGGCTTCT*		
Blocker 1	*CCCAGTCTTTGCCTTGCACAAACCTATATG*		
Blocker 2	*AGAAGCCCGTTGGCCTCATCCAAAAGAAAA*		

7.4 Results

7.4.1 AR-V7 biomarker detection - Probe Design

As highlighted previously, due to aberrant splicing events, AR-V7 contains an exclusive region present adjoining E3 to a cryptic exon CE3 [Figure 7.1A]. Thus, targeting this region with probes originating in CE3 would potentially allow for discriminate detection of AR-V7 mRNA within samples.

Nanovery probes consists of two oligonucleotides that have either a FAM or Dabcyl modification at their 3' or 5' end and are named the Fluorophore (F) (FAM modification) or the Quencher (Q) (Dabcyl modification). The Quencher sequence will be complementary to the Fluorophore while the sequence of the Fluorophore is complementary to the target region of interest, however, longer than the Quencher to provide a 'toehold' to facilitate strand exchange upon interaction with the target sequence. To make the final usable probe [Figure 7.1B], both these Fluorophore and Quencher oligonucleotides are annealed together at 95 ^{o}C. The Fluorophore alone emits a constant florescent signal that is silenced by the Quencher when annealed in the final probe. However, following the TMSD reaction, the Quencher is displaced by the target RNA and the fluorescent signal is emitted which can be detected using a plate reader. 3 probes were used that targeted CE3 with lengths of 24, 30 and 40 nucleotides, and 1 probe was designed to target the upstream E3, encoded in full-length and AR-V7 isoforms, which was 30 nucleotides long. All 4 probes cross the E3-CE3 junction and share some sequence similarity. Of the 4 different probes tested, each shared the Quencher being 8 nucleotides shorter than the Fluorophore, resulting in an 8 nucleotide 'toehold'. The Fluorophore-tagged oligonucleotide lengths were 24, 30 and 40 nucleotides, with two 30 nucleotide long Fluorophores each with the FAM signal either at the 3' or 5' end. Following annealing, probes were referred to as CE3-24P, CE3-30P and CE3-40P where the 'toehold' was positioned within CE3, and the E3-30P where the 'toehold' was positioned within E3 [Figure 7.1B]. Synthetic target sequences were also generated with their 5' end originating in CE3 that were 24, 30, 40 and 60-nucleotides (CE3-24T, CE3-30T, CE3-40T and CE3-60T) long as well as a 30-nucleotide (E3-30T) sequence with 5' end originating in E3.

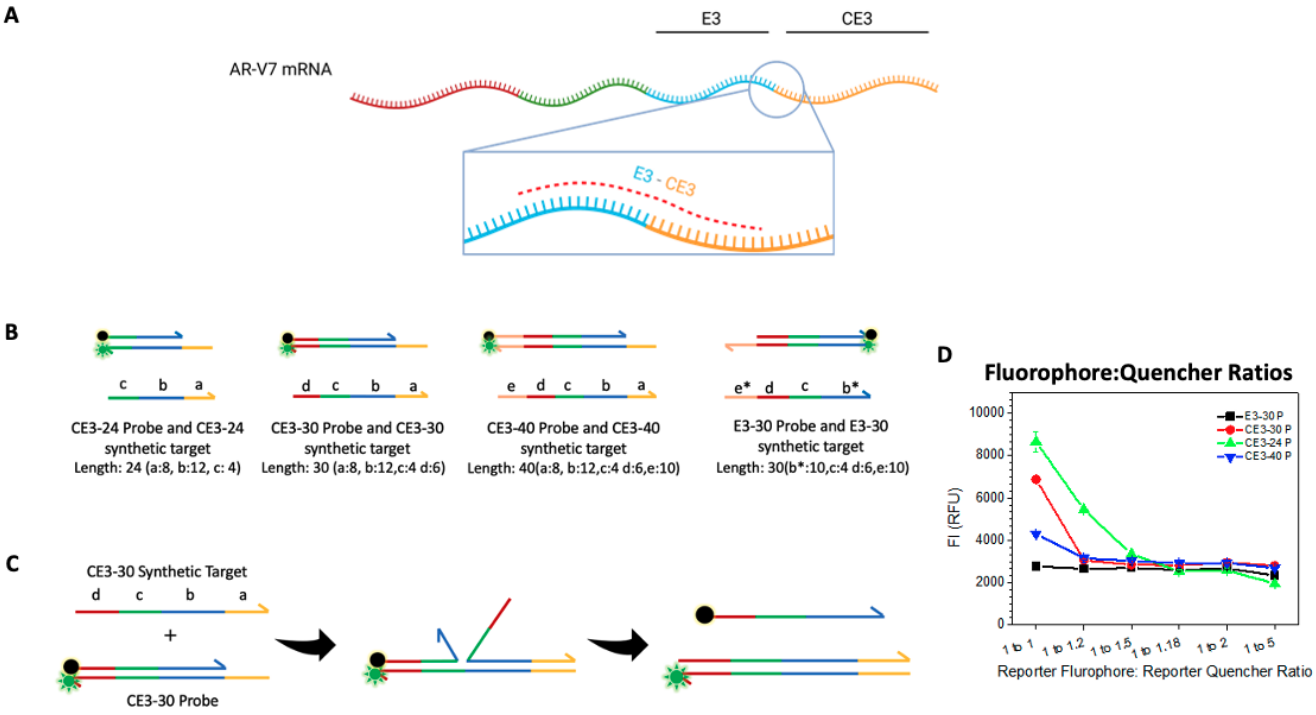


Figure 7.1 - Rationale for AR-V7 targeted probe designs. A. *Schematic representation of AR-V7 mRNA, with the highlighted E3-CE3 junction and dashed red line indicating probe target region.* ***B.*** *Schematic representation of CE3-24, CE3-30, CE3-40 and E3-30 probes with their associated synthetic targets. Regions shared across synthetic targets are denoted a-e, with their associated lengths denoted.* ***C.*** *Schematic representation of CE3-30P interacting with the CE3-30T and undergoing TMSD reactions to produce detectable fluorescent signal.* ***D.*** *Endpoint RFU output from a range of Flurophore:Quencher ratios in the CE3-24P, CE3-30P, CE3-40P and E3-30P following 180 fluorescence detection window. Data is representative of 3 individual experiments ± SEM.*

First, before the probe's ability for target detection could be assessed, the optimum ratio of Fluorophore to Quencher needed to be determined. The probes were assembled with F:Q ratios ranging from 1:1 (100 nM F: 100 nM Q) to 1:5 (100 nM F: 500 nM Q). The optimum ratio would contain the lowest amount of unbound material and return the most stable, yet minimal background fluorescence across the detection period. To do this, the background fluorescence of various F:Q ratios was detected across 180 minutes and RFU produced at the endpoint taken [Figure 7.1D]. These results show that the ratio that returned the least background fluorescence was the 1:2 F:Q ratio, as the background across all 4 probes was between 2000 and 3000 RFU, whilst containing the least material and remaining stable across the 180 minutes. Ratios with increased Quencher produced a similar background fluorescence, although they could result in excess material being present that may interfere with detection in later experiments. Using these results, a 1:2 ratio was used in all subsequent experiments, and final probe concentration of 10 nM was used unless stated otherwise.

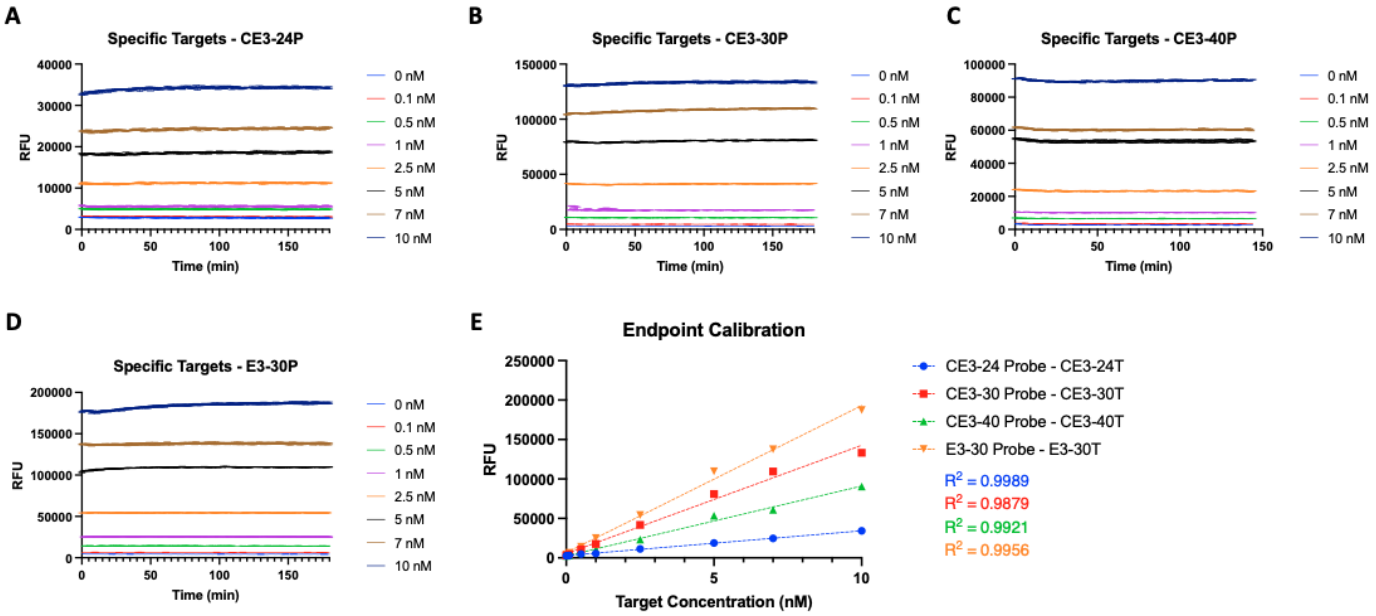


Figure 7.2 - CE3-24P, CE3-30P, CE3-40P and E3-30P all react with their cognate inputs and produce high-confidence standard curves. *Concentration ranges of 0– 10 nM of each probes specific input was added to the CE3-24P (**A**), CE3-30P (**B**), CE3-40P (**C**) and E3-30P (**D**) and RFU output was assessed over 180 minutes. Data is representative of 3 individual experiments ± SEM. **E.** Endpoint values from each of the previous target concentration ranges were taken for each probe and its specific target and plotted to determine R2 value. Data is representative of 3 individual experiments ± SEM.*

7.4.2 Initial probe validations and determination of the effects on probe action with increasing input length

To firstly validate the probes were able to undergo the expected strand displacement reactions, assays were set up utilising the specific synthetic target for each probe at a range of concentrations [Figure 7.2A-D]. In all cases the specific target for each probe was able to trigger the reaction within each probe. In the case of the CE3-24P, the maximal RFU returned from the 10 nM target input was 34277 RFU and 18764 from the 5 nM input; approximately half of the maximal value. This trend was observed in all other probes tested. Finally, using the endpoint RFU from each reaction, standard curves were generated [Figure 7.2E]. Linear regressions were plotted using this endpoint data to give an idea of the expected R^2 values from an optimum reaction for each of these probes. Each of these standard curves returned R^2 values of greater than 0.98, with CE3-30P returning the lowest of 0.9879 and the CE3-24P displaying the highest R^2 of 0.9989. Also of note, each probe returned different maximum RFU outputs at the 10 nM input level, with an RFU returned of 133267 and 187454 for the CE3-30P and E3-30P, respectively.

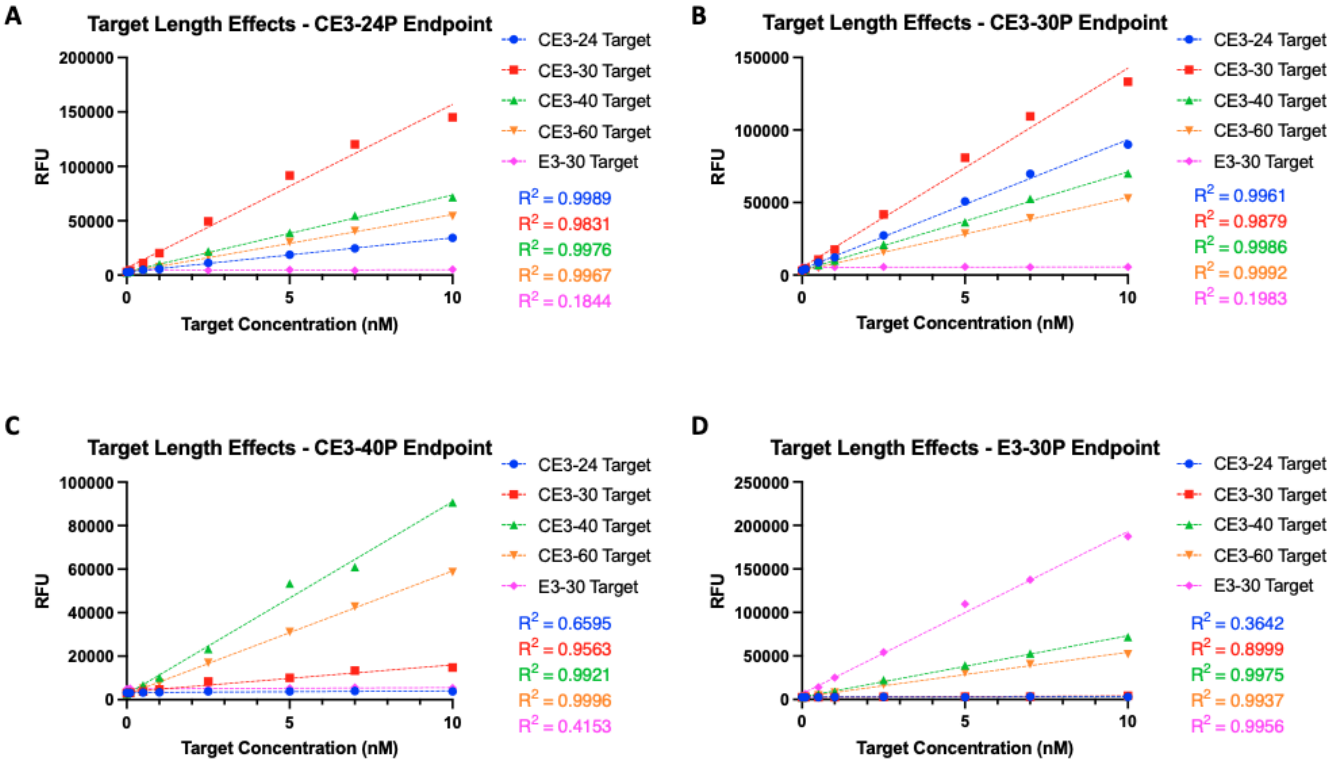


Figure 7.3 -– Input length determines ability to trigger probes. *Endpoint RFU output following 180-minute fluorescence detection from 0– 10 nM concentration ranges of CE3-24, CE3-30, CE3-40, CE3-60 and E3-30 targets with the CE3-24P (**A**), CE3-30P (**B**), CE3-40P (**C**) and E3-30P (**D**) probes. Average RFU from each concentration's endpoint was used to plot standard curve and calculate R2 value. Data is representative of 3 individual experiments ± SEM.*

Next to assess the effects of varying lengths of targets on the action of the different probes, reactions were set-up with each probe alongside a range of concentrations of each target [Figure 7.3]. The CE3-24P was able to be triggered by all targets regardless of length, but as expected, was unable to be triggered by the E3-30 target due to the lack of complementarity within the CE3-24P toehold region [Figure 7.3A]. Alongside this, there was minimal changes on the resultant R^2 values calculated showing that the length of CE3 target has no effect on the calibration of the standard curves produced. Interestingly the maximal RFU output changed with the length of target used. When using the CE3-30 probe the maximal RFU output was seen with the fully complementary CE3-30 target (CE3-30T) [Figure 7.3B]. A similar effect on probe calibration was seen as before in all the CE3 targets, with the E3-30 target not being able to trigger a reaction in the CE3-30P. Also of note, the shorter CE3-24T was able to trigger the CE3-30P, implying that the 6 nucleotides remaining bound upon target interaction within the probe were displaced by the complementary binding by the 24 nucleotides facilitated by the toehold resulting in complete removal of the quencher. The ability for the CE3-24T to

trigger the displacement reaction was not continued as the probe length increased in the CE3-40P [Figure 7.3C]. Further, although able to trigger the displacement reaction, the CE3-30T target did not produce as high a maximal RFU output as seen with the CE3-24P and CE3-30P. Again, the E3-30T was unable to trigger the displacement reaction at any of the target concentrations tested. In the case of the E3-30 probe, none of the CE3- targets under 40 nucleotides in length were able to trigger a reaction [Figure 7.3D]. However, the 40 and 60 nucleotide targets were able to facilitate strand displacement with minimal impact on probe calibration in regard to the returned R^2 and final RFU output. Encouragingly, the E3-30T triggered the E3-30 probe as expected. Together, this data shows that there is a minimum length of complementary nucleotides between the target and the probe to facilitate the strand displacement reactions which is between 6-10 nucleotides. However, once passed this threshold, there is minimal effect on the probe's calibration although maximal RFU output may be affected.

7.4.3 Limits of detection of probes are dependent on input length

Satisfied that all probes were functioning as expected, the CE3-30P and E3-30P probes were selected as candidate probes and taken for further analysis (experimental runs were also performed in the CE3-24P and CE3-40P). To further assess the probe capabilities as a new platform for biomarker detection, the LOD of these simple duplex probes were assessed, first with the CE3-30P with each target [Figure 7.4]. In isolation, the probes in buffer without a target (probe alone) produced a background fluorescence which was used as the baseline RFU. Significant increases in RFU above this level being used as a positive output. In the case of the CE3-24T, a significant increase from a baseline of 3306 RFU to 4380 RFU was seen upon interaction with 0.1 nM of target, with a stepwise increase seen up to the maximum RFU of 92718 output at 10 nM input concentration [Figure 7.4A]. A similar trend was seen with increased input length in the CE3-30T with a significant increase of 1461 RFU over the baseline RFU 0.1 nM input concentration, and again with the CE3-40T which returned a significant 896 RFU increase over the baseline at 0.1 nM input concentration [Figure 7.4B&C]. In contrast, when the target length increased further in the CE3-60T, no significant detection was observed at the 0.1 nM input concentration, although a slight increase of 372 RFU over the baseline was

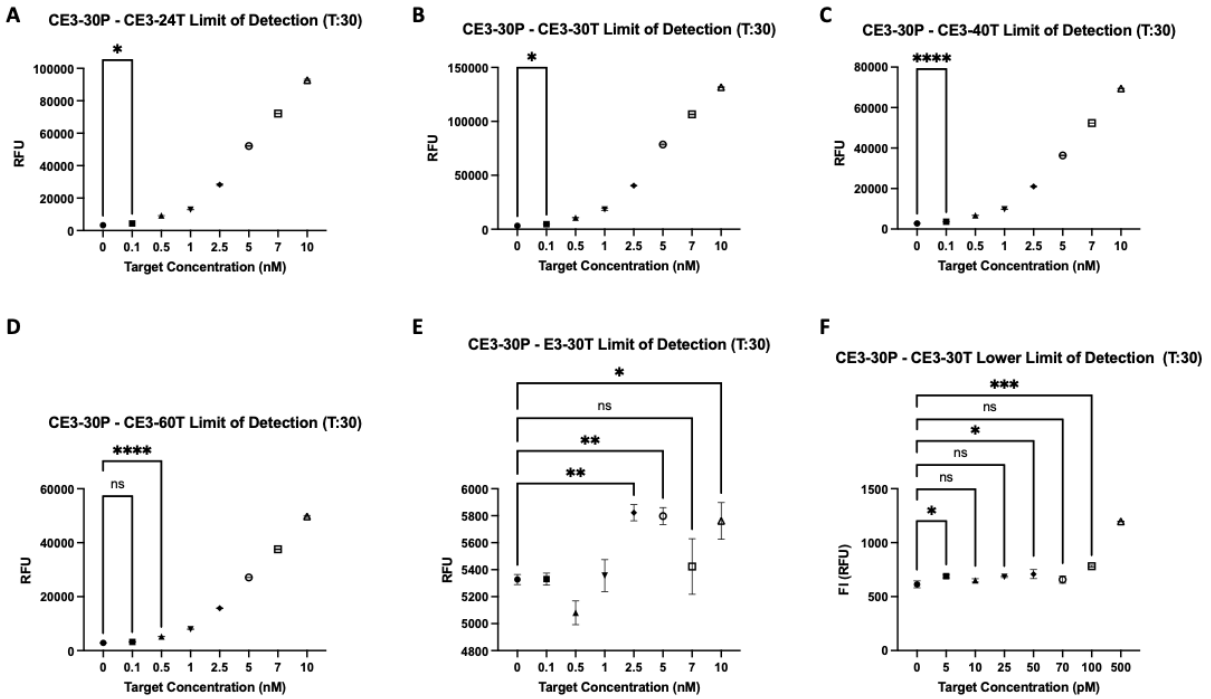


Figure 7.4 – Limit of detection determination for CE3-30P with each input. *CE3-30P tested against 0 – 10 nM target concentration range of CE3-24T (**A**), CE3-30T (**B**), CE3-40T (**C**), CE3-60T (**D**) and E3-30T (**E**) based of RFU output at 30 minutes. Data is representative of 3 individual experiments ± SEM. Statistical significance was determined by a One-way ANOVA with Dunnett's multiple comparison test (*, **, *** and **** denote P-values <0.0332, <0.0021, <0.0002 and <0.0001 respectively, ns denotes not significant).* ***F.*** *RFU output from CE3-30P and 0 - 500 pM concentration range of CE3-30T at 30 minutes. Data is representative of 3 individual experiments ± SEM. Statistical significance was determined by a One-way ANOVA with Dunnett's multiple comparison test (*, **, *** and **** denote P-values <0.0332, <0.0021, <0.0002 and <0.0001 respectively, ns denotes not significant).*

seen [Figure 7.4D]. Only upon elevating the input concentration to 0.5 nM was significant detection observed, with an RFU of 5213 returned when compared to the baseline RFU of 2871. When using the E3-30 target, significant detection was seen upwards of 2.5 nM although the RFU increases were minimal when compared to previous results [Figure 7.4E]. With this being a non-specific input, and not observing linear increases in RFU proportional to increasing input concentrations, as detected for the other targets tested, the results here indicate no true reaction being triggered.

Finally, with the LOD for the CE3-30P determined to be at least 0.1 nM target concentration, an experiment was conducted using the CE3-30T in conjunction with utilising an increased gain on the plate reader and reduced probe concentration to 1 nM to establish if the sensitivity of the assay can be increased while reducing the total probe background RFU which would enhance signal to noise ratio [Figure 7.4F]. Considerable variance between 613 and 708 RFU was seen in the target concentrations below 100 pM. However, consistent with previous data,

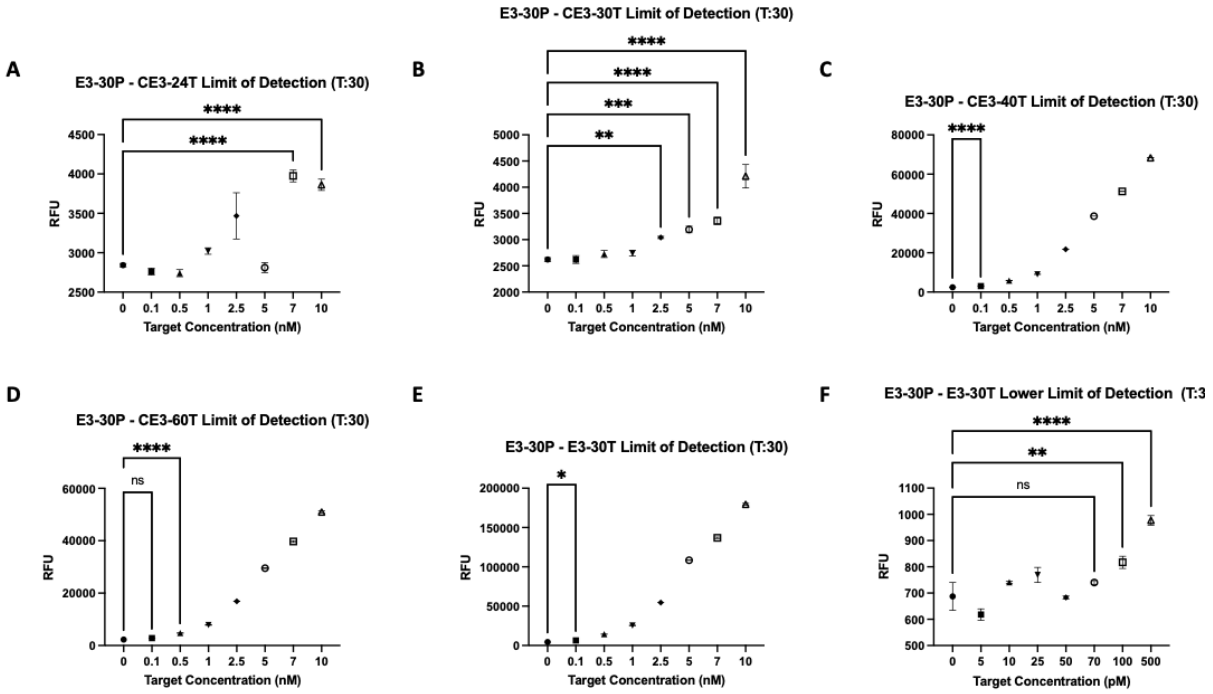


Figure 7.5 – Limit of detection determination for E3-30P with each input. *E3-30P tested against 0 – 10 nM target concentration range of CE3-24T (**A**), CE3-30T (**B**), CE3-40T (**C**), CE3-60T (**D**) and E3-30T (**E**) based of RFU output at 30 minutes. Data is representative of 3 individual experiments ± SEM. Statistical significance was determined by a One-way ANOVA with Dunnett's multiple comparison test (*, **, *** and **** denote P-values <0.0332, <0.0021, <0.0002 and <0.0001 respectively, ns denotes not significant).* ***F.*** *RFU output from E3-30P and 0 - 500 pM concentration range of CE3-30T at 30 minutes. Data is representative of 3 individual experiments ± SEM. Statistical significance was determined by a One-way ANOVA with Dunnett's multiple comparison test (*, **, *** and **** denote P-values <0.0332, <0.0021, <0.0002 and <0.0001 respectively, ns denotes not significant).*

significant detection over the probe alone 613 RFU was observed at 100 pM (781 RFU) which further increased to 1197 at 500 pM. Taken together, these results indicate that for the CE3-30 probe, the LOD was around 0.1 nM, which decreased when tested with the longer CE3-60T.

As before, the LOD for the E3-30P was tested across the range of synthetic inputs available [Figure 7.5]. With the non-complementary CE3-24T, as in Figure 7.4E, significant detection above the background was seen at the higher 7 nM and 10 nM target concentration, although the total RFU changes across range of concentrations was minimal [Figure 7.5A]. Furthermore, the RFU obtained from the 7 nM input sample was 3975 which saw a decrease to 3862 at 10 nM input. Thus, the CE3-24T data indicates that the shorter target was unable to trigger the expected reaction. In the case of the CE3-30T, the RFU saw a liner increase with increasing input concentration. However, the total RFU changes when compared to the reaction with the CE3-30P are much lower, potentially indicating a partial reaction occurring [Figure 7.5B]. Here, only significant increases above the 2621 RFU obtained from the probe alone was observed at 2.5 nM target with an RFU output of 3042. The CE3-40T looked able to trigger the strand

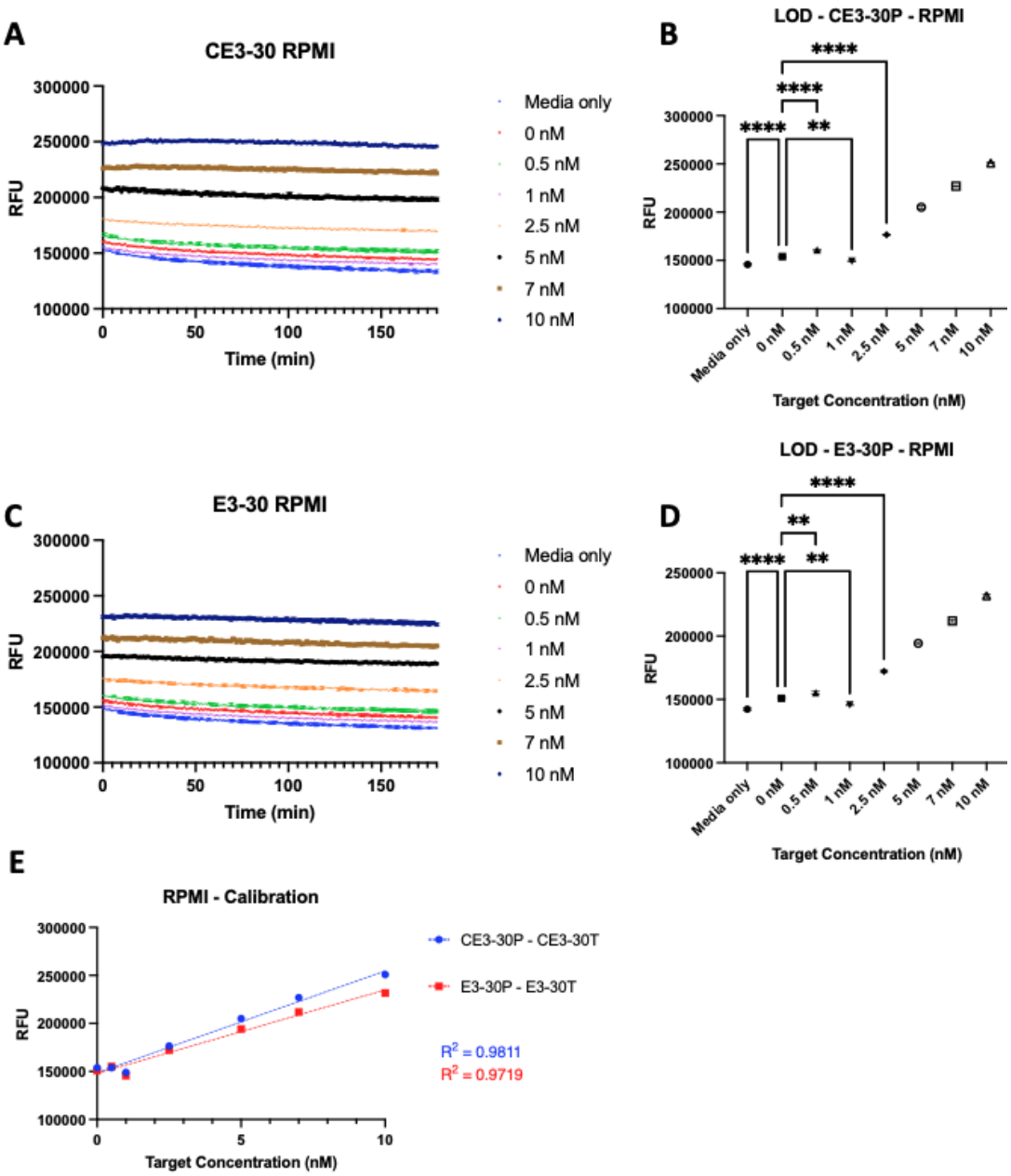


Figure 7.6 – Effects of RPMI cell culture media on calibration and limits of detection on CE3-30P and E3-30P. A. *RFU output from CE3-30P across 180-minute detection window in RPMI media in response to 0 – 10 nM concentration range of CE3-30T alongside RFU output from media alone. Data is representative of 3 individual experiments ± SEM.* ***B.*** *RFU output from CE3-30P in response to CE3-30T in RPMI media at 30 minutes. Data is representative of 3 individual experiments ± SEM. Statistical significance was determined by a One-way ANOVA with Dunnett's multiple comparison test (*, **, *** and **** denote P-values <0.0332, <0.0021, <0.0002 and <0.0001 respectively, ns denotes not significant).* ***C.*** *RFU output from E3-30P across 180-minute detection window in RPMI media in response to 0 – 10 nM concentration range of E3-30T alongside RFU output from media alone. Data is representative of 3 individual experiments ± SEM.* ***D.*** *RFU output from E3-30P in response to E3-30T in RPMI media at 30 minutes. Data is representative of 3 individual experiments ± SEM. Statistical significance was determined by a One-way ANOVA with Dunnett's multiple comparison test (*, **, *** and **** denote P-values <0.0332, <0.0021, <0.0002 and <0.0001 respectively, ns denotes not significant).* ***E.*** *Endpoint RFU taken from 180-minute detection windows for CE3-30P and CE3-30T, and E3-30P and E3-30T in RMPI media. Average RFU from each concentration's endpoint was used to plot standard curve and calculate R2 value. Data is representative of 3 individual experiments ± SEM.*

displacement reaction, showing a significant increase from a baseline of 2450 to 3025 RFU at 0.1 nM target concentration [Figure 7.5C]. As the target length increased in the CE3-60T, the LOD increased to 0.5 nM of target, although a marginal increase of 532 RFU was seen from the

baseline to 0.1 nM target concentration [Figure 7.5D]. Finally, with the fully complementary E3-30T, significant detection above the probe-only baseline was seen at 0.1 nM, as the RFU increased from 4349 to 6437 [Figure 7.5E]. Consistent with the CE3-30P, probe concentration was decreased to 1 nM and gain increased to assess if the lower LOD was below 0.1 nM [Figure 7.5F]. In the case of the E3-30 probe, in-line with previous data, no significant detection above the baseline was seen until the target concentration hit 100 pM with an RFU increase from 687 to 817. Results from these LOD experiments alongside those performed in the CE3-30P indicate that the current simple duplex probe design boasts a LOD of ~0.1 nM as long as target length does not exceed 40 nucleotides.

7.4.4 Probes are able to detect synthetic targets within more convoluted media conditions

With current data for the probes in TNaK buffer showing promising results, steps were taken to check probe stability, LODs and calibration impacts in more complex media with sub-optimal reaction conditions; the first tested being supplemented RPMI full media [Figure 7.6]. Here, the CE3-30P displayed minimal reductions in RFU across the 180-minute time course, with a maximal reduction of 22% seen from the first to last timepoint [Figure 7.6A]. As before, using T:30 data, the LOD was calculated [Figure 7.6B]. Of note here, the media without addition of the probe demonstrated a much higher background RFU of 145782, which increased to 153817 with the addition of the probe. A significant increase above this probe-only reading was observed at 0.5 nM target concentration, returning an RFU of 160858. However, a decrease to 149124 was recorded when the target concentration was increased to 1 nM. RFU then continued to increase in-line with increasing target concentration, indicating that the LOD within RPMI media is around 2.5 nM. Results with the CE3-30P were mirrored in the E3-30P with LODs observed at 2.5 nM [Figure 7.6C&D]. Although impacting the LOD, the change of background media had minimal effects on the probe calibration when compared to that in buffer [Figure 7.6E].

Expanding this further, similar experiments were set up in human serum [Figure 7.7]. Looking at the CE3-30P a much higher initial background RFU was observed starting at 743142 in the media only which saw a decrease over the total time course [Figure 7.7A]. Upon addition of the probe alone, the background increased further to 772267 RFU showing an increase to

777157 by the final timepoint. Of note, within serum, the kinetics of the toehold displacement reaction was captured, which was previously unseen in buffer, implying serum as a background media influencing reaction speed. Also of note, in all experimental arms, the reactions hit their peak before starting to decrease, implying that the probe and target interactions are much less stable within human serum when compared to similar experiments in buffer. With respect to the LODs, the lower target concentrations suggested no significant detection above the probe only background until the target concentration was increased to 2.5 nM [Figure 7.7B]. Similar

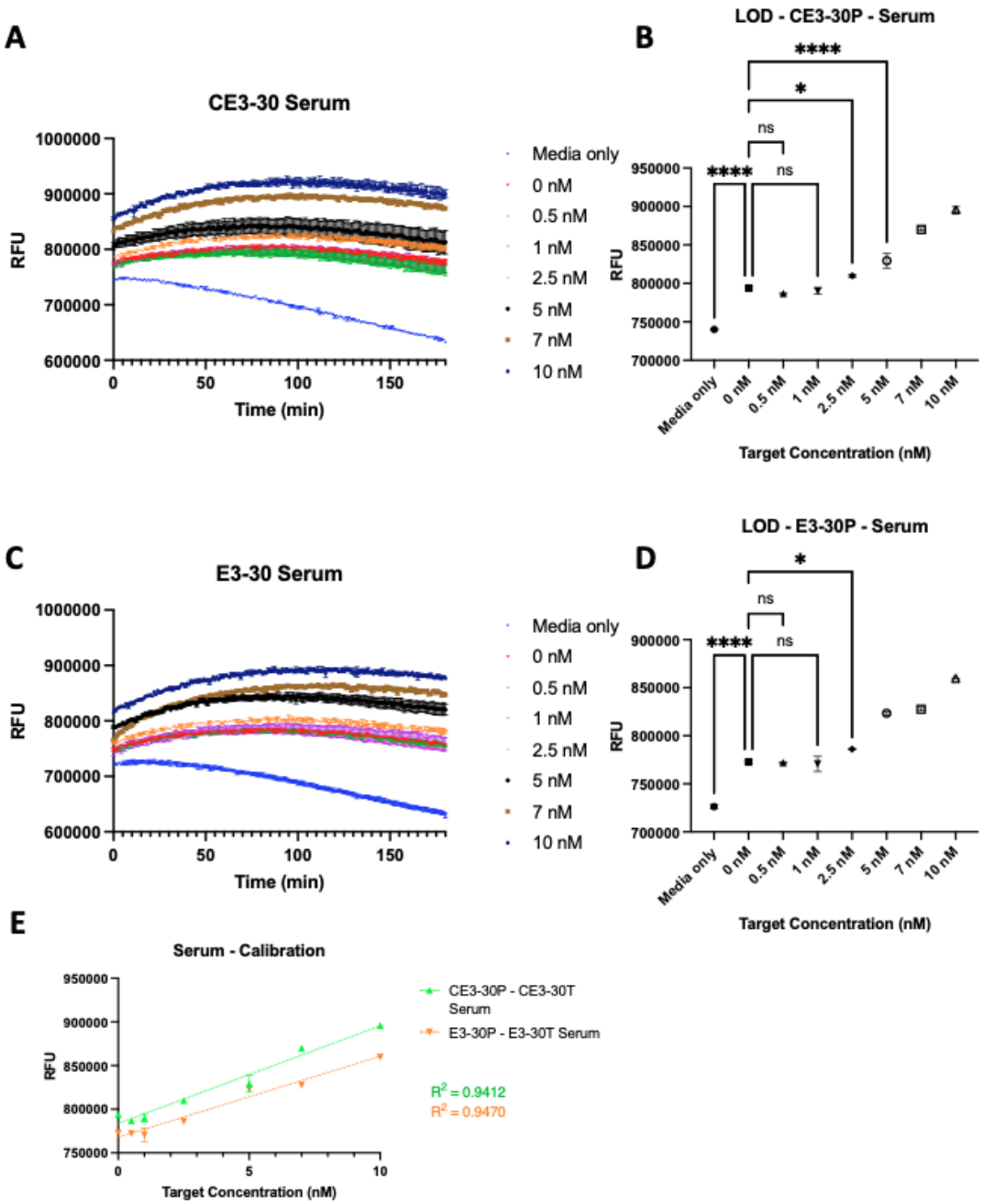


Figure 7.7 – Effects of human serum on calibration and limits of detection on CE3-30P and E3-30P. A. *RFU output from CE3-30P across 180-minute detection window in serum media in response to 0 – 10 nM concentration range of CE3-30T alongside RFU output from serum alone. Data is representative of 3 individual experiments ± SEM.* ***B.*** *RFU output from CE3-30P in response to CE3-30T in serum at 30 minutes. Data is representative of 3 individual experiments ± SEM. Statistical significance was determined by a One-way ANOVA with Dunnett's multiple comparison test (*, **, *** and **** denote P-values <0.0332, <0.0021, <0.0002 and <0.0001 respectively, ns denotes not significant).* ***C.*** *RFU output from E3-30P across 180-minute detection window in serum in response to 0 – 10 nM concentration range of E3-30T alongside RFU output from serum alone. Data is representative of 3 individual experiments ± SEM.* ***D.*** *RFU output from E3-30P in response to E3-30T in serum at 30 minutes. Data is representative of 3 individual experiments ± SEM. Statistical significance was determined by a One-way ANOVA with Dunnett's multiple comparison test (*, **, *** and **** denote P-values <0.0332, <0.0021, <0.0002 and <0.0001 respectively, ns denotes not significant).* ***E.*** *Endpoint RFU taken from 180-minute detection windows for CE3-30P and CE3-30T, and E3-30P and E3-30T in serum. Average RFU from each concentration's endpoint was used to plot standard curve and calculate R2 value. Data is representative of 3 individual experiments ± SEM.*

effects on probe stability and reaction speed were seen in the E3-30P, with the media-only control decreasing across the total time points tested, with all experimental arms taking longer

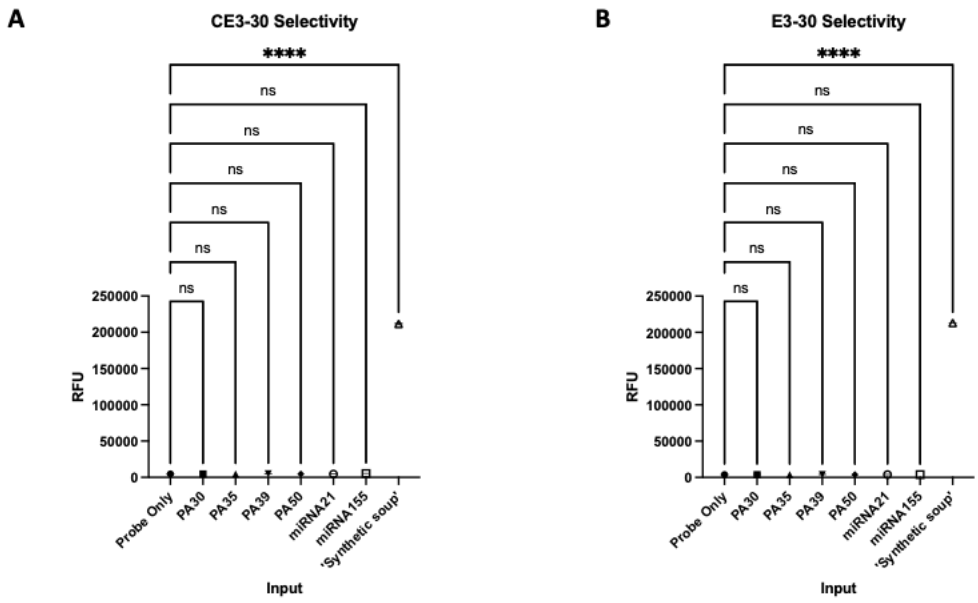


Figure 7.8 – CE3-30 and E3-30 are selective for their respective specific inputs in the presence of high concentrations of non-specific inputs. *Endpoint RFU from 180-minute detection windows from the probe alone, alongside a range of non-specific 10 nM inputs and a final combined 10 nM 'synthetic soup' containing specific targets from the CE3-30P (**A**) and E3-30P (**B**). Data is representative of 3 individual experiments ± SEM. Statistical significance was determined by a One-way ANOVA with Dunnett's multiple comparison test (*, **, *** and **** denote P-values <0.0332, <0.0021, <0.0002 and <0.0001 respectively, ns denotes not significant).*

to hit peak RFU output before ultimately showing a RFU drop-off [Figure 7.7C]. As before, significant increases in RFU over the baseline were seen only at 2.5 nM target input [Figure 7.7D]. Alongside this decrease in LOD seen in both probes, using human serum as a background media also saw reductions in the probe calibration, with R^2 values 0.9879 of in the CE3-30P and 0.9856 in the E3-30P in buffer dropping to 0.9412 and 0.9470 respectively in serum [Figure 7.7E].

7.4.5 Probe selectivity and detection ability within PC cell lines

Until this point, probe selectivity had only been assessed using synthetic targets with varying complementarity to the probes. However, it remains unclear whether the probes would react and trigger strand displacement reactions with non-specific targets sharing no complementarity. To validate the specificity of the probes for the target strands of interest, the probe systems were tested against 10 nM of 6 non-specific synthetic DNA targets, alongside a pool of all these targets containing spiked-in CE3-30T or E3-30T [Figure 7.8]. For the CE3-30P, no fluorescent output from the TMSD reaction was seen with any of the non-specific targets resulting in an RFU ~4700, with a significant increase to 211503 RFU only seen with the CE3-

30T spike-in [Figure 7.8A]. As with the CE3-30P, similar effects were observed in the E3-30P, with significant increases in RFU only seen in response to the E3-30T spike-in [Figure 7.8B].

Finally, both the CE3-30P and E3-30P were tested using PC cell line lysates (normalised to 500 ng input) to see if this basic duplex system was able to detect endogenous AR-FL or AR-V7 mRNA [Figure 7.9]. Here, VCaP, CWR22Rv1-AR-EK and HEK293T whole cell RNA lysates were generated and ran with each respective probe alongside the probe alone, a positive control of probe and target as well as a HEK293T RNA lysate with the synthetic CE3-30T or E3-30T spiked-in. Here, probe concentration was also decreased, as in the LOD experiments, to reduce the amounts of background to better detect RFU changes. Of the cell lines tested, positive detection of the CE3-30P was expected in the CWR22Rv1-AR-EK and VCaP lines, with only positive detection expected in the VCaP cell line for the E3-30P due to the AR-FL amplification present in the line. As before, following probe addition, RFU output was detected over 180 minutes, with the endpoint RFU from each being taken for analysis. Due to the significant increase in RFU over the probe-only background seen in the HEK293T cell line lysates from both probes, potentially due to the increase in material present in the well, these were used as the baseline RFU for comparisons. The CE3-30P displayed no significant detection of the AR-V7 transcript from either cell line lysates tested [Figure 7.9A]. There was, however, an increase seen when the CE3-30T was spiked into an HEK293T RNA lysate which saw an increased RFU when compared to a positive buffer control CE3-30P + CE3-30T, set up with equivalent synthetic target, suggesting that the presence of increased amounts of RNA within the system causes a general increase in background RFU, potentially having implications on the LOD of the system. The effects on both the increased RNA amount on background RFU in the probe only vs the HEK293T samples, and the increase in RFU between the positive E3-30P +E3-30T samples and the spiked HEK293T samples was mirrored using the E3-30P [Figure 7.9B]. Interestingly, there was significant detection seen within the VCaP RNA lysate when tested with this probe, with an RFU returned of 1204 and a background of 884 RFU. Contrasting this, a decrease to 809 RFU was seen in response to the CWR22Rv1-AR-EK RNA lysate, potentially indicating that the E3-30P successfully detected the amplified AR-FL within the VCaP lysate where the CE3-30P did not. Taken together, these results display the potential of these simple duplex probes, not only in sensitive and specific detection of synthetic targets within a range of background media but also indications of their functionality and detection capability within complex whole cell RNA lysates.

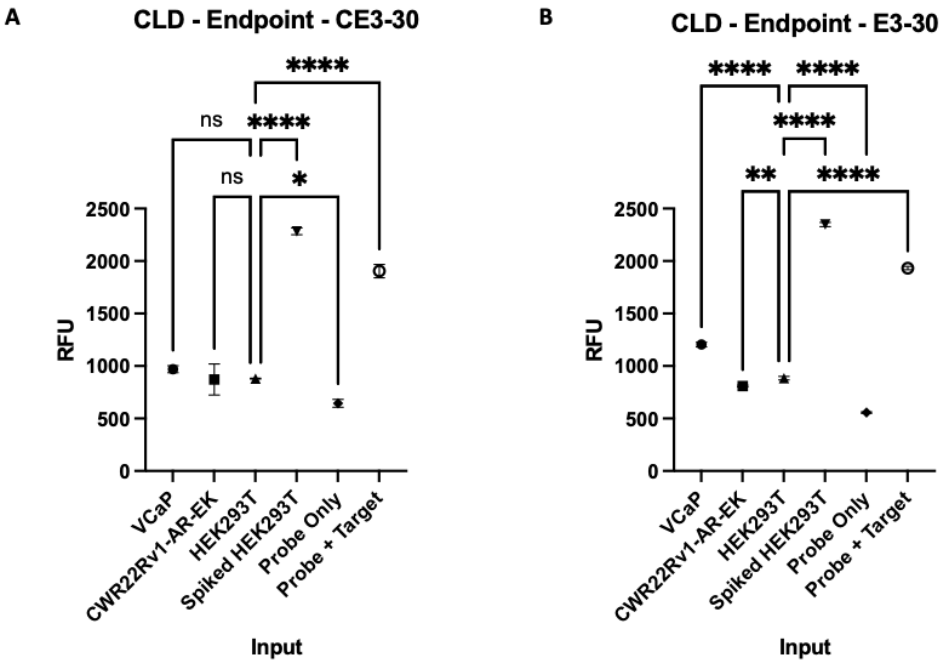


Figure 7.9 – CE3-30P and E3-30P detection capabilities within PC cell line whole RNA lysates. *RFU output at 30 minutes following addition of CE3-30P (**A**) and E3-30P (**B**) to VCaP, CWR22Rv1-AR-EK, HEK293T whole RNA lysates alongside HEK293T RNA lysates spiked with each probes respective specific input as well as a probe only negative control and a probe+target positive control in buffer. Data is representative of 3 individual experiments ± SEM. Statistical significance was determined by a One-way ANOVA with Dunnett's multiple comparison test (*, **, *** and **** denote P-values <0.0332, <0.0021, <0.0002 and <0.0001 respectively, ns denotes not significant).*

7.4.6 Single nucleotide variant detection using SPOP as a candidate biomarker Probe design

With aims to utilise the previous simple duplex probe designs in SNV detection, SPOP was used as a candidate due to the previously presented work showing its importance within PC and potential as a biomarker [Figure 7.10]. Looking at the SPOP mRNA transcript, alongside data from TCGA (TCGA, 2015) denoting common mutations, all *SPOP* loss of function mutations within its MATH domain are caused by single base changes.

To allow for detection of the point mutations present within the SPOP transcript a new approach to probe design was required utilising the idea of probe 'doping' [Figure 7.11]. Through designing the probe to be specific for the SPOP-MT mRNA, a single base mismatch would be introduced upon interaction with the SPOP-WT transcript, slowing the reaction kinetics and increasing discrimination factors. With further 'doping', it is expected that further mismatches will occur with the wild-type (WT) sequence when compared to the mutant (MT) sequence, meaning that the reaction kinetics would be slowed further. The aim here would be to use the slowed kinetics upon SPOP-WT interaction to discriminate between reactions containing SPOP-MT transcripts and those not; ideally to a level that supersedes the %MT transcript presence detectable via qPCR-based techniques. For these experiments, probes

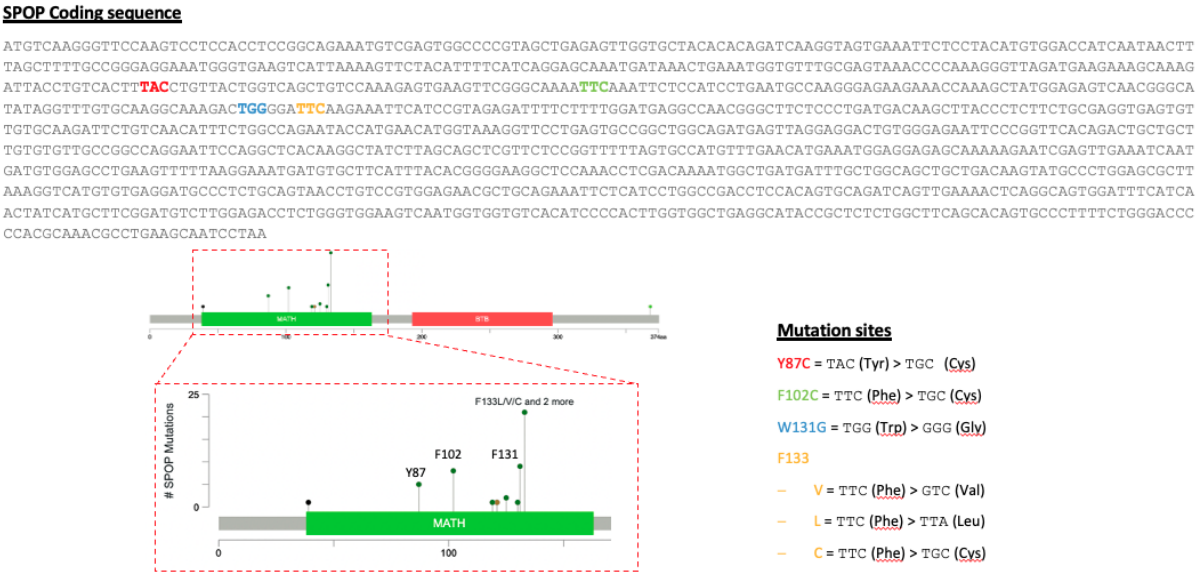


Figure 7.10 - SPOP mutations within the MATH domain are all single base changes. *SPOP DNA coding sequence with common mutation sites from the TCGA highlighted alongside the single base changes at mutation sites that are responsible for the amino acid changes seen in patient data.*

were designed specific for the SPOP-MT transcript with the single base change situated within the 'toehold'.

To decide on which of the common mutations within SPOP to target, the secondary structure of the SPOP transcript was mapped using the NUPack software for DNA/RNA secondary structure prediction [Figure 7.12A]. As expected, the full SPOP transcript formed a large number of secondary structures. To attain a more focal conformation at the target site within the MATH domain of SPOP, this view was simplified to show specifically whether each base was bond or unbound, represented by a period or bracket respectively [Figure 7.12B]. Of the *SPOP* mutation sites, most showed predictions of all bases being fully bound with the exception of the F133-encoding site which showed two open bases. As such, the F133 site was targeted as a candidate for further downstream probe development. With multiple residue changes present within the MATH domain, the F133V mutation was selected based on it showing the most promise following probe design as well as being the most commonly identified *SPOP* mutation PC patients.

Considering the high degree of secondary structures present within the SPOP mRNA as a whole, attempts were made to open the target region through the use of complimentary RNA blockers up- and down-stream of the mutation site. These were designed and the resultant

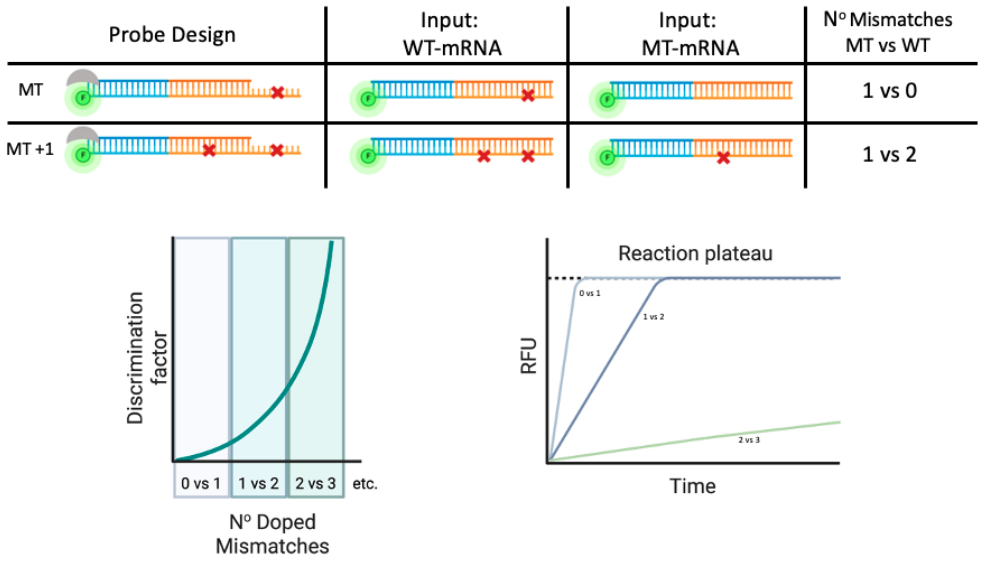


Figure 7.11 – Doping probes with mismatches allows for greater discrimination factors and longer time to reach reaction plateau. *Schematic illustration of probe designs targeted towards the mutant to allow for integrated mismatches when interacting with WT-mRNA, creating differential numbers of mismatches and thus increased discrimination factors which will have effects on probe kinetics. Mismatches can be increased further to allow for enhanced discrimination factors between detection of wild-type and mutant transcripts.*

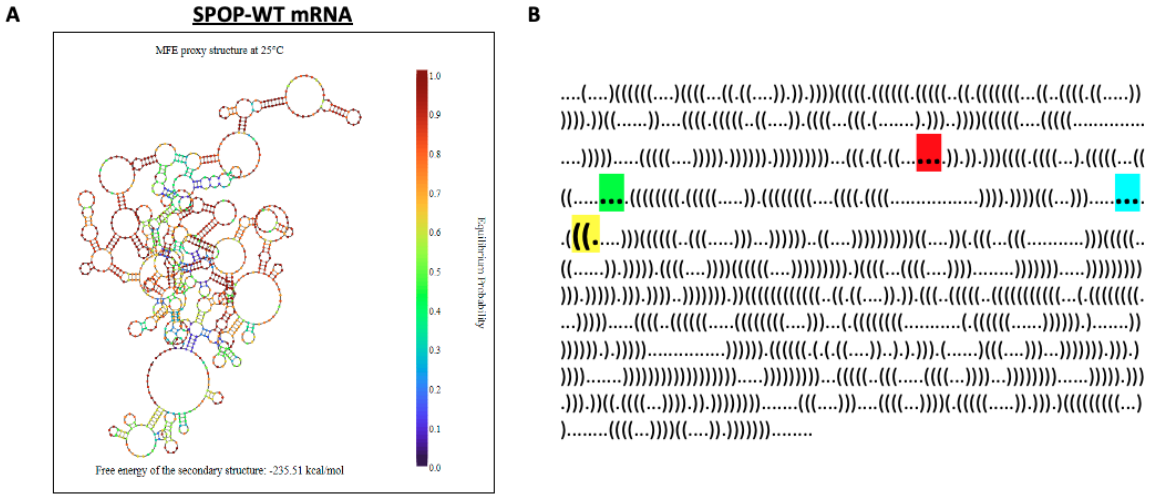


Figure 7.12 – SPOP mRNA forms complex secondary structures that block access to mutation sites. A. *NUPack modelling of secondary structures present within the whole SPOP wild type mRNA.* ***B.*** *Representative SPOP mRNA sequence with conformation at each base annotated.* ***.*** *represents closed conformations and* ***)/(*** *represent bases that are open. SPOP mutation sites at Y87, F102, F131 and F133 are annotated in red, green, blue and yellow respectively.*

predicted structure mapped using NUPack [Figure 7.13]. In the case of the SPOP-WT mRNA, structural predictions showed that the blockers were able to bind around the probe target site although resulted in small secondary structures being formed within this 'open' region caused by the blockers [Figure 7.13A]. When running the structural predictions again, this time using the SPOP-F133V mRNA sequence alongside the blockers, the probe target site was fully opened [Figure 7.13B]. Together, this indicates that the probes may be unable to trigger reactions with the WT SPOP mRNA due to lack of access to the target site, while the reaction could continue unhindered upon the presence of the MT mRNA.

To validate the *in-silico* probe designs, a synthetic system was developed [Figure 7.14]. Here, short 23 nucleotide inputs (MT-IS/WT/IS) which were directly complementary to the probe were tested alongside 150 nucleotide inputs (MT-IL/WT-IL). In each case, inputs were generated for both the MT and WT sequences. Alongside this, the long inputs were annealed to the blockers for both the MT and WT long inputs (MT-IL+Bl and WT-IL+Bl), the predicted structures of which represented that seen when tested with the full SPOP mRNA. To further iterate on previous probe designs, alongside the duplex probes that had been used in the previous AR-V7 mRNA detection work, a second design was implemented utilising a single strand with fluorophore and quencher annealed on the 5' and 3' respectively in a hairpin

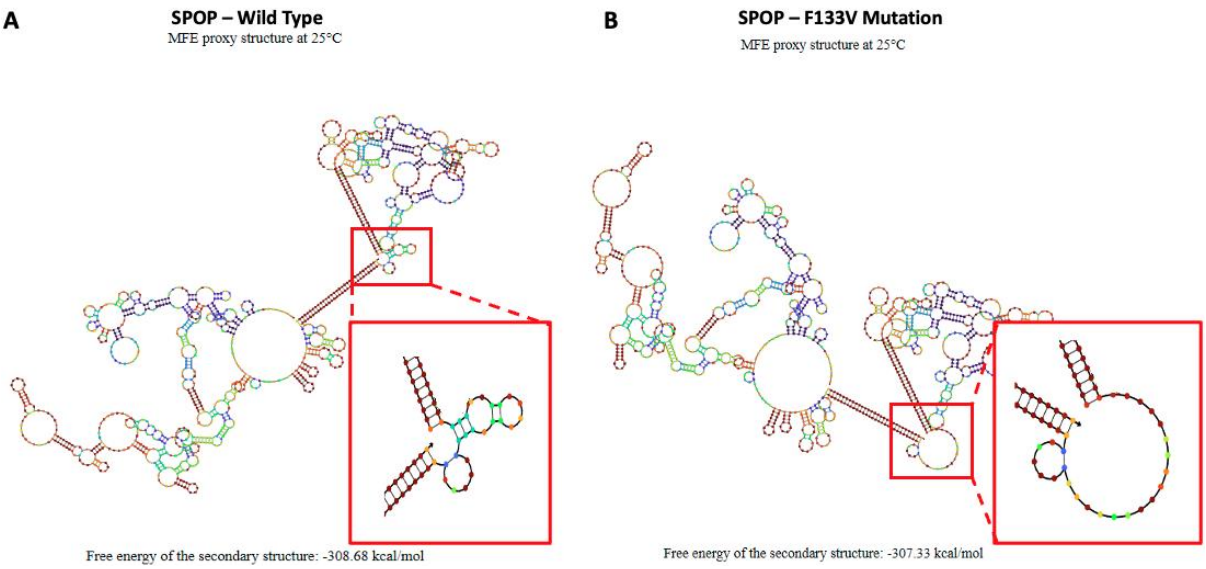


Figure 7.13 – in silico addition of oligonucleotide blockers cause opening of secondary structures in the F133V mutant SPOP mRNA but not the SPOP-WT. *mRNA input into NUPACK with predicted secondary structure formed from SPOP-WT mRNA (**A**) and SPOP-F133V (**B**) mRNA upon the addition of two oligonucleotide blockers. Secondary structure formed around the F133 mutation site is highlighted in red.*

conformation [Figures 7.14B and 7.14C respectively]. As in the duplex probe, target interaction with the toehold would facilitate strand displacement and removal of the quencher from the vicinity of the fluorophore resulting in a fluorescent signal produced. Probes were designed targeting either the SPOP-WT or SPOP-F133V sequences, although due to time constraints, experimental work was only conducted with the SPOP-MT targeting probe.

7.4.7 Initial MT-Duplex and MT-Hairpin probe validations

Firstly, it was critical to validate that the fully complementary MT-S inputs were able to trigger stand displacement reactions with the hairpin and duplex MT-probes. To do this, initially 5 nM of probe was tested against a range of input concentrations [Figure 7.15]. The hairpin probes displayed fast reaction kinetics that, once reaching plateau, fluctuated in signal in all tested concentrations [Figure 7.15A]. Furthermore, with 5 nM probe present in each reaction, the resultant maximum RFU output was attained from all input concentrations above 5 nM. In contrast, this was not the case with the hairpin probe, as at the endpoint of the experiment, the RFU outputs looked to increase as input concentration increased [Figure 7.15B]. This may be due to the design of the hairpin probe, where the 'blocker' strand is tethered to the target interacting with the toehold, and upon fraying of the bases at the ends of the complementary region, displacement of the target and subsequent re-quenching of the fluorophore may take

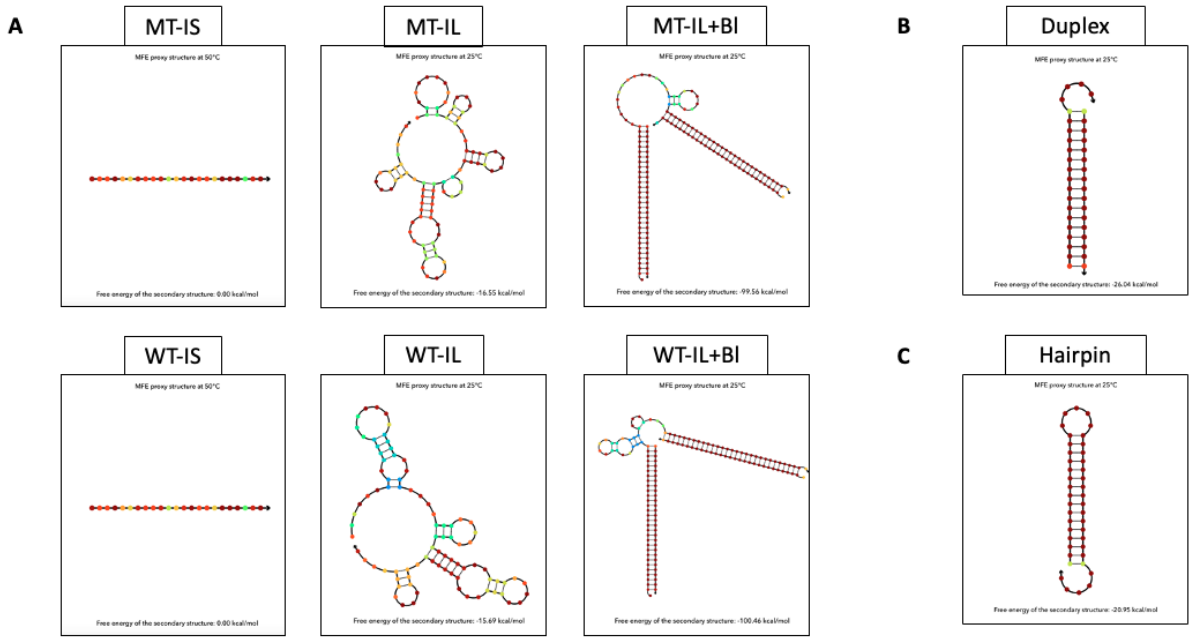
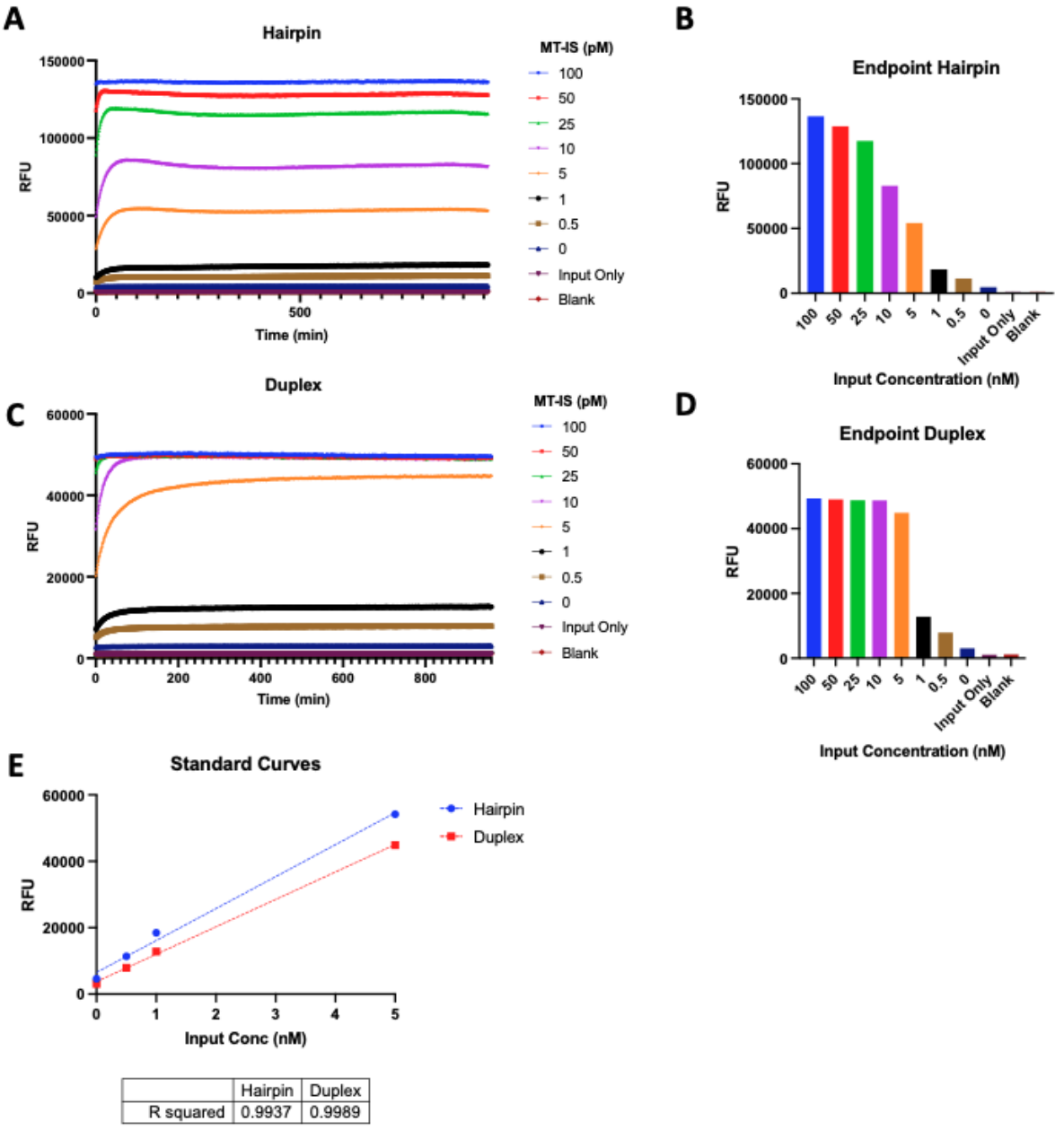


Figure 7.14 - Development of a synthetic system representative of full SPOP mRNA. A. *NUPACK secondary structure predictions of the range of inputs used comprising of either the SPOP-MT sequence or SPOP-WT sequence. IS represents the short input with shares full complementarity with the probes used. IL represents the long inputs that are able to form similar secondary structures to that predicted from the full SPOP mRNA and IL+Bl are the long inputs with blockers annealed. Also shown are NUPACK secondary structures representing the both the Duplex (**B**) and the Hairpin (**C**) probes used in downstream experiments, both targeting the SPOP-MT sequence.*

place. This indicates that, unlike previously seen in the duplex probes, this reaction does not favour full strand displacement and is actually in an equilibrium state with the presence of a backwards reaction resulting in the fluctuations in RFU output across the time course, thus requiring a much greater amount of input to fully trigger the system. As observed previously in the AR-V7 detection using the duplex probe design, no fluctuations in signal were seen across the total time-course tested [Figure 7.15C]. Moreover, assessing the endpoint RFU outputs from each of the input concentrations tested, all reactions above 5 nM produced a similar maximum RFU output at 46000 RFU. When taking these lower end input concentrations and plotting the standard curves for each, the R^2 values were 0.9937 and 0.9989 for the hairpin and duplex, respectively [Figure 7.15E]. From these initial experiments, the MT-Duplex was taken as the candidate probe for more in-depth downstream validation experiments, although experiments were continued using the MT-Hairpin probe.

To further validate the MT-Duplex probe and to help inform the probe concentration being used in downstream experiments, a similar experiment was set-up utilising a lower probe concentration of 2 nM [Figure 7.16]. Here, across the timepoints, all reactions containing inputs showed stable final outputs, where the input only produced RFU in-line with the



	Hairpin	Duplex
R squared	0.9937	0.9989

Figure 7.15 - Determination of optimum probe concentration and maximal RFU output for downstream experiments. A. *RFU output from MT-IS input (0 – 100 nM) concentration range alongside input only background added to 5 nM of the Hairpin probe alongside blank reads across a 960-minute detection window. Data is representative of 3 individual experiments ± SEM.* ***B.*** *Endpoint RFU from MT-IS input (0 – 100 nM) concentration range alongside input only background added to 5 nM of the Hairpin probe alongside blank reads. Data is representative of 3 individual experiments ± SEM.* ***C.*** *RFU output from MT-IS input (0 – 100 nM) concentration range alongside input only background added to 5 nM of the Duplex probe alongside blank reads across a 960-minute detection window. Data is representative of 3 individual experiments ± SEM.* ***D.*** *Endpoint RFU from MT-IS input (0 – 100 nM) concentration range alongside input only background added to 5 nM of the Duplex probe alongside blank reads. Data is representative of 3 individual experiments ± SEM.* ***E.*** *Standard curve generated from both Duplex and Hairpin probes from 0 – 5 nM MT-IS concentration range alongside R2 calculations generated from interpolated best-fit. Data is representative of 3 individual experiments ± SEM.*

background or 'blank' readings [Figure 7.16A]. The baseline RFU increased again with the addition of the probe alone, however once inputs were added, evidence of the strand displacement reaction occurring was evident, with clear kinetic curves attained albeit slower due to the reduced probe concentration. The time for each of these reactions to reach a

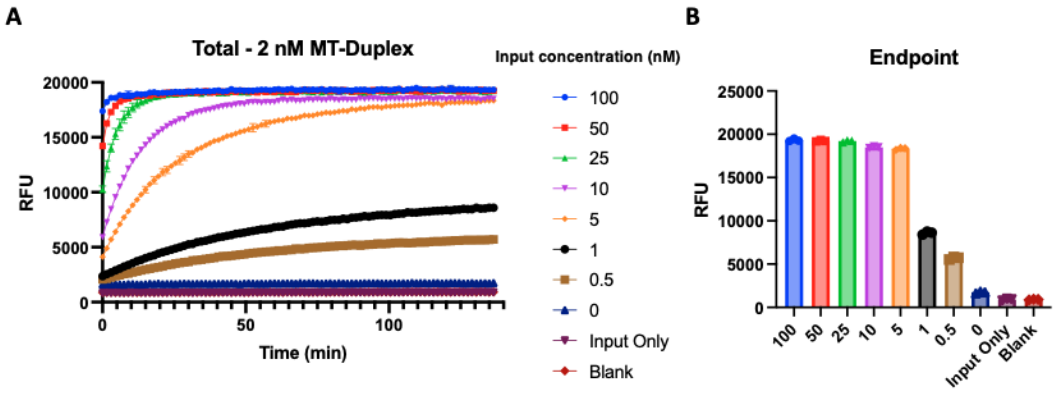


Figure 7.16 – Reducing duplex probe concentration shows more favourable response to MT-IS input concentration. A. *RFU output from MT-IS input (0 – 100 nM) concentration range alongside input only background and blank reads added to 2 nM of the Duplex probe across a 180-minute detection window across a 180-minute detection window. Data is representative of 3 individual experiments ± SEM.* ***B.*** *Endpoint RFU from MT-IS input (0 – 100 nM) concentration range alongside input only background and blank reads added to 2 nM of the Duplex probe. Data is representative of 3 individual experiments ± SEM.*

plateau was largely dependent on input concentration. When looking at the endpoint RFU from each of the of the input concentrations, all concentrations over 5 nM produced the maximal RFU signal, with proportional reductions in signal output observed in those under 5 nM. For example, 1 nM input returned 8762 RFU which is around half of the maximal output from the 5 nM input of 18337 [Figure 7.16B]. With this data, combined with the previous runs at 5 nM probe concentration, 2 nM was taken further as the probe concentration used in all downstream experiments.

7.4.8 Blocker sequences do not trigger MT-Duplex probes

Satisfied that the MT probes were working as expected, it was necessary to confirm that the blockers themselves did not trigger the probes [Figure 7.17]. To do this, large amounts of unbound blockers were added to 2 nM MT-duplex probes alongside a concentration range of MT-S inputs to act as a control and confirm the probes were still functioning as expected. As before, at 10 nM and 100 nM MT-S input concentration produced a maximal signal of 102603 and 105374 respectively, with the 1 nM input producing 40481. Importantly, the data suggested no evidence of a strand displacement reaction taking place in response to addition of the blocker strands alone regardless of the input concentration; with no additional RFU over the probe-only control observed when correcting for the 'blank' (probe alone) readings [Figure 7.17B].

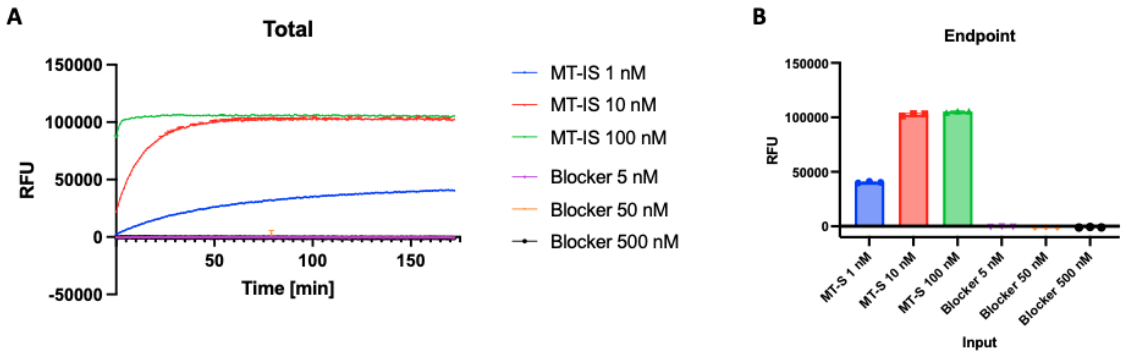


Figure 7.17 – Oligonucleotide blockers used to open secondary structures do not trigger Duplex probes. A. *RFU output from MT-IS input (1 – 100 nM) concentration range alongside Blocker (5- 500 nM) concentration range added to 2 nM of the Duplex probe across a 180-minute detection window. Data is representative of 3 individual experiments ± SEM.* ***B.*** *Endpoint RFU from MT-IS input (0 – 100 nM) and Blocker (5 – 500 nM) concentration range added to with 2 nM of the Duplex probe. Data is representative of 3 individual experiments ± SEM.*

7.4.9 Calculation of discrimination factors of both MT-Duplex and MT-Hairpin probes

To assess the effects of each of the synthetic inputs on the MT-Duplex probe, experiments were set up using 2 nM probe alongside 1 nM of either the WT or MT input and a positive control of 10 nM MT input [Figure 7.18]. This was performed for each of the IS, IL and the IL+Bl inputs in attempts to determine the ability of the MT-Duplex probe to selectively detect the MT over WT input, and assess the effects of the blockers in overcoming the secondary structure present in the long inputs and ability of facilitating probe access. Firstly, the RFU output was tracked over the 960-minute time-course for each of the inputs [Figure 7.18A-C]. With regards to the IS inputs, a maximal RFU output of 34000 was attained, with clear kinetic curves seen in response to the addition of 1 nM of the WT input, reaching a plateau of ~10500 RFU after 180 minutes [Figure 7.18A]. Upon addition of 1 nM WT input, a minimal RFU increase was seen form 3693 at 0 minutes to 5077 at the end of the experiment, ultimately implying that the strand displacement reaction is taking place, however at a much slower rate. A similar trend was seen when input secondary structures were introduced in the IL inputs [Figure 7.18B]. Here, a slow reaction was observed with addition of the IL-MT input over the total time-course, with a total increase in RFU from 3444 to 8090 [Figure 7.18B]. The reaction was further slowed with the addition of the IL-WT where a 221 total RFU increase was observed. As expected, the IL-positive input was able to facilitate a complete reaction, implying that the

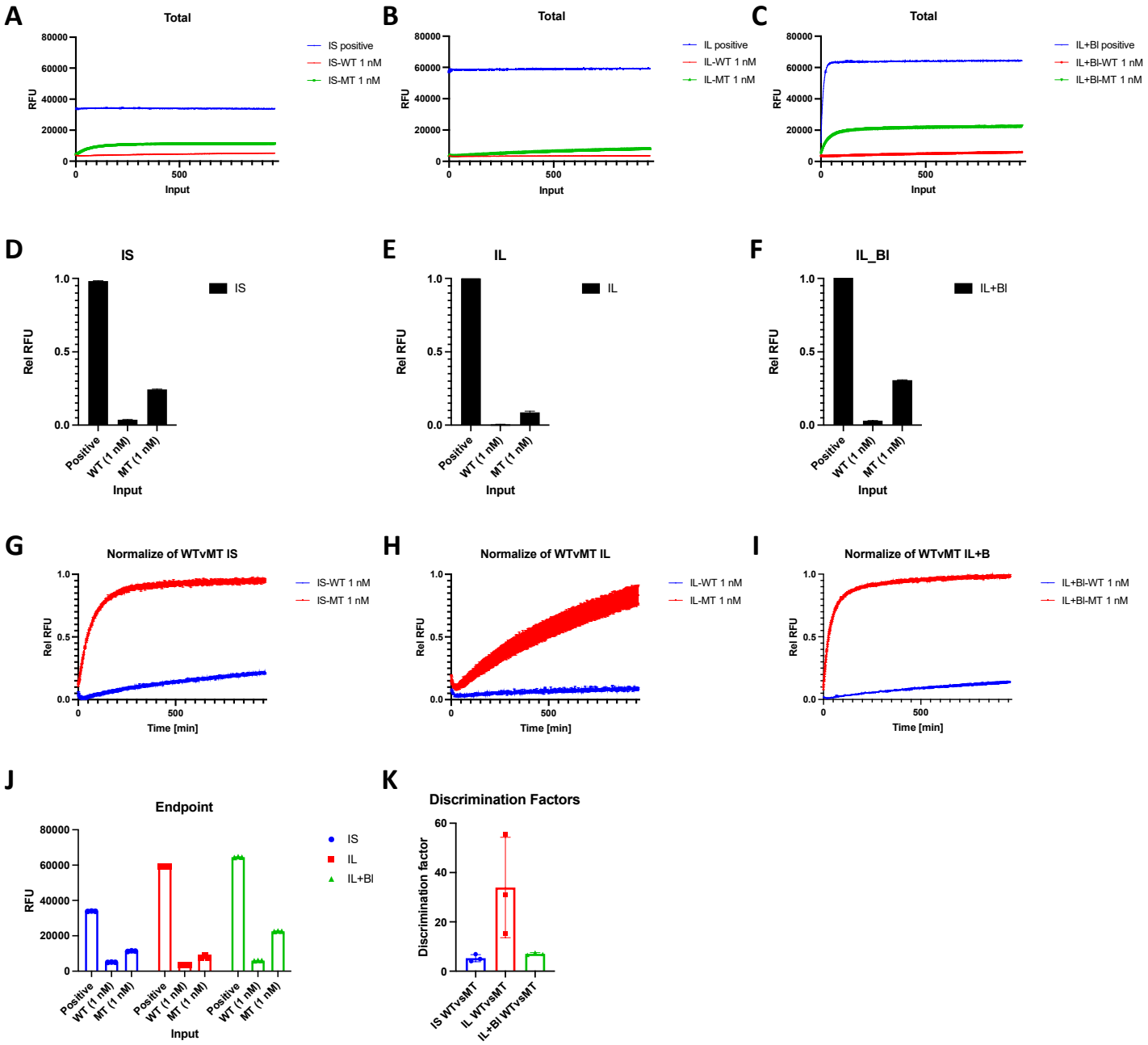


Figure 7.18 – Determination of MT-Duplex discrimination factor for MT-inputs over WT-inputs. *RFU output from 1 nM MT- and WT- IS (**A**) IL (**B**) and IL+Bl (**C**) inputs alongside a 10 nM MT input positive controls of the Duplex probe window across a 960-minute detection window. Data is representative of 3 individual experiments ± SEM. Endpoint RFU from IS (**D**), IL (**E**) and IL+Bl (**F**) inputs normalised to positive controls. Data is representative of 3 individual experiments ± SEM. Normalised RFU output across 960-minute detection window reaction kinetics for MT- and WT- IS (**G**), IL (**H**) and IL+Bl (**I**) inputs. Data is representative of 3 individual experiments ± SEM. **J**. Endpoint RFU from each of the inputs tested above alongside positive controls. Data is representative of 3 individual experiments ± SEM. **K.** Discrimination factor calculations from endpoint RFU for MT inputs over WT inputs. Data is representative of 3 individual experiments ± SEM.*

1 nM MT and WT inputs would eventually reach completion, although the length of time necessary would be considerable longer. With the addition of blockers to the long inputs, results showed clear evidence of strand displacement reactions occurring and reaching completion within the 960-minute time-course [Figure 7.18C]. Here, when the MT-Duplex probes were overloaded within the IL+Bl positive input, the reaction was evident, with the

reaction plateau being reached at 60 minutes. In response to the 1 nM IL+Bl-WT input, a marginal increase from 3503 to 6005 RFU over the total reaction window was observed. Upon the addition of the IL+Bl-MT input, a reaction plateau is reached after 140 minutes with a higher RFU of ~21000 attained. Of note here, in response to 1 nM of either the WT/MT inputs, different RFUs were returned from the IS-MT and IL+Bl-MT and different maximum RFUs were demonstrated from the positive controls.

With this in mind, to better compare the specific effects of these inputs, the data from these runs was normalised to their internal positive control and the endpoints plotted [Figure 7.18D-F]. With regards to the IS inputs, 24% and 0.04% of the maximal RFU output was returned from the IS-MT and IS-WT, respectively. By the end of the time-course the IL-MT displayed 0.09% of the RFU produced by its positive control, where the IL-WT saw 0.01%. Finally, with the addition of blockers in the IL+Bl-MT and IL+Bl-WT inputs, an increase to 31% and 0.03% endpoint RFU respectively was seen. Of note, following normalisation, RFU output remained increased in the IL+Bl-MT input compared to the IS-MT input. Further analysis of this data through normalisation to the 1 nM MT inputs, direct comparisons between the effects of the MT/WT inputs can be made across the IS, IL and IL-Bl versions of each [Figure 7.18G-I]. Through this normalisation, clear similarities in response to 1 nM of the MT inputs was observed between the IS and IL+Bl inputs, which contrast with the much slower kinetic rate and variance in the IL-MT. In concordance with this, a similar response to the WT inputs was seen in the IS and IL+Bl systems with the IS showing an increased relative RFU across all timepoints, potentially due to the lack of secondary structures that are present in the IL+Bl. Although hard to discriminate, a minimal increase in RFU was detected in response to the addition of the IL-WT over the total time-course, further implying that the strand displacement reaction is taking place although at a much slower rate than in the IL-MT inputs. Finally, the non-normalised RFUs were taken at the endpoint of the experiments and discrimination factors calculated for the probe's selectivity for the MT input vs the WT input [Figure 7.18J-K]. This was calculated by dividing the increase in RFU seen over the total time-course in response to the MT version of the input by the same increase seen in response to the WT input, essentially giving a fold change representation of the preference for the MT strand over the WT. Here, the discrimination factor calculated for the IS inputs was 5.3, with an increase to 34.0 calculated for the IL inputs and 7.1 with the IL+Bl inputs. Taken together, these data confirm the hypothesis postulated during probe development, where secondary structures present in the

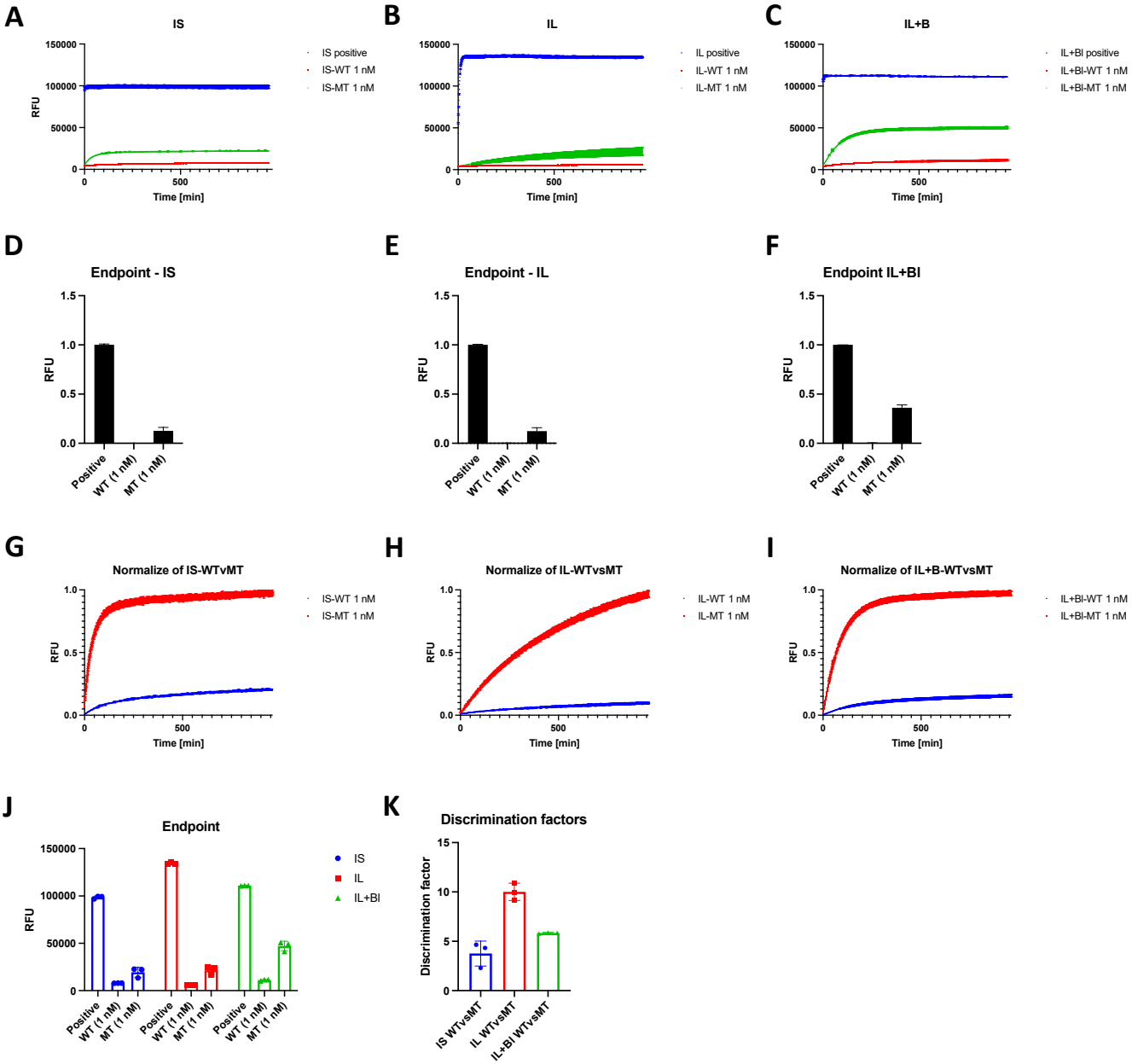


Figure 7.19 - Determination of MT-Hairpin discrimination factor for MT-inputs over WT-inputs. *RFU output from 1 nM MT- and WT- IS (**A**) IL (**B**) and IL+Bl (**C**) inputs alongside a 10 nM MT input positive controls of the Hairpin probe window across a 960-minute detection window. Data is representative of 3 individual experiments ± SEM. Endpoint RFU from IS (**D**), IL (**E**) and IL+Bl (**F**) inputs normalised to positive controls. Data is representative of 3 individual experiments ± SEM. Normalised RFU output across 960-minute detection window reaction kinetics for MT- and WT- IS (**G**), IL (**H**) and IL+Bl (**I**) inputs. Data is representative of 3 individual experiments ± SEM. **J**. Endpoint RFU from each of the inputs tested above alongside positive controls. Data is representative of 3 individual experiments ± SEM. **K.** Discrimination factor calculations from endpoint RFU for MT inputs over WT inputs. Data is representative of 3 individual experiments ± SEM.*

IL inputs would impact the MT duplex probe's ability to bind and that these can be overcome with the addition of blocker strands to open this secondary structure and allow for stable strand displacement reactions to take place this enabling an increase in discrimination between the IS and IL+Bl inputs.

Next, experiments were conducted to assess the ability of the MT-Hairpin probe to selectively detect MT compared to WT inputs within the synthetic system with the ultimate goal of comparing against the previous data from the duplex probes [Figure 7.19]. Similar experiments as before were conducted with 1 nM of wither the WT or MT form of the IS, IL or IL+Bl inputs alongside an internal positive control with 10 nM MT input. In the IS input, the maximum RFU detected from the positive control was 99129 RFU; with the 1 nM IS-MT input reaching a plateau of ~22000 after 260 minutes and the IS-WT showing a minor increase from 4065 to 7893 RFU across the total time-course [Figure 7.19A]. The IL inputs showed a greater max RFU at 136057, but did not show evidence of a complete reaction in either the WT or MT inputs across all the time-course tested, although a steady increase was observed from 3971 to 21650 RFU by the endpoint in response to 1 nM MT-IL [Figure 7.19B]. In contrast, upon addition of the RNA blockers in the IL+Bl inputs, the max RFU of the system decreased to 112851 RFU, although a clear plateau of ~48000 RFU was detected after 340 minutes in response to 1 nM IL+Bl-MT inputs [Figure 7.19C]. As in the IS system, the IL+Bl-WT showed a slight increase during the experiment from 3976 to 11192, although no clear plateau can be seen. Again, to compare the IS, IL and IL+Bl systems against each other, the total data was normalised to the internal positive control and relative RFU at the experimental endpoint quantified [Figures 7.19D-F]. Here, the IS-MT made up 20% of the maximum RFU, with the IS-WT input returning 8% [Figure 7.19D]. Generally seeing a decrease compared to the IS inputs, the IL-WT and IL-MT inputs saw 4% and 16% of the maximum RFU output, respectively [Figure 7.19E]. In contrast to this, addition of the blockers in the IL+Bl-MT saw an increase to 42% of the maximum RFU retuned, while the IL+Bl-WT made up 10% relative to the positive control [Figure 7.19F]. As before, RFU was normalised to the 1 nM MT inputs for each of the systems to allow for comparison of probe selectivity across the IS, IL and IL+Bl systems [Figure 7.19G-I]. Here, clear differences in the kinetic profiles of the MT hairpin probe can be observed between the IS and IL systems, indicating again that the secondary structure present within the IL inputs is causing a much slower strand displacement reaction. Furthermore, this is overcome upon the addition of the blockers which facilitate the opening of this region of secondary structures and result in a kinetic output similar to that of the IS inputs, which have no secondary structures. For the non-normalised RFU output at the endpoint of the experiments in response to 1 nM MT input, the IL+Bl demonstrated 47176 RFU while the IS and IL returned 19413 and 21650 RFU, respectively [Figure 7.19J]. Moreover, minimal differences in RFU output was attained from the 1 nM WT inputs for all systems; with the 7893

RFU for the IS, 5857 RFU for the IL and 11192 RFU for the IL+Bl. Finally, the discrimination factors for each of the systems were calculated as before by dividing the increase in RFU across the total time-course in the MT by the increase in the WT reaction [Figure 7.19K]. The IL system provided the highest discrimination factor of 10, with the IS and IL+Bl systems displaying lower values at 3.8 and 5.8, respectively, due to the much smaller RFU increase seen in response to the 1nM IL-WT than in the IS and IL+Bl systems.

7.4.10 Determining probe %MT input sensitivity

As previously described, current methods of SNV detection via RT-qPCR boast sensitivity in relation to the percentage of MT 'transcripts' able to be detected amongst WT transcripts. In an attempt to create comparable outputs, several inputs were generated containing different ratios of MT:WT IL+Bl inputs, all at 1 nM total input concentration and all were added to 2 nM of either the hairpin or duplex MT targeting probes before RFU detection across a 960-minute time-course [Figure 7.20]. Firstly, the RFU outputs from the total 960-minute window was assessed using the MT-Duplex probe [Figure 7.20A]. As expected, across all inputs tested, excluding the 100% MT input, no reaction reached a plateau indicating in-complete reactions across the time range tested. This is potentially indicating that the initial increases in RFU are from the much quicker strand displacement with the MT strands within each input and the increases seen after this being the slower interactions with the WT strands. Of note here, all reactions looked to finish the fast kinetic phase at the start of the time-course around 180 minutes, meaning time to reach plateau may not be used as a marker for significant detection in this instance.

To assess if positive detection was displayed from the range of inputs, a significant increase above the baseline in RFU at a range of timepoints during the fast kinetic phase was used to denote a positive readout [Figure 7.20B&C]. Here, due to an increase in RFU observed upon interaction with the WT input, the 0:100 input was used as a baseline rather than the probe alone. After 30 minutes, a significant increase from 3592 RFU in the 0:100 input to 3761 RFU in the 0.5:99.5 input was detected, which saw further increases as the percentage of MT input increased [Figure 7.20B]. Again, a significant increase was seen after 60 minutes from 3860 to 4057 in response to the 0:100 and 0.5:99.5 inputs, respectively [Figure 7.20C]. Finally, at the endpoint of the total time-course, the final RFU for each of the inputs was plotted as %MT

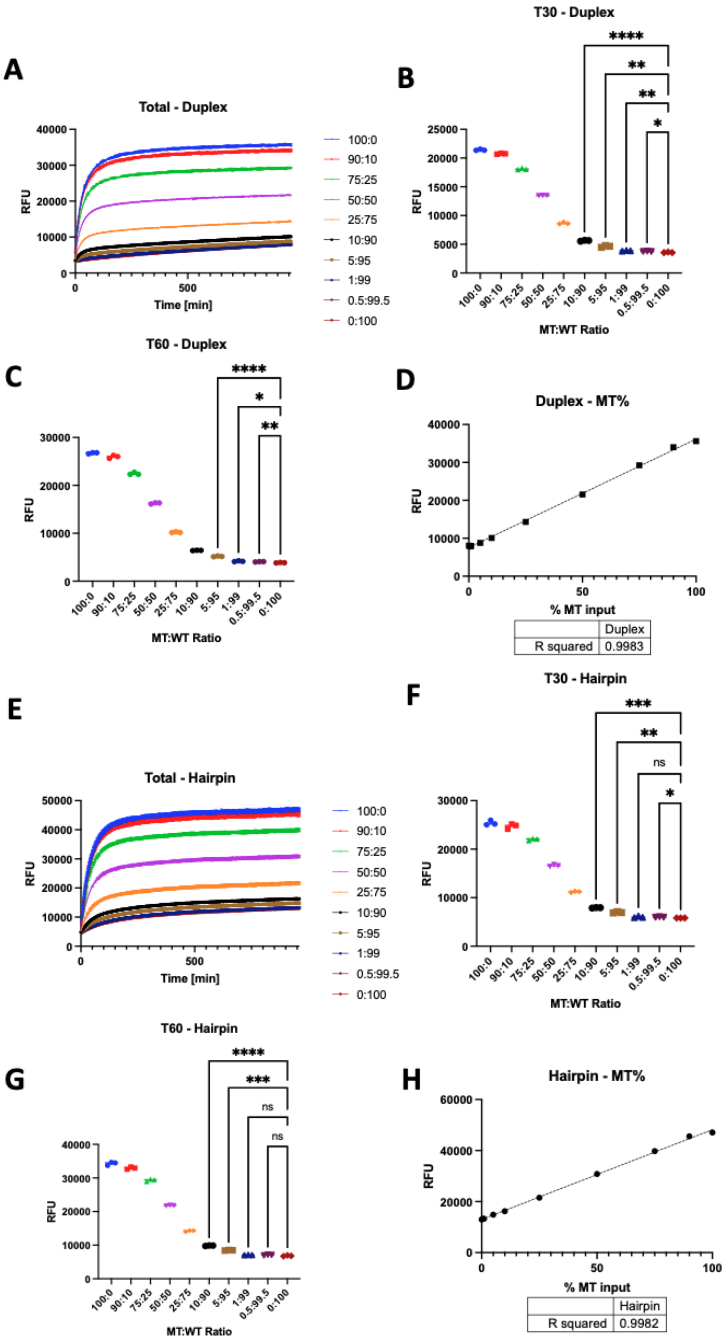


Figure 7.20 – Detection sensitivity of Duplex and Hairpin probes. A. *RFU output from a range of MT+Bl:WT+Bl inputs at 1 nM total input concentration added to 2 nM Duplex probe across a 960-minute detection window. Data is representative of 3 individual experiments ± SEM. RFU taken from 30 minutes (**B**) and 60 minutes (**C**) taken from range of MT+Bl:WT+Bl inputs added to 2 nM Duplex probe. Data is representative of 3 individual experiments ± SEM. Statistical significance was determined by a One-way ANOVA with Dunnett's multiple comparison test (*, **, *** and **** denote P-values <0.0332, <0.0021, <0.0002 and <0.0001 respectively, ns denotes not significant).* ***D.*** *Standard curve interpolated with associated R2 value from the range of MT+Bl:WT+Bl inputs added to 2 nM Duplex probe at the experimental endpoint. Data is representative of 3 individual experiments ± SEM.* ***E.*** *RFU output from a range of MT+Bl:WT+Bl inputs at 1 nM total input concentration added to 2 nM Hairpin probe across a 960-minute detection window. Data is representative of 3 individual experiments ± SEM. Data taken from 30 minutes (**F**) and 60 minutes (**G**) taken from range of MT+Bl:WT+Bl inputs added to 2 nM Hairpin probe. Data is representative of 3 individual experiments ± SEM. Statistical significance was determined by a One-way ANOVA with Dunnett's multiple comparison test (*, **, *** and **** denote P-values <0.0332, <0.0021, <0.0002 and <0.0001 respectively, ns denotes not significant).* ***H.*** *Standard curve interpolated with associated R2 value from the range of WT+Bl:MT+Bl inputs added to 2 nM Hairpin probe at the experimental endpoint. Data is representative of 3 individual experiments ± SEM.*

input to assess if the MT-Duplex probe could be used in %MT transcript quantification if the amount in the original input was unknown through use of a standard curve [Figure 7.20D].

With the linear regression fit from the data presented returning an R^2 value of 0.9983, an argument could be made for faithful %MT quantification from an unknown sample, so long as there is greater than 0.5% total MT concentration.

To see if the hairpin probe was able to recapitulate that seen in the MT-Duplex probe, the same experimental set-up was implemented using the MT-Hairpin probe [Figure 7.20E-H]. Assessing the full kinetic curves produced from the MT-Hairpin probes, the incomplete reaction seen in the MT-Duplex probe was mirrored, although a much greater variance between RFU output at each timepoint was observed [Figure 7.20E]. As before, two individual timepoints were taken within the fast kinetic phase of the strand displacement reactions, initially at 30 minutes where no clear significant detection above the 100% WT background was seen until %MT increased to 5% [Figure 7.20F]. At the 60-minute time-point, a background RFU of 6851 was observed and a significant increase to 8465 RFU was only seen at 5% MT input [Figure 7.20G]. Finally, the endpoint RFU was taken and plotted as %MT input to generate a standard curve with the aim to investigate the MT-Hairpin probes ability to quantify %MT transcripts, returning an R^2 value of 0.9982.

Finally, to further validate taking single time-points during the fast kinetic phase to determine positive read-outs from the assay, and assess the MT-Duplex probes capabilities for rapid %MT quantification, a final experiment was conducted which only captured the RFU output across the initial 30 minutes of the strand displacement reaction using 2 nM MT-Duplex and a range of 1 nM inputs each harbouring different rations of MT:WT IL+Bl [Figure 7.21]. As before, each input resulted in evidence of strand displacement reactions taking place with an increase in RFU across 30 minutes, with the rate and amount of increase seen by the endpoint dependant on the %MT within the input [Figure 7.21A]. Importantly, significant increases in end-point RFU were detected above the background RFU as before [Figure 7.21B]. To better assess if a significant increase above the background RFU can be seen across the total 30-minute window, data was normalised to the 100:0 %MT:WT input at every timepoint and the datapoints for relative RFU for the 0.5:99.5 inputs and 0:100 inputs calculated [Figure 7.21C]. When assessing the last 15 minutes of the detection window, a significant increase was observed from a mean of 0.19 relative RFU from the background to 0.21 relative RFU for 0.5% MT input (when extended further to the total data for the full 30 minutes similar trends are seen). Together, these data suggest that the MT-Duplex probe is the best design for use in SNV detection, and

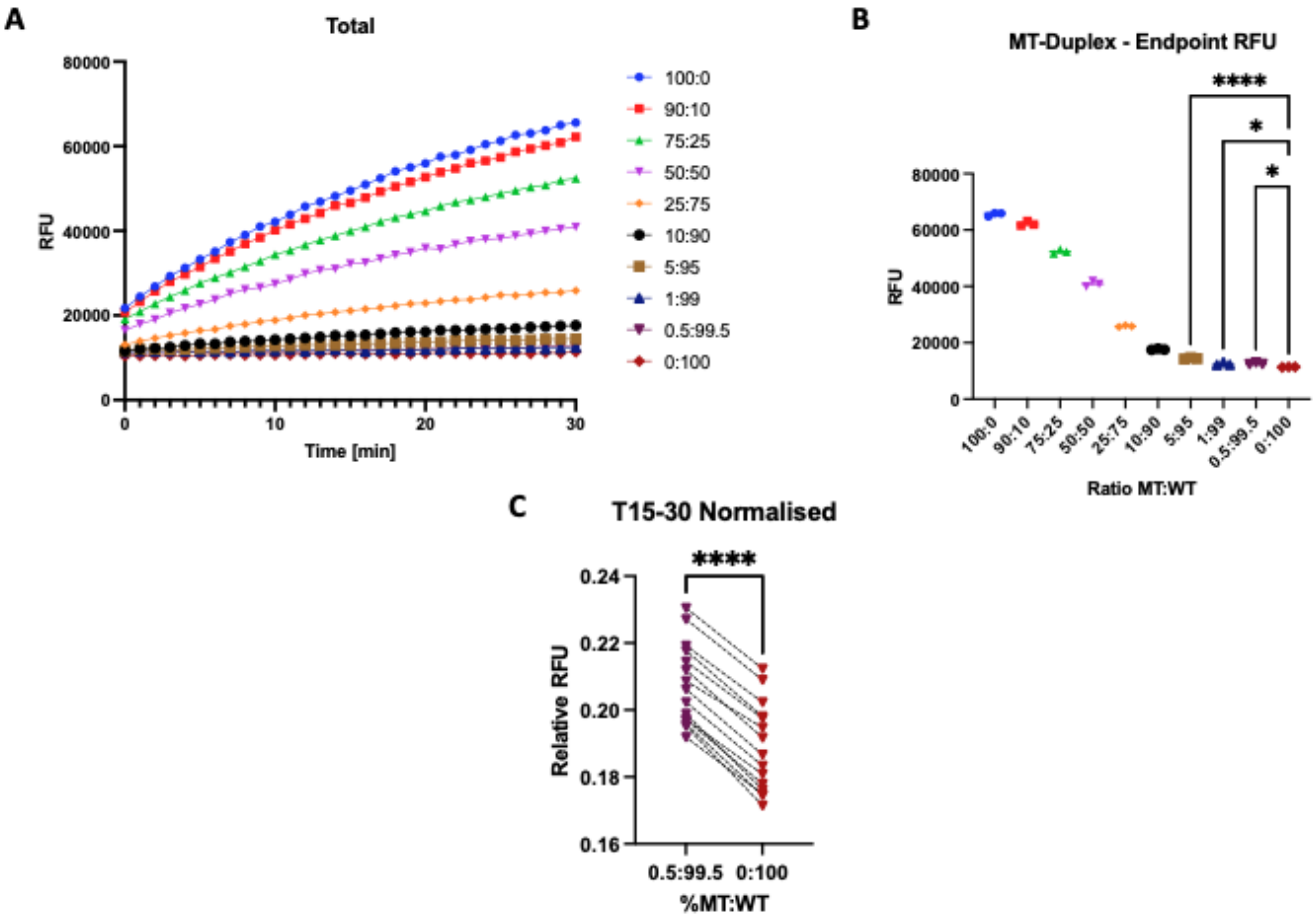


Figure 7.21 – Duplex probes are able to detect 0.5% MT input amongst WT inputs. A. *RFU output from a range of MT+Bl:WT+Bl inputs at 1 nM total input concentration added to 2 nM Duplex probe across a 30-minute detection window. Data is representative of 3 individual experiments ± SEM.* ***B.*** *RFU taken from experimental endpoint in response to a range of MT+Bl:WT+Bl inputs added to 2 nM Duplex probe. Data is representative of 3 individual experiments ± SEM. Statistical significance was determined by a One-way ANOVA with Dunnett's multiple comparison test (*, **, *** and **** denote P-values <0.0332, <0.0021, <0.0002 and <0.0001 respectively, ns denotes not significant).* ***C.*** *RFU from each timepoint following 15 minutes to experimental endpoint in response to 0.5:99.5 and 0:100 MT+Bl:WT+Bl added to 2 nM Duplex probe. Data is representative of 3 individual experiments ± SEM. Statistical significance was determined by T test (*, **, *** and **** denote P-values <0.0332, <0.0021, <0.0002 and <0.0001 respectively, ns denotes not significant).*

that within the synthetic system presented here, the probes are capable of more sensitive detection than the current presented RT-qPCR systems.

7.5 Discussion

Current methods of PC biomarker detection are limited by many facets, but mainly it comes down to their sensitivity and selectivity. There are few PC specific biomarkers, with PSA being the basis of many early, and since iterated upon, biomarker detection platforms to assist in the diagnosis of PC, although comes with many associated drawbacks as has been found to not be truly disease specific. Instead looking into AR-V7, an AR-V that sees increased expression as PC progresses to the more aggressive treatment resistant phenotype could better assist in patient stratification and treatment decisions. Other biomarkers of interest could be mutations in key proteins that are common in PC and potentially exploitable through drug treatments, such as *SPOP* mutations in the work presented previously. Currently available techniques for detection of these oncogenic SNVs boast promising results although experimental data suggests their limits of detection may be misleading. This alongside the limited adaptability of these assays show there is an unmet need for highly adaptable and accurate detection of SNVSs from patient samples. In response to mRNA-based biomarker detection, such as AR-V7 transcript levels and SPOP SNVs, DNA-based nanotechnology techniques offer a new avenue for biomarker detection platform development that may overcome the drawbacks with the current standard of techniques.

To assess if DNA-based nanotechnology can be utilised for biomarker detection the work enclosed within this chapter aimed to validate a biomarker detection platform using AR-V7 and *SPOP* mutations as two forms of mRNA detection. This was carried out through designing simple DNA duplex probes targeting AR-V7 and assessing their abilities in the detection of endogenous transcript levels within PC cell line whole RNA lysates. As well as this, duplex and hairpin probe designs were generated that aimed to distinguish between SPOP wild type and mutant sequences in a synthetic system and test the sensitivity and selectivity in relation to that based by currently available qPCR-based SNV detection systems.

With regards to the general detection of mRNA, the early detection using synthetic inputs provided a great deal of information into how the DNA-based nanotechnology probes would function and specifically provide insights into the length of input necessary to trigger the probe and output positive signal. With the minimum complementarity necessary to displace the blocker strands in this system being 6-10 nucleotides, this means that most endogenous transcripts would be able to facilitate the strand displacement reaction and result in a positive signal regardless of potential fragmentation of endogenous inputs. Also of note from these

early experiments was the speed of which the reactions took place. In all of the concentrations and probe/target combinations tested, strand displacement reactions looked to have reached their endpoints by the time the fluorescence detection has begun. Ultimately, this displayed rapid detection of target strands and positive reads able to be output from the data within minutes of setting up the assays.

The experiments conducted looking into AR-V7 mRNA, specifically the unique E3-CE3 boundary, showed promising early proof of concept results. Across the PC cell lines tested, significant changes in probe-based detection were noted. From the early insights from this detection system, the E3-30P contained enough nucleotides sharing complementarity with E3 of the AR, which in theory should detect total AR mRNA from the PC cells. Similar detection can be expected from the CE3-30P form PC cell lines, with it being specific for AR-V7 as it is the only AR-V containing this region but does not share enough complementarity with the E3 region to trigger a strand displacement reaction. Data from the cell line detection systems looks to confirm this theory by comparisons with the HEK293T (AR negative) and VCaP (AR amplified) cell lines. Here, the CE3-30P showed no significant changes in detection of AR-V7, whereas the E3-30P displayed a significant increase in detection, indicating that the probes were able to detect the amplified AR-FL mRNA via E3-30P, whereas the CE3-30P specific for AR-V7 saw no reactions. Unfortunately, the same results were not mirrored when comparing positive detection in the HEK293T and CWR22Rv1-AR-EK (AR-V only) cell lines. Neither probe was able to produce significant positive detection indicating as expected. This is likely due to the probes LOD not being able to detect AR-V7 expression within the CWR22Rv1-AR-EK line, with VCaP *AR* amplification driving AR expression over the LOD and resulting in a positive signal.

Another main positive drawn from this experimental package was the ability for probe detection in more complex media. Seeing positive detection in serum with the synthetic system was greatly encouraging as to this point all experiments had been conducted in optimum buffer conditions to facilitate the dynamic strand displacement reactions. This informs potential downstream applications of the final assay as time consuming sample preparation steps may be unnecessary, with probes potentially being able to be added directly to liquid biopsy samples. Unfortunately, the change in background media did confer a loss in LOD of the probe systems, ultimately causing a 25-fold increase in LOD. Considering the initial LODs are already much higher than the transcript abundance described from liquid biopsy

samples, seeing increases in LOD due to more convoluted background media is a problem. This highlights a major next step in assay development if this platform was to be taken further as a detection technique. Promisingly, this simple duplex probe was initially generated to validate strand detection, and to ultimately act as the target strand recognition component or detector of a larger DNA-based nanotechnology system. Due to the high adaptability and programmability of DNA-based TMSD reactions, the blocker strand output from interaction with the specific target strand could be used to trigger a downstream reaction cascade aimed to amplify the signal seen. If this was to be developed, the presented work gives a great deal of encouragement in the selectivity of the current detector.

Unfortunately, use of AR-V7 as a biomarker for the detection of PC comes with complications out with the that of the biology and technology of detection platforms. With the ADNA test and other commercial detection methods developed for AR-V7 detection comes associated issues of patents and the ability to take development of new AR-V7 detection platforms further. Due to the potential of infringing on patents currently in place for AR-V7 detection, this line of work was not taken any further. However, if taken further potential work would have included the aforementioned iterations on the probe designs and creating a larger more complex DNA-based nanotechnology system and subsequent head-to-head testing of this system in more representative samples against the current commercially available assays.

Ultimately the use of this DNA-based nanotechnology system would likely not see routine use in gene expression studies in place of routine RT-qPCR based approaches. However, due to the speed of results and cost effectiveness, there may be other uses of this technology in respect to detection of mRNA targets. Potential uses of this technology may be better suited for providing quick snapshots of certain mRNA levels within patient samples or monitoring response to treatments through changes in expression levels pre- and post-drug treatment.

The other main work package enclosed within this chapter is the detection of SNVs using DNA-based nanotechnology probes, specifically looking into detection of SPOP mutations. Here, experiments looked to use mismatches generated upon probes interacting with target strands to slow reaction speeds and exploit differences in mismatch number by designing probes specific for mutant transcripts. This would ideally increase the discrimination factors of the probes and allow them to selectively distinguish the presence of SNV containing mRNA

amongst wild type transcripts. As before, once probes had been designed, extensive work into validating the newly generated synthetic system was conducted which generated lots of data into how these probe systems may act and how potential downstream assays would be designed.

One of the key insights from this work was the ability to overcome the complex secondary structures that are present in mRNA transcripts, if un-fragmented, through the use of blocker strands. Through the addition of blockers to the synthetic systems, opening of the specific SPOP-F133V target region was possible, as indicated by secondary structure simulations. This was validated through assessing reaction speeds in the 'closed' IL inputs compared to the 'opened' IL+Bl inputs as without the blockers present probe access to the target region was denied with ultimately led to slower reactions speeds. This gives great confidence in the translatability of this probe design system to other SNVs relevant in tumorigenesis or disease progression as mRNA secondary structures can easily be overcome.

Also tested here was a new hairpin probe design where the quencher is attached to the fluorophore in on continuous strand with internal complementarity, so a hairpin is formed, and the out signal is sequestered in the absence of the target strand. Initial experiments with this probe design showed an output in signal produced proportional to the amount of input added which fluctuated at its maximum across the total time-course, unlike what was observed in the duplex probes where the signal output was limited to the concentration of probe present. An explanation for the observations in the hairpin probe is that instead of a complete TMSD reaction taking palace, as in the duplex probes, the hairpin probes are undergoing toehold exchange reactions, where an equilibrium state is formed between the final output and intermediate steps, resulting in continuous displacement and re-binding of target input strands and resultant fluctuations in signal seen. Only at the 100 nM input overloading the system was maximal output seen across for the full time-course. With this in mind, later experiments were run with the hairpin probes to assess discrimination factors within the synthetic system as well as %MT discrimination, however results were not expected to be less positive than that of the duplex probes tested.

Key data here was found into the investigations into the discrimination factors of these probe systems, as it looked to confirm the probe-access hypothesis generated from *in silico* designs of the synthetic system. However, the mutation of interest causing opening of conformation originally closed in the wild-type transcript may not be present in other mutations of interest. Data presented looking at the Is-MT and Is-WT inputs would suggest that this may not be an issue across alternate SNV targets as with these inputs there is no secondary structures present and will have no impact on probe accessibility. This highlights the effects of doping the probes with a single mismatch when binding wild-type transcripts, which could be used to increase discrimination factors by increasing the mismatch doping further. Again, this gives great confidence in the adaptability of this detection system as through probe mismatch doping it should be possible to drive selective detection of SNVs from any relevant target.

The key finding from the SNV detection project using this synthetic system was the LODs and how they compare to that described from other detection platforms, either commercially available or from published articles. These boasted selectivity of their systems in an expression of MT:WT transcripts ratios as detectable percentages, with 0.95% MT being the most sensitive system seen, although there are caveats to this, namely that these are from spiking-in the MT transcripts prior to RT cDNA preparation, meaning that the actual amount within the detected sample is much higher than that originally spiked-in (Yang et al., 2018). In an attempt to recapitulate these read-outs, using the synthetic IL+Bl inputs a range of MT:WT ratios were made up, with significant positive detection above the 'background' seen as low as 0.5% MT giving early indications that this system was operating to the standard of, if not better than that previously described in the literature. Although initially promising results were generated here, future investigation into this sensitivity in the background of whole cell RNA lysates to simulate complex sample inputs would need to be done and assess if this has any impact on the current sensitivity.

One of the main variables that was necessary to determine from these sensitivity determination experiments was to decide what would be used as a positive read-out for the assay. Although much slower, strand displacement reactions were taking place with the WT input, meaning that direct RFU readouts may not be directly representative of MT levels as

there would still be WT contributions to total RFU. As such, it was thought that due to slower reactions with WT inputs, reactions would reach completion at a faster rate based on the MT content of each sample, but would all eventually plateau at the same total RFU. In reality, across all of the MT:WT ratios tested, no plateau was seen within the 960-minute detection window, so this was ruled as being used to detect positive read-outs. Instead, significant increases above the 'background' 0:100 MT:WT ratio was used with final validation showing this was possible if data was taken across the entire 30-minute initial detection window.

Although many positives can be taken from this proof-of-concept work, there are many further experiments necessary if this method of SNV detection was to be taken further. Alongside that previously mentioned of spiking in the synthetic system to whole cell RNA lysates, use of either $SPOP^{WT}$ or $SPOP^{MT}$ overexpression systems previously validated (as in Section 4.4.4) could be used to assess detection of SPOP point mutations within samples generated form PC cell lines. Further, testing the DNA-nanotechnology-based system against commercially available systems could provide greater confidence in the assay as a whole and encourage further development of the assay into a commercial option. Currently to do this, a different target would have to be chosen and new probes designed as there is no commercially-available option for SPOP mutation detection, highlighting the previously described lack of adaptability of these commercial options. Although probes were designed with complete complementarity to the SPOP WT transcripts, resulting in a mismatch upon MT detection, these were not used within the project as results from the MT-probes were promising. However, as described in Chapter 8, DNA-based nanotechnology detection systems are capable of multiplexed detection using probes harbouring alternate fluorophores. By annealing alternate fluorophores to a WT targeting probe and factoring this into detection alongside the MT targeting probes, selectivity and specificity of this probe system could be increased through WT-probes acting as a 'sink' capable of reducing the off-target reactions of the MT-probe with WT transcripts, as described in Wang and Zhang, 2015. Here, using more complex 'X-probe' designs with conjugated fluorophores targeting the MT inputs alongside unconjugated duplex probes targeting the WT inputs, a 1%MT sensitivity was achieved. Inputs generated here were from extracted PCR amplicons from cells expressing either the WT or MT forms of SMAD7, essentially making a clean %MT:%WT system to assess the technology similar to the synthetic input system used here (Wang and Zhang, 2015). Utilising a fluorophore conjugated WT 'sink'

could provide alternate read-outs of the %WT transcript within a given sample which could be cross-referenced with %MT input giving more strength to the results of the assay.

With both projects initially outlined as proof-of-concept studies, work enclosed within this chapter has not only satisfied the initial aims but has provided critical insights into how these methods of mRNA detection could be utilised and perform if taken further. The main positive that underlies all data presented is that these duplex probes, whilst looking to function alone, can be utilised as the initial detector component of a much larger DNA-based nanomachine with relative ease, meaning that if iterated upon could definitely outperform current probe designs but also current industry and clinical standard detection methods. These initial experiments have also given great insight into both basic applications of DNA-based nanotechnology in biomarker detection but also have given insights into the development of complete assays able to produce sensitive, selective and reliable routine tools for either target mRNA or SNV detection within a range of sample inputs. With one of the main benefits of DNA-based nanotechnology being its adaptability and programmability, although data presented here focuses on PC related mRNA transcripts, the eclosed data confirms that there is scope to utilise this technology into wider pan-cancer detection platforms.

Chapter 8 – Multiplexed miRNA detection and quantification using DNA-based nanotechnology.

8.1 Introduction

Since its introduction PSA testing saw massive benefits to detection of PC within patients and remains the most prominent biomarker for PC detection and disease monitoring (Etzioni et al., 2008). However, a key drawback with PSA testing is that it is not cancer specific with levels fluctuating due to other common diseases of the prostate, such as BPH, which results in increased false positive rates (Thompson et al., 2004). As a result, PSA testing has been incorporated and combined with a range of other biomarkers to increase its effectiveness and accuracy in PC diagnosis (Section 1.10.2). These current clinical tests look at levels of key PC-associated proteins and mRNA although there is currently a growing body of evidence focusing on the potential of miRNA as new candidate biomarkers for PC (see below).

Following RNA polymerase II- or III-mediated transcription in the nucleus, long primary miRNA transcripts (pri-miRNA) are produced which are subsequently processed by the microprocessor complex into 70-100 nucleotide long precursor RNAs (pre-miRNA). Following this, the pre-miRNA is transported to the cytoplasm for processing by Dicer and TRBP which produces 19-25 nucleotide miRNA duplexes. Of note here, the miRNA duplexes can either originate from the 5’ or 3’ arm of the pre-miRNA hairpins, denoted as -5p or -3p miRNA, respectively, with one of the arms utilised to facilitate gene regulation and the other often degraded (Kozomara and Griffiths-Jones, 2014, Mitra et al., 2015, Fabris et al., 2016) Mature miRNA duplexes then unwind into mature miRNA and direct assembly of the RNA-induced silencing complex (RISC) and ultimately regulate mRNA translation. This translation regulation commonly takes place through interaction with 3’UTR regions of mRNA to reduce stability and ultimately lead to degradation or translational repression. Further, miRNAs have also been found to induce gene expression through indirect or direct means and as such have been found to play key roles in many cellular processes such as apoptosis and proliferation (Fabris et al., 2016).

Many of the properties of miRNAs make them ideal candidates as molecular biomarkers. Due to the processing steps in miRNA biogenesis, miRNAs are relatively heterogenous and are readily found in common forms of liquid biopsies, such as plasma and serum, as well as urine and semen (Cortez et al., 2011, Mall et al., 2013). Furthermore, they can be easily extracted from a range of biological samples that may be taken, including liquid biopsy, and are both molecularly stable in various storage conditions. Finally, due to playing key roles in 'normal' cell homeostasis, miRNAs show clear changes in expression signatures in specific diseases states, including PC (Volinia et al., 2006, Ambs et al., 2008).

Methods of detecting and quantifying miRNA within a given sample includes bespoke microarrays which although useful for generating high-throughput data offer low selectivity and sensitivity. Additionally, miRNA-sequencing, which utilises NGS technology targeting small RNA which is typically missed using conventional RNAseq, offers highly reliable data although comes with increased financial cost. For routine miRNA quantification, RT-qPCR is undertaken using either poly(A) tailing or traditional stem-loop methodologies for amplification of the short RNA, although this is limited largely to relative quantification of target miRNAs using a reference 'housekeeping' miRNA or spike-in control. However, absolute quantification of miRNA within a sample can be achieved using digital droplet PCR (ddPCR), which boasts higher sensitivity than traditional qPCR and removes the need to include housekeeping genes and their associated issues (Saliminejad et al., 2019). Using these techniques, a range of studies have endeavoured to capture a full 'miRNAome' of PC using bespoke arrays, with the make-up of these array platforms rapidly evolving across the years, the ability to make direct comparisons between publications is diminished and, in some cases, produces contrasting results between studies. Ultimately results from these wider studies are also unclear due to inconsistencies with miRNA nomenclature as a consequence of generalisation and lacking -3p/-5p denotations. More recent studies have utilised a range of techniques to investigate miRNA changes in PC from liquid biopsy samples and have shown a range of key PC-associated miRNAs with altered expression, summarised in Abramovic et al 2020 (Abramovic et al., 2020) and described below.

miR-221 is encoded alongside another cancer associated miRNA, miR-222, on the X chromosome short arm (Song et al., 2017). miR-221 is commonly described as a cancer associated miRNA which has been described to be over-expressed and downregulated in a variety of cancers, including PC (Ambs et al., 2008, Sun et al., 2012, Goto et al., 2015, Song et al., 2015, Song et al., 2017). With these contrasting reports, there is also evidence suggesting large downregulation of miR-221 in PC blood and tissue samples, making it a potential candidate PC biomarker (Leidinger et al., 2016, Spahn et al., 2010). Although literature is available for both miR-221-5p and -3p, miR-221-3p has been the focus of a majority of publications (Kiener et al., 2019).

As seen with other PC-associated miRNA, published literature show both tumour suppressive and oncogenic roles of miR-221-3p. Within PC, miR-221-3p has been shown to be significantly downregulated in malignant tissue compared to benign tissue (Zhang et al., 2018b). Furthermore, miR-221-3p has been shown to reduce proliferation through SOCS3 suppression, as well as inhibiting cell migration and invasion through Ecm29 modulation (Goto et al., 2015, Zhang et al., 2018b). Contrasting this, miR-221-3p has been found to target a range of PC tumour suppressors, such as PUMA and p27/Kip1, ultimately inducing proliferative and oncogenic effects in both *in vivo* xenografts and PC cells (Galardi et al., 2007, Mercatelli et al., 2008, Zhang et al., 2010, Kneitz et al., 2014). Within PC cell lines, contrasting evidence indicates miR-221-3p acts as both a tumour suppressor and oncogene to control progression to CRPC (Krebs et al., 2020). Studies of miR-221-5p have demonstrated tumour suppressive roles of the miRNA, and as such increased miR-221-5p expression inversely correlates with disease progression (Kiener et al., 2019). Moreover, miR-221-5p directly targets the PC tumour suppressor SOCS1, supressing protein expression and thus resulting in promotion of cell proliferation and migration through loss of regulation of the MAPK/ERK signalling pathway (Shao et al., 2018).

miR-16-5p functions to regulate cell survival, cell cycle and has the potential to induce apoptosis in PC cells and as such is generally downregulated within PC. One of its direct targets AKT3 is associated with PC carcinogenesis and thus a potential method of miR-16-5p eliciting its cellular functions is through modulating activity within the PI3K/AKT/mTOR pathway (Wang

et al., 2020). Other cellular responses described are an upregulation of miR-16-5p within PC cell lines in response to IR exposure, via its modulation of the Cyclin D1/E1-pRb-E2F1 pathway, leading to enhanced radiosensitivity (Wang et al., 2019).

miR-145 functions as a tumour suppressor and has been shown to play roles in gene silencing in many cancer subtypes, including PC, and is transcriptionally controlled by p53 (Seven et al., 2014, Karatas et al., 2016, Zeng et al., 2021, Tohidast et al., 2023). Specifically, miR-145 has been described to regulate key cancer-associated processes, such as cancer stem cell proliferation, angiogenesis; whilst also implicated in growth, metastasis, invasion and chemotherapeutic resistance (Ye et al., 2019, Zeng et al., 2021). Direct targets of miR-145 that highlight its PC tumour suppressor functions include AR and SOX2 (Larne et al., 2015, Ozen et al., 2015). Of interest, SPOP has been found to be post-translationally regulated through miR-145, with the radiosensitivity conferred by SPOP deregulation or mutation able to be recapitulated in *SPOP* wild-type PC cell lines upon addition of miR-145 (Huang et al., 2014, El Bezawy et al., 2020). Finally, there is lack in denotation of whether miR-145-5p or -3p is responsible for the described oncogenic or tumour suppressive effects within the literature. With relation to PC, if noted at all, miR-145-5p is most commonly described.

8.2 Aims

Due to the drawbacks of current protein- and mRNA-based detection platforms used in PC, not only are new, more predictive biomarkers required, but so too are methods for improved detection. The inherent biology properties of miRNA, including robust stability, and high relative abundance make these molecules an attractive candidate for molecular biomarkers. As such, much work has been undertaken over the past decade to understand how miRNA-mediated gene regulation is aberrantly altered within the cancer setting, utilising techniques such as NGS, RT-qPCR and ddPCR. However, these detection methods come with associated drawbacks. Therefore, new detection platforms, which hold the potential to overcome these limitations are of great interest. As such, utilising DNA-based nanotechnology, a new method of direct miRNA detection and quantification, could be highly effective and beneficial if miRNAs are taken further as clinical biomarkers.

With the literature unclear on key candidate miRNAs to focus a specific detection platform around, a more targeted approach on target selection was used; focusing on PC-associated miRNAs that show clear expression changes between PC cell lines. As such, miR-221-3p and miR-16-5p were selected which show clear expression changes between PC3 and LNCaP cell lines (Santos et al., 2014, Dart et al., 2019). Further miR-145 was selected as it was shown as it has been shown to target SPOP and regulates its activity (El Bezawy et al., 2020), whilst also showing marginal expression changes between the PC3 and LNCaP cell lines (Santos et al., 2014).

Using these targets, a DNA-based nanotechnology approach was applied to develop a new detection platform for miRNA quantification. The aims of this chapter were therefore to;

1. Develop and validate a new probe design within a synthetic system, that holds the capability of multiplexed miRNA detection within PC cell line lysates.
2. Use this new method and compare against available NGS data and commercially available miRNA quantification tools using RT-qPCR.

8.3 Methods

8.3.1 Sequences used for miRNA-based detection.

Table 8.1 - *Oligonucleotide sequences and associated 5' and 3' modifications for 2xSEP probes*

Oligo	**Sequence**	**5' Modifier**	**3' Modifier**
P221-3p F	*GGGTAGGGATTCCTGGGAAAACTGGACTGCG*	FAM	FAM
P221-3p Q1	*GAATCCCTACCC*		Dabcyl
P221-3p Q2	*CGCAGTCCAGTT*	Dabcyl	
I221-3p	*GTCCAGTTTTCCCAGGAATCCCT*		
P145 F	*GCGTAAATCTACATTGTATGCCAGGTTGCG*	Cy3	Cy3
P145 Q1	*GTAGATTTACGC*		BHQ2
P145 Q2	*CGCAACCTGG*	BHQ2	
I145	*ACCTGGCATACAATGTAGATTT*		
P221-5p F	-	FAM	FAM
P221-5p Q1	-		Dabcyl
P221-5p Q2	-	Dabcyl	
I221-5p	-		
P16-5p F	-	Cy3	Cy3
P16-5p Q1	-		BHQ2
P16-5p Q2	-	BHQ2	
I16-5p	-		
Carrier	*TTTTTTTTTTTTTTTTTTT*		

Assay set-up.

Stoichiometric Entropy Probes were annealed from 90°C to room temperature for over 1 hour and stored in fridge for a minimum of 2 h. Experiments are prepared in 0.5 mL low-binding RNase free tubes and were supplemented with 500 nM Carrier strand. Input is added first followed by TNaK 1000 buffer (20 mM Tris-HCl, 1000 mM NaCl), followed by addition of the probe. Solutions are analysed immediately following plating. Experiments ran with RNA extracts replaces TNaK 1000 buffer with 5x concentrate TNaK buffer to normalise the ionic content of extracts to that of standard curve present on the plate.

8.3.2 Probe Fluorescence detection.

The fluorescent signal produced by probes upon binding of the target was detected on 96-well plates (Greiner) using a CLARIOstar Plus plate reader (BMG LABTECH). The experimental settings for each run were as follows:

Table 8.2 - *CLARIOstar Plus plate reader set-up for multiplexed 2xSEP probe runs*

Plate	Greiner 96-well, black, F-bottom, full-area, polystyrene, non-binding (655900)
Excitation/Emission	**FAM**: 483 nm (14 nm bandwidth)/ 530 nm (30 nm bandwidth) **Cy3:** 530 nm (bandwidth 20 nm)/ 580 nm (bandwidth 30 nm)
Shaking	10 seconds, double orbital, 500 rpm, before each cycle
Optic	Top
Temperature	25 °C
Well Volume	100 µL
Settle Time	1 sec
Cycles/Cycle Time	961/90
Measurement Time	8 hours
Flashes	20

Z height	8 mm
Gain	Fixed Gain (2000)

8.3.3 RT-qPCR analysis of miRNA targets

RT-qPCR analysis of PC3 and LNCaP whole cell RNA extracts was performed using the miRCURY LNA SYBR® Green PCR Kit (QIAGEN) for the selected targets as per manufacturers protocol and plates ran on a QuantStudio™ 6 and 7 Flex Real-Time PCR System plate reader (Thermofisher).

8.4 Results

8.4.1 2xSEP Probe Design for PC-associated miRNA

With aims of testing DNA nanotechnology-based probes capabilities for multiplexed biomarker detection, probes were designed for two targets each with different conjugated fluorophores. Alongside this, with promising results shown from duplex probes (see Chapter 7) the basic design was iterated upon further to the stochiometric entropy probe design (SEP), the main feature of which being an internal toehold which allows for conjugation of fluorophores and quenchers to the 5' and 3' end of the complementary strand to the target [Figure 8.1]. This ultimately allows for the potential of double the fluorescent output upon target strand detection than that seen from the basic duplex or hairpin designs. Using the growing evidence of miRNAs as new biomarker targets, the PC associated miRNAs miR-221-3p, miR-221-5p, miR-145-5p and miR-16-5p were chosen and probes designed for their detection.

8.4.2 2xSEP probes for miR-145-5p and miR-221-3p are selective for their specific input

Firstly, to assess if the inputs were able to trigger strand displacement reactions within the system, 5 nM of probe was run alongside 0, 2 and 4 nM of the input specific to each probe, as well as the effects of addition of 0, 2 and 4 nM of non-specific inputs [Figure 8.2]. Here, both the P145-5p and P221-3p were used, which were annealed to Cy3 and FAM fluorophores, respectively. As such, fluorescent output for each probe was measured across 65 minutes in those respective channels. As expected, the P145-5p probe displayed clear evidence of strand

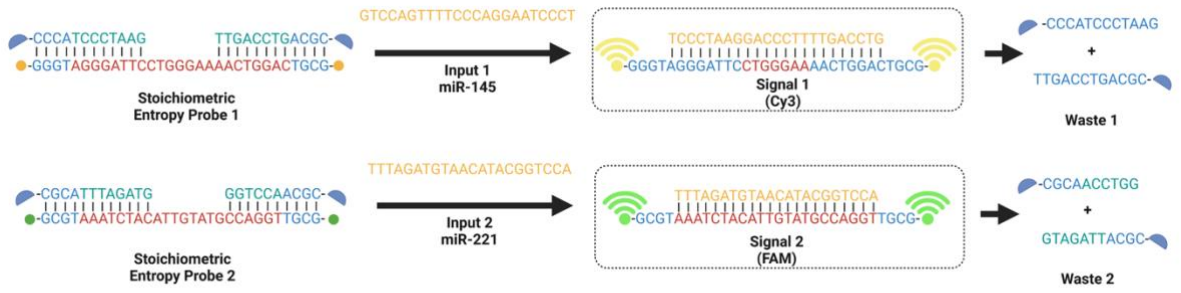


Figure 8.1 - Schematic representation of multiplexed detection using the 2xSEP probe system targeting miR-145 and miR-221. *TMSD reaction between two example probes targeting miR-145 and miR-221 with resultant Cy3 or FAM fluorescent signal produced respectively following displacement of quenchers. Fluorophores represented by yellow (Cy3) and green (FAM) circles and quenchers by blue semi-circles.*

displacement reactions taking place with a clear kinetic curve reaching a plateau by the end of the timepoints tested [Figure 8.2A]. Here, in response to 0 nM I145-5p input (probe only background) 3382 endpoint RFU was produced. With 2 nM of I145-5p the P145-5p returned 36379 RFU at the endpoint, with 69401RFU detected in response to 4 nM of the same input. Next, the P145-5p was tested with the same concentration range of I221-3p, looking again for a positive RFU output in the Cy3 channel [Figure 8.2B]. Across all input concentrations tested, no RFU increases were seen above the background of 3362 RFU, with 3552 and 3381 RFU returned as a result of the addition of 4 and 2 nM I145-5p. A similar result was observed across the concentration range upon the addition of the I145-5p to the P221-3p probe in the FAM channel, as no RFU output was seen above the baseline of 19088 RFU [Figure 8.2C]. Of note here is the increase of 15726 RFU in background signal produced by the P221-3p probe compared to background of the P145-5p probe. Upon addition of the I221-3p to the P221-3p probe, evidence of the strand displacement reactions can be seen again [Figure 8.2D]. Although the reaction did not seem to reach a definitive plateau in the time-course tested, a proportional increase in RFU output was observed with 4 nM input displaying 93878 RFU and 2 nM 56642 RFU. Together, this data demonstrates each probe is able to be triggered by their complementary targets with no off-target reactions caused by the non-specific inputs.

8.4.3 2xSEP probes do not bleed-through into opposite channels upon positive detection.

To further evaluate if the new probes would be suitable for multiplexed detections, experiments were set-up to test if a positive signal in response to one probe activation in its respective channel caused any bleed-through of signal into the other detection channel. To do this, first probe concentration was decreased to 2 nM and a range of input concentrations from 0-400 pM were tested with their specific probe and the RFU output from both the FAM and Cy3 channels was measured across 960 minutes [Figure 8.3]. With the P145-5p probe, across the range of I145-5p input concentrations, a RFU output was observed in the Cy3 channel indicating a positive strand displacement reaction. Critically, this elevated RFU was not detected in the FAM channel [Figure 8.3A]. When taking the RFU values output at the end of

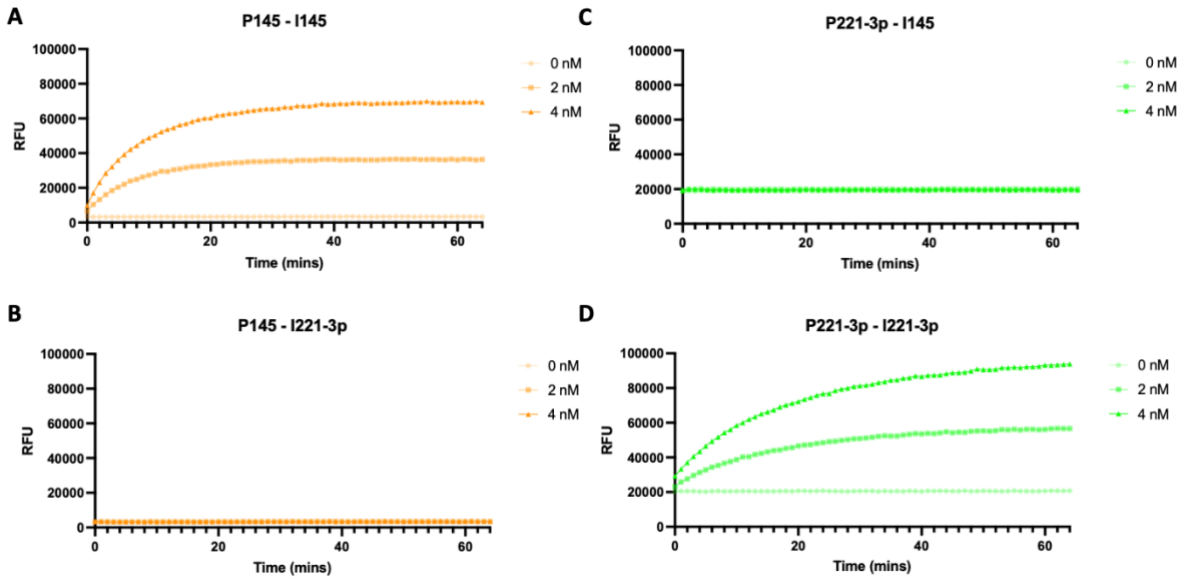


Figure 8.2 – Both P145 and P221-3p are selective for their specific input strand. A. *RFU output from Cy3 detection channel in response to 0, 2 and 4 nM I145 added to 4 nM P145 across 65-minute detection window. Data is representative of 3 individual experiments ± SEM.* ***B.*** *RFU output from Cy3 detection channel in response to 0, 2 and 4 nM I221-3p added to 4 nM P145 across 65-minute detection window. Data is representative of 3 individual experiments ± SEM.* ***C.*** *RFU output from FAM detection channel in response to 0, 2 and 4 nM I145 added to 4 nM P221-3p across 65-minute detection window. Data is representative of 3 individual experiments ± SEM.* ***D.*** *RFU output from FAM detection channel in response to 0, 2 and 4 nM P221-5p added to 4 nM P221-3p across 65-minute detection window. Data is representative of 3 individual experiments ± SEM.*

the total time-course from the Cy3 channel and plotting a standard curve, the resultant linear regression fit reports an R^2 value of 0.9918 [Figure 8.3B]. Further, the RFU output seemed to scale proportionally as before to the input concentration, with RFU ranging from 6228, 11649, 18016 and 26959 in response to 0, 100, 200 and 400 pM inputs, respectively. Finally, using the 0 pM input as the background signal and classing the limit of detection as significant RFU output over the background, these endpoint RFU can be used to calculate a 40 pM LOD for the P145-5p [Figure 8.3C].

In response to the I221-3p input concentration range, the P221-3p probe displayed evidence of the expected reactions taking place in the FAM channel. However, a minimal increase in RFU from 1230 to 1450 by the end of the time-course was observed in the Cy3 channel [Figure 8.3D]. As before, the endpoint RFU were taken, and standard curve plotted from the FAM channel outputs to return a R^2 value of 0.9913 from the regression fit [Figure 8.3E]. Again, the LOD for the P221-3p probe was quantified from these endpoint values and was found to match that seen form the P145-5p probe at 40 pM [Figure 8.3F]. Collectively this data shows that through changing the probe designs to incorporate two fluorophores per positive strand

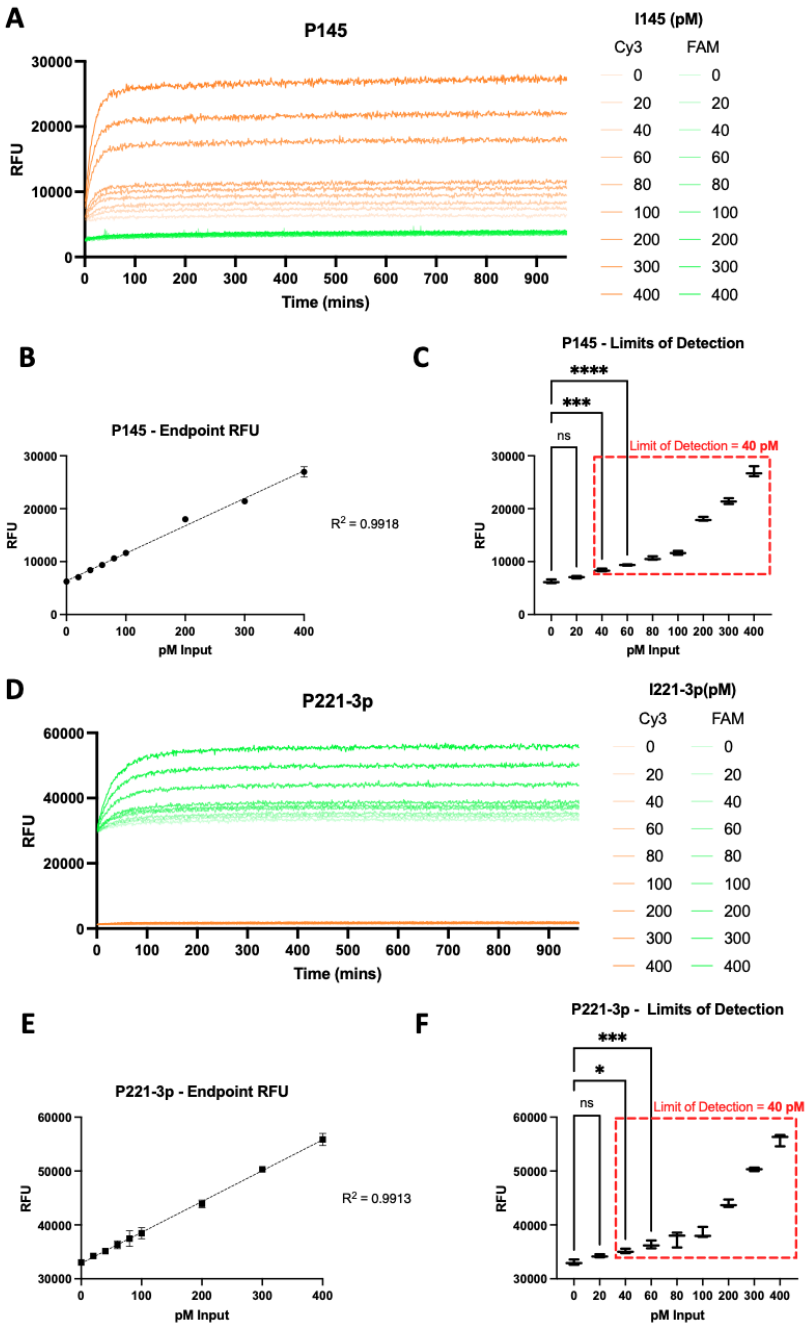


Figure 8.3 – P145 and P221-3p show no signal bleed through into alternate detection channels. A. *RFU seen in both Cy3 and FAM detection channels in response to concentration range (0– 400 pM) of I145 added to 2 nM P145 across 960-minute detection window. Data is representative of 3 individual experiments ± SEM. Endpoint RFU from Cy3 detection channel in response to concentration range (0– 400 pM) of I145 added to 2 nM P145 and interpolated standard curve with associated R2 value* ***(B)*** *as well as calculation of limits of detection* ***(C)****. Data is representative of 3 individual experiments ± SEM. Statistical significance was determined by a One-way ANOVA with Dunnett's multiple comparison test (*, **, *** and **** denote P-values <0.0332, <0.0021, <0.0002 and <0.0001 respectively, ns denotes not significant).* ***D.*** *RFU seen in both Cy3 and FAM detection channels in response to concentration range (0– 400 pM) of I221-3p added to 2 nM P221-3p across 960-minute detection window. Data is representative of 3 individual experiments ± SEM. Endpoint RFU from FAM detection channel in response to concentration range (0– 400 pM) of I221-3p added to 2 nM P221-3p and interpolated standard curve with associated R2 value* ***(E)*** *as well as calculation of limits of detection* ***(F)****. Data is representative of 3 individual experiments ± SEM. Statistical significance was determined by a One-way ANOVA with Dunnett's multiple comparison test (*, **, *** and **** denote P-values <0.0332, <0.0021, <0.0002 and <0.0001 respectively, ns denotes not significant).*

displacement reaction, LODs have increased over that seen from previous simple duplex designs. Furthermore, experiments here display that new probes show both minimal bleed-

through outside of their respective detection channel meaning that they would be suitable for multiplexed detection. Combined with a high degree of confidence in the standard curves produced; these new tools have the capacity to potentially allow for direct quantification of miRNA within a given sample.

8.4.4 2xSEP probes are capable of multiplexed detection of two inputs within the same sample.

With confidence in the new 2xSEP probe designs, further experiments were conducted with the aim of multiplexing detection of two inputs by two probes within the same solution using RFU measurements concurrently within the Cy3 and FAM channels. With evidence from previous experiments that reactions were reaching completion ~100 minutes, the RFU measurement time was reduced to 180-minutes. For these experiments, 2 nM of each probe was added to a range of inputs from 0 – 100 pM, individually at first, before adding equal amounts of both to the probes [Figure 8.4]. First, I145-5p was added to both probes and evidence of strand displacement reactions with the P145-5p probe were observed within the Cy3 channel. Also observed was an increase in RFU from 26659 to 30171 in the FAM channel across 180-minute detection although this did not appear to increase as input concentration increased. Of note, the maximal RFU produced in the Cy3 channel in response to 100 pM of I145-5p was 10761, much lower than the background RFU of 28977 detected in the FAM channel. Taking the endpoint RFU from the Cy3 channel, standard curves were able to be generated and resultant R^2 value of 0.9694 calculated, providing a marginal decrease from the R^2 value of 0.9918 seen in response to the single probe [Figure 8.4B]. Similar effects were seen upon addition of the I221-3p to both probes, with evidence of reactions taking place within the FAM channel. However, in contrast to these observations, I221-3p showed no RFU increase, remaining around 5054 RFU across the 180-minute window [Figure 8.4C]. Again, decreases in the R^2 value was seen with two probes within the system when taking the endpoint RFU from the FAM channel [Figure 8.4D]. Finally, the same pM range of input concentrations of both the I145-5p and I221-3p were added to the dual probe system [Figure 8.4E]. Crucially, both probes look to have triggered a strand displacement reaction, showing kinetic curves in each of their respective channels, with the RFU range from 6545 RFU and 28972 RFU at the start of the time-course and 8477 and 37328 RFU produced at the endpoint in the Cy3 and FAM channels respectively in response to 100 pM input. Taking the endpoint RFU from each probes respective channel standard curves for each can be interpolated,

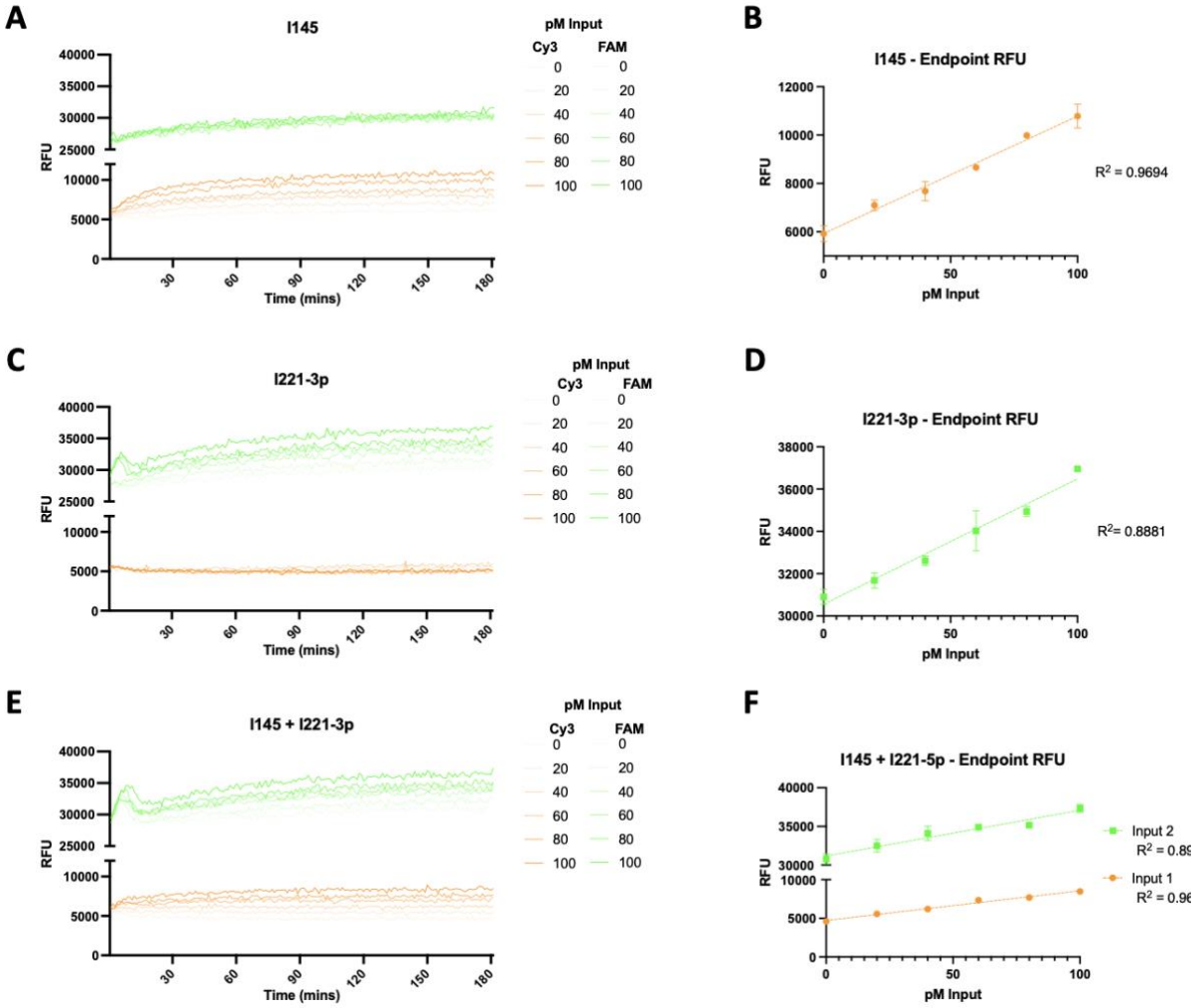
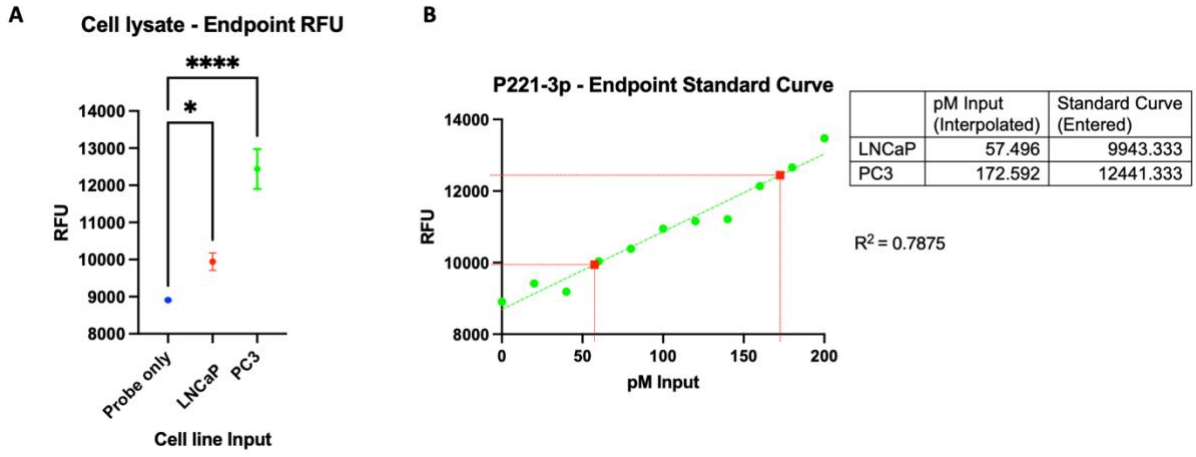


Figure 8.4 -– Both probes are able to function in the presence of each other and selectively detect their specific inputs. A. *RFU output in both Cy3 and FAM channels from concentration range of I145 (0 -100 pM) added to 1 nM P145 and 1 nM P221-3p across 180-minute detection window. Data is representative of 3 individual experiments ± SEM.* ***B.*** *Endpoint RFU from Cy3 detection channel in response to concentration range (0– 100 pM) of I145 added to 1 nM P145 and 1 nM P221-3p and interpolated standard curve with associated R2 value. Data is representative of 3 individual experiments ± SEM.* ***C.*** *RFU output in both Cy3 and FAM channels from concentration range of I221-3p (0 -100 pM) added to 1 nM P145 and 1 nM P221-3p across 180-minute detection window. Data is representative of 3 individual experiments ± SEM.* ***D.*** *Endpoint RFU from FAM detection channel in response to concentration range (0– 100 pM) of I221-3p added to 1 nM P145 and 1 nM P221-3p and interpolated standard curve with associated R2 value. Data is representative of 3 individual experiments ± SEM.* ***E.*** *RFU output in both Cy3 and FAM channels from concentration range of I145 and I221-3p (0 -100 pM) added to 1 nM P145 and 1 nM P221-3p across 180-minute detection window. Data is representative of 3 individual experiments ± SEM.* ***F.*** *Endpoint RFU from Cy3 detection channel in response to concentration range (0– 100 pM) of I145 and I221-3p added to 1 nM P145 and 1 nM P221-3p and interpolated standard curve with associated R2 value. Data is representative of 3 individual experiments ± SEM.*

returning R^2 values of 0.9962 and 0.8996 for the P145-5p and P221-3p probes, respectively, which indicates minimal change between values for individual versus multiplexed probe experimental set-ups [Figure 8.4F]. Together, this demonstrates that the probes are able to detect their own respective inputs when in solution together, with minimal bleed through into the opposite detection channel and are able to facilitate multiplexed detection which, through use of calibration curves, would be able to quantify miRNA content within a given sample.

8.4.5 Quantification of miR-221-3p within PC3 and LNCaP whole RNA cell lysates

Continuing towards the aim of multiplexed biomarker detection within cell lysates, it was important to provide evidence of the current probes working within whole cell RNA lysates. With previous literature reporting an increase in miR221 within PC3 cells compared to LNCaP cells (Santos et al., 2014, Dart et al., 2019), equal quantities of total RNA from whole cell RNA lysates from each of these lines was used. Following assay set-up, final input concentration of these lines were 0.6 µg/µl of RNA, to which 2 nM of P221-3p was added and RFU output detected over 960 minutes in the FAM channel [Figure 8.5]. When taking the RFU values at the endpoint of the experiment, the background fluorescence produced by the probe alone was observed at 8909 RFU, with significant increases to 9943 and 12441 RFU produced by the LNCaP and PC3 cell lines, respectively [Figure 8.5A]. Alongside the whole cell RNA lysates ran within these experiments, an internal standard curve was set-up with an input concentration range from 0 to 200 pM using the synthetic inputs as before [Figure 8.5B]. Although the standard curve generated returned the lowest R^2 value seen at 0.7875, the RFU output from the whole cell RNA lysates were able to be input and the pM of miR221-3p interpolated. The results of this returned 57.7 pM of miR221-3p within the LNCaP lysate and 172.6 pM miR221-3p in the PC3 lysate, showing the similar trend of greater quantity miR221-3p in PC3 cells than LNCaPs, to that previously reported in the literature.



	pM Input (Interpolated)	Standard Curve (Entered)
LNCaP	57.496	9943.333
PC3	172.592	12441.333

Figure 8.5 – 2xSEP probes are capable of detecting endogenous miRNA within PC cell lysates. A. *Endpoint RFU from Probe only background, LNCaP and PC3 whole cell RNA extracts added to 2 nM P221-3p. Data is representative of 3 individual experiments ± SEM. Statistical significance was determined by a One-way ANOVA with Dunnett's multiple comparison test (*, **, *** and **** denote P-values <0.0332, <0.0021, <0.0002 and <0.0001 respectively, ns denotes not significant).* ***B.*** *Standard curve generated from endpoint RFU (on same plate as whole cell RNA lysates) from concentration range (0 - 200 pM) I221-3p used to interpolate pM miR221-3p within whole cell RNA lysates. Data is representative of 3 individual experiments ± SEM*

8.4.6 Validation of 2xSEP probes for multiplexed detection of miR-221-5p and miR-16-5p

Satisfied that the P221-3p probe was able to detect changes within miRNA levels that had previously been reported in the literature, this was taken on for further experimentation. With there being minimal differences in miR-145-5p levels reported previously within PC cell lines a new probe was designed targeting miR16-5p whose abundance had been previously reported to change substantially within PC3 and LNCaP cell lines (Santos et al., 2014). As well as this, to investigate whether this DNA-based nanotechnology system would be able to detect between miRNA 5p/3p arms, another probe was designed targeting miR221-5p. The new P16-5p was conjugated to Cy3 fluorophores so that it could be used alongside the current miR-221-3p probe for multiplexed detection, and the P221-5p probe conjugated to FAM fluorophores. As before, experiments were set up to validate the new probes multiplexing capabilities by adding a range of input concentrations to 2 nM of both probes and detecting RFU output in both the Cy3 and FAM channels across 960 minutes [Figure 8.6]. Firstly, the two new probes were used in combination with their respective I221-5p and I16-5p inputs [Figures 8.6A&B]. In the P221-5p probe reactions a plateau was reached by the end of the time-course tested, with a maximal RFU of 47536 produced in response to 500 pM I221-5p [Figure 8.6A]. For the P16-5p probe, a high degree of variance was observed in results and is reflected in the traces produced across the 960-minute time-course [Figure 8.6B]. As expected, when taking the endpoint RFU from these probes together, clear increases in RFU can be seen with increasing input concentration [Figure 8.6C]. The resultant R^2 values calculated from the standard curves were 0.9915 and 0.9970 from the P221-5p and P16-5p probes, respectively, giving great confidence in the ability for these probes to precisely quantify RNA from a given sample. Before continuing further, the ability of the P16-5p probe to function alongside the P221-3p that had been used previously was also validated [Figure 8.6D-F]. As seen previously, the P221-3p probe returned a much higher RFU range than other probes tested and displayed a similar time to reach plateau as in previous experiments [Figure 8.6D]. Of note here, there was higher background signal in this system, with much greater increases seen in the probe only control than in previous experiments and those observed in the other probes tested. In the case of the P16-5p probe, results almost mirrored that seen from the experimental runs performed with the P221-5p probe, with similar RFU outputs and time to reach a reaction plateau [Figure 8.6E]. Again, the endpoint RFU from each probe was taken and standard curve plotted [Figure 8.6F]. The P221-

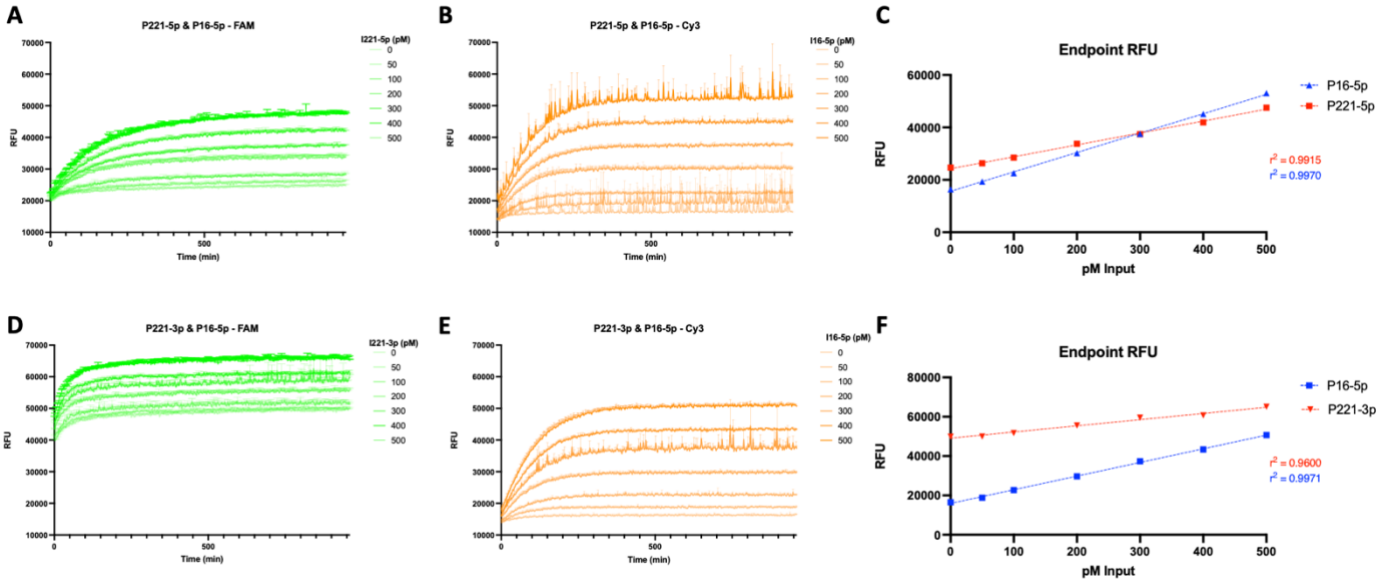


Figure 8.6 – Validation of new 2xSEP probes targeting miR221-5p and miR-16-5p. A. *RFU output in FAM detection channel from concentration range (0 – 500 pM) I221-5p added to 1 nM P221-5p and 1 nM P16-5p across 960-minute detection window. Data is representative of 3 individual experiments ± SEM.* ***B.*** *RFU output in Cy3 detection channel from concentration range (0 – 500 pM) I16-5p added to 1 nM P221-5p and 1 nM P16-5p across 960-minute detection window. Data is representative of 3 individual experiments ± SEM.* ***C.*** *Endpoint RFU from Cy3 detection channel in response to concentration range (0– 500 pM) of I221-5p added to 1 nM P221-5p and 1 nM P16-5p and FAM detection channel in response to concentration range (0– 500 pM) of I16-5p with interpolated standard curves with associated R2 values for both. Data is representative of 3 individual experiments ± SEM.* ***D.*** *RFU output in FAM detection channel from concentration range (0 – 500 pM) I221-3p added to 1 nM P221-3p and 1 nM P16-5p across 960-minute detection window. Data is representative of 3 individual experiments ± SEM.* ***E.*** *RFU output in Cy3 detection channel from concentration range (0 – 500 pM) I16-5p added to 1 nM P221-3p and 1 nM P16-5p across 960-minute detection window. Data is representative of 3 individual experiments ± SEM.* ***F.*** *Endpoint RFU from Cy3 detection channel in response to concentration range (0– 500 pM) of I221-3p added to 1 nM P221-3p and 1 nM P16-5p and FAM detection channel in response to concentration range (0– 500 pM) of I16-5p with interpolated standard curves with associated R2 values for both. Data is representative of 3 individual experiments ± SEM.*

3p probe returned greater values than that reported from previous multiplexing experiments and, importantly, the P16-5p probe produces near-identical R^2 values to that seen in the previous experiments with the P221-5p probe. This finding, alongside the data from the P221-3p probe, suggest that individual probes will function in multiplex experiments, regardless of which probe it is paired with as long as distinct conjugated fluorophores are utilised, and hence enable modular assay designs using these probes.

8.4.7 Multiplexed detection of miR-221-3p, miR-221-5p and miR-16-5p within PC3 and LNCaP whole cell RNA lysates.

Confident in the ability for both combinations of probes to enable multiplexed detection of targets within our synthetic system, they were next tested using whole cell RNA lysates from PC3 and LNCaP PC cell lines as before [Figure 8.7]. Here, a combination of either the P16-5p

and P221-5p or the P16-5p and P221-3p were added to 1 µg/µL RNA from each PC cell line and RFU captured as before across 960 minutes in each respective channel. An internal standard curve was also generated on the plate for each probe with its complementary input. The endpoint RFU produced from P221-5p and P16-5p were taken and compared to a probe only background showing significant increases in RFU in both lines tested [Figure 8.7A]. The P221-5p returned an increase from a background of 24699 RFU to 26453 RFU in response to the PC3 input and 27138 RFU in response to the LNCaP input. A similar trend was seen in the P16-5p with 16347, 24232 and 28935 seen from the probe alone, PC3 and LNCaP inputs, respectively. Taking the endpoint RFU from the P221-3p and P16-5p combination, the P16-5p probe shows similar trends as in the previous combination, with minimal differences seen in RFU output in response to each input. However, the P221-3p probe produced a higher RFU of 55138 in response to the PC3 input compared to the LNCaP input, producing 52732 RFU, with both being significant increases over the 49891 RFU probe only background.

8.4.8 miR-221-3p, miR-221-5p and miR-16-5p multiplexed quantification using 2xSEP probes partially aligns with RT-qPCR and public NGS data.

To validate the findings from the probes data, RT-qPCR experiments were conducted, and the total results were compared to publicly available NGS data [Figure 8.8]. Using the internal standard curves from the previous experimental set-up [Figure 8.7], standard curves were generated, and the pM levels of each miRNA were interpolated [Figures 8.8A&B]. Here miR221-5p levels were the lowest reported, with 48.6 pM observed in the PC3 cells an increase to 63.7 pM in the LNCaP inputs. Contrasting this, although greater amounts generally quantified, 191.2 and 115.1 pM levels of miR221-3p were reported from the PC3 and LNCaP input respectively. Within both of the experiments, the levels of miR-16-5p interpolated remained relatively constant, with 116.2 and 135.8 pM reported from the PC3 inputs and 179.8 and 179.9 pM from the LNCaP inputs. Although results to this point look to align with the trends reported within the literature, the data presented in the literature and methods

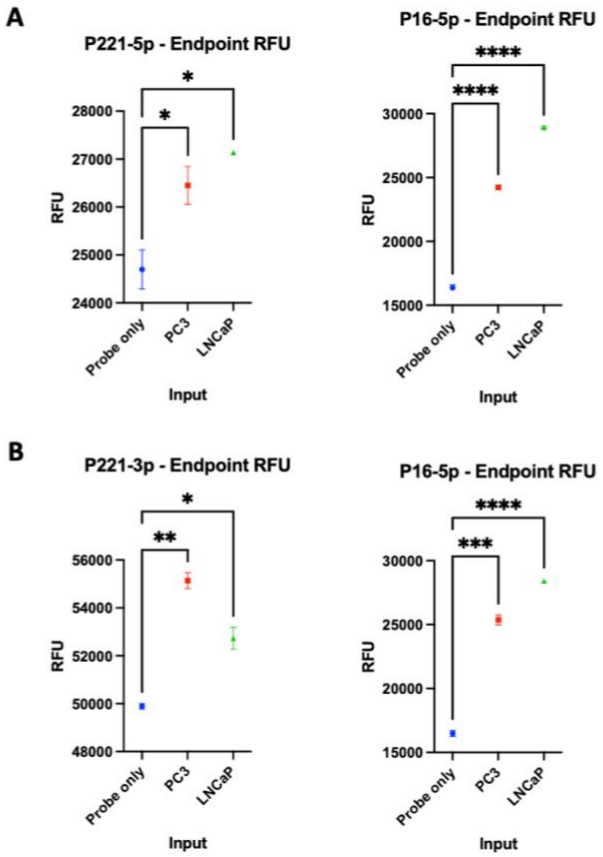


Figure 8.7 – 2xSEPs are able to detect differential expression of miRNA within PC3 and LNCaP cells. *Endpoint RFU from Probe only background, LNCaP and PC3 whole cell RNA extracts added to 1 nM P221-5p and 1nM P16-5p (**A**) and 1 nM P221-3p and 1nM P16-5p (**B**). Data is representative of 3 individual experiments ± SEM. Statistical significance was determined by a One-way ANOVA with Dunnett's multiple comparison test (*, **, *** and **** denote P-values <0.0332, <0.0021, <0.0002 and <0.0001 respectively, ns denotes not significant).*

used vary between reports, thus RT-qPCR experiments looking into the selected miRNA levels were performed using a commercially available kit in attempts to further validate the findings from the probes [Figure 8.8C]. Investigated here were miR-16-5p, miR-221-3p and miR-221-5p levels as performed for the probes, alongside a UniSp6 spike-in control and miR-103a which is used as a miRNA standard control. The spike-in controls between the PC3 and LNCaP inputs generated mean Ct values of 22.2 from both cell line inputs however differences were seen in miR-103a levels returning a mean Ct of 17.6 in the PC3 inputs and 18.7 from the LNCaP inputs. With regards to miR-16-5p Ct values were 16.1 and 16.6 for the PC3 and LNCaP inputs. miR-221-3p and miR-221-5p both show the same trends when comparing relative levels in PC3 and LNCaPs, with it being more abundant in PC3s than in LNCaPs and there was generally less of miR-221-5p within both line than that of miR-221-3p. Finally, using publicly available NGS data utilising a bespoke CombiMatrix 4x2k human microRNA microarray platform for detection, raw

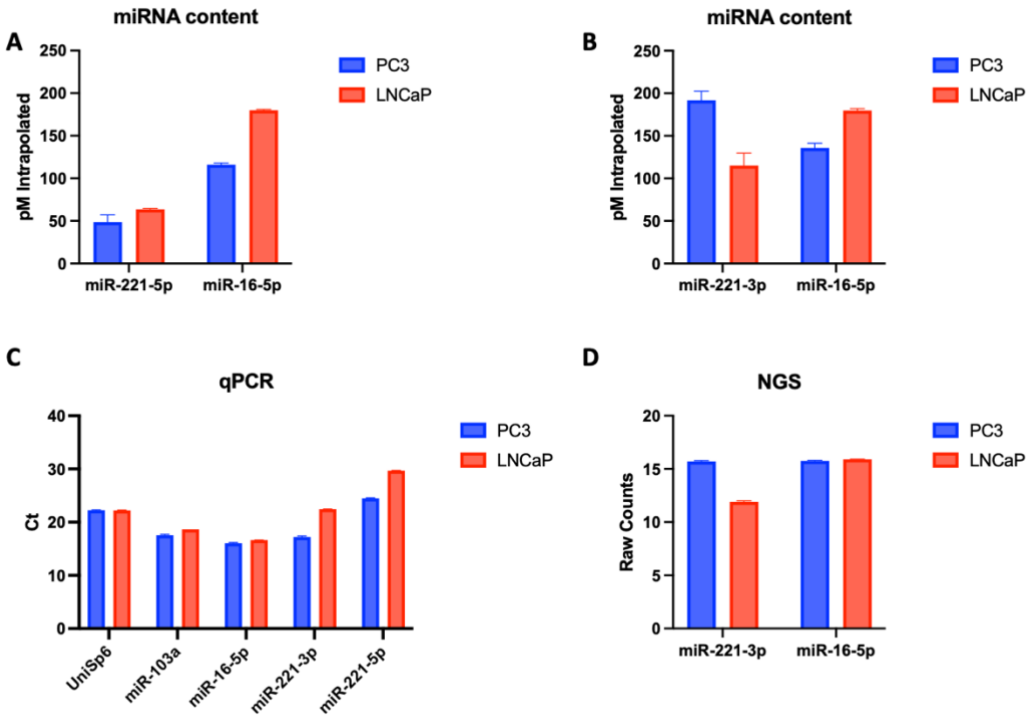


Figure 8.8 – 2xSEP probe experimental results compared to RT-qPCR and NGS miRNA analysis. A. *pM miR-221-5p and miR-16-5p content interpolated from standard curves as before (Figure 8.7) from endpoint RFU from both PC3 and LNCaP whole cell RNA inputs. Data is representative of 3 individual experiments ± SEM.* ***B.*** *pM miR-221-3p and miR-16-5p content interpolated from standard curves as before (Figure 8.7) from endpoint RFU from both PC3 and LNCaP whole cell RNA inputs. Data is representative of 3 individual experiments ± SEM.* ***C.*** *qPCR Ct values from PC3 and LNCaP whole cell RNA extracts for UniSp6 spike-in and miR-103a endogenous controls alongside miR-16-5p, miR-221-3p and miR-221-5p. Data is representative of 3 individual experiments ± SEM.* ***D.*** *Raw counts from NGS data from both PC3 and LNCaP cell lines for mir-221-3p and miR-16-5p. Data is representative of 4 individual experiments ± SEM.*

counts for miR-221-3p and miR-16-5p were compiled [Figure 8.8D](Boll et al., 2013). As seen within the probes and as seen from the RT-qPCR experiments, there was greater amounts of miR-221-3p seen in the PC3 line when compared to the LNCaP line, although not to the same extent as indicated by the RT-qPCR. miR-16-5p levels were seen to be the same between the two lines, contrasting the results from the probes but aligning with that seen from the RT-qPCR based experiments.

8.5 Discussion

As noted previously, there are many known disadvantages to utilising PSA as a biomarker for the detection of PC from patient samples. As such, there have been multiple attempts to either combine PSA testing with other biomarkers or validate new PC specific biomarkers for routine use. One field of research which has seen extensive work is that of gene silencing by miRNAs, in which altered expression in PC can drive disease progression. Alongside their known effects on gene silencing, miRNA hold many attractive qualities as biomarkers namely their sequence heterogeneity and bioavailability in liquid biopsy samples. As such, methods of miRNA detection and quantification have been developed, however these each have their own associated drawbacks and also fall victim to the issue of miRNA generalisation and lacking specific nomenclature which can confuse findings when looking across publications. Regardless, specific miRNAs have been found to see altered expression in the PC3 and LNCaP PC cell lines (as well as PC samples) such as miR-221-3p, miR-221-5p, miR-16-5p and miR-145 which have been shown to display tumour suppressive and oncogenic roles.

Due to literature remaining unclear, the above miRNAs were selected as candidate miRNA for designing and validating detection platforms as they see clear expression differences across select PC cell lines. With DNA nanotechnology-based approaches already showing positive early results in previous work focusing on mRNA-based detection, work enclosed within this chapter looked to adapt this detection technique, iterate on the current prove designs and repurpose the technology for direct detection and quantification of multiple miRNAs within given samples through a multiplexed detection system. Alongside this, due to the availability of both NGS data for miRNA content of PC cell lines and commercially available kits capable of detecting the above miRNA, work outlined here also aimed to compare the sensitivity and selectivity of the newly developed miRNA detection platform using DNA-based nanotechnology probes against the results from NGS and RT-qPCR.

One of the first notable findings from the enclosed work was the iteration on the probe designs presented in Chapter 7, moving from the duplex to the 2XSEP probes with two conjugated fluorophores. Encouragingly, these additional fluorophores resulted in an LOD of 40 pM, a marked decrease over the 100 pM displayed from the duplex probes. Further, the strand displacement reaction in this instance results in two blocker strands being released allowing for double the amount of downstream amplification reactions if these designs were to be taken further.

The main novel application of these DNA-based nanotechnology systems discovered from this project was the ability to multiplex detection from multiple probes within the same reaction; producing data comparable to that of the probes alone with no impact on limits of detection or standard curve calibration. Importantly, there was no signal bleed through into the other detection channels meaning this could be, in theory, expanded to factor in multiple other probes assuming correct fluorophore combinations were selected. This unlocks a range of benefits to the technology that can be factored into downstream applications such as direct quantification of miRNA content of given samples through use of internal standard curves present on plates. This direct quantification was only previously available through dd-RT-qPCR. Another application possible would-be relative quantification of miRNA targets to a housekeeper gene, similar to quantification performed in RT-qPCR, through comparing relative fluorescence of target miRNA to that of probes specific to housekeeper miRNA.

The final experiments in this chapter looked to compare the DNA-based nanotechnology detection system to other techniques for miRNA detection and quantification. Of those available, the QIAGEN RT-qPCR kit was used and specific primers selected for miR-221-5p, miR-221-3p and miR-16-5p. Also available was published NGS data looking at miRNA expression across a range of PC cell lines. QIAGEN RT-qPCR experiments were run in-house and compared to experiments set-up as similar as possible using the 2xSEP probes, with both collated and compared to the raw counts extracted for the available targets from the NGS data (miR-221-3P and miR-16-5p).Unfortunately there was no data available from miR-221-5p from the NGS data, likely due to the lack of specific nomenclature available. From this data, there was only data for ‘miR-221’ and only through analysis of the sequence prescribed to this was it uncovered that this was for miR-221-3p. Of note from these comparisons, initially there looked to be less miR-221-5p expressed from samples than miR-221-3p which correlated well with the qPCR data. However, miRNA content generated from the 2xSEP probes indicated that there was less miR-221-3p in LNCaP cells compared to the PC3 cells, the opposite of which was observed via qPCR analysis. Though expression was lower, qPCR analysis would indicate a substantially higher presence of miR-221-5p in PC3 cells than LNCaP cells, which again the opposite was seen from quantification using the 2xSEP probe. One potential explanation for the differences seen through the qPCR analysis could stem from the specificity of the primer set used, as when looking into the melt curves available from each primer set there looks to be unexpected peaks, indicating the presence of primers binding to alternate targets. Looking

specifically at the temperatures these peaks appear; it would suggest that the primers are binding to both miR-221 isoforms leading to potentially unreliable results. Frustratingly, the NGS data available looked to confirm the miR-221-3p expression from the 2xSEP system as well as the miR-16-5p expression from the qPCR analysis meaning that it sheds little light onto which is the 'correct' output from either the 2xSEP system or the qPCR assay.

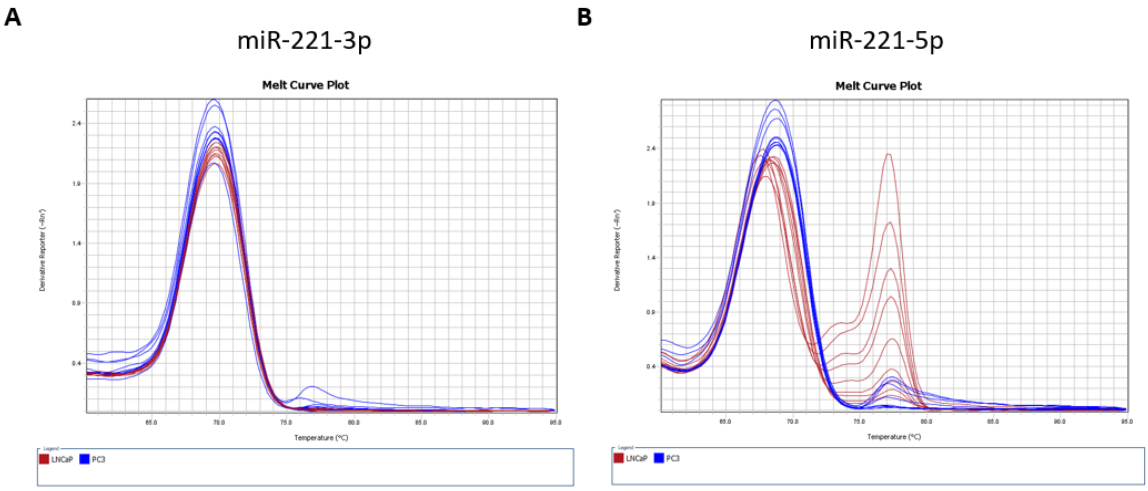


Figure 8.9 – Melt curve analysis of qPCR primer sets for miR-221-3p and miR-221-5p quantification show amplification of alternate isoforms. *Melt curve graphs for primers for both miR-221-3p (A) and mir-221-5p (B) from qPCR quantification conducted within above results. Data is representative of 3 independent repeats.*

Looking to the published data available for miRNA content from both cell lines in clear similarities between the data presented from the 2xSEP probes can be seen (Santos et al., 2014, Dart et al., 2019). Firstly, using qPCR Santos et al., described am 11.3-fold increase in miR-221 expression in PC3 cells over LNCaP cells (Santos et al., 2014). In line with these qPCR experiments, normalising against U6 and U8 RNA, described a ~30 fold change increase in PC3 cells compared to ~4 fold change in LNCaP cells (Dart et al., 2019). Although these fall victim of the issue with nomenclature, the trends in expression changes between the PC3 and LNCaP cell lines are recapitulated in the data from the 2xSEP probe system. Further, Dart et al., described an increase in miR-16 in the PC3 cells compared to LNCaP cells, aligning with the data presented from the 2xSEP probe system and contradicting that of the QIAGEN qPCR analysis (Dart et al., 2019).

Although the current work package presented within this chapter makes a compelling argument for the use of DNA-based nanotechnology in miRNA detection and quantification, more work is necessary for this detection platform to be fully realised. From the outlined work, a necessary experiment would have been to send the same samples used for qPCR and 2xSEP analysis for short-RNAseq and perform in-house NGS analysis. This would allow for definitive answers to the questions raised by the differences in miRNA expression seen from the 2xSEP and QIAGEN qPCR analysis that the available NGS data wasn't able to confirm. Another line of work would be looking into cross reactivity of the miR-221-5p and miR-221-3p probes by seeing if there was any triggering of reactions with non-specific inputs for each. Further multiplexing miR-221-5p and miR-221-3p detection within the same system to see if data aligns with current outputs, would provide great insights into the specificity of the probes and take the technology to a higher level than that of the qPCR techniques used. Of interest would be increasing the probes present within each system to 3, with alternate conjugated fluorophores, and assessing the effects on detection sensitivity and selectivity would be of great interest as this would allow for comparisons of multiple targets against a given housekeeper in the same well. If results from this look promising, then the full limits of this multiplexed detection can be tested by increasing probe number until signal bleed-through becomes too high and has negative impacts on results. Finally, as mentioned previously these initial experiments, although generating positive results alone, were focussing on validating the initial detector components of potential larger DNA-based nanomachines. With 2 strands displaced upon target interaction downstream signal amplification cascades are again possible and allow for potentially greatly increased sensitivity than that of the previous duplex probes and are more likely to be able to detect physiological levels of miRNA within liquid biopsy samples.

Another potential application of the technology is quantifying miRNA expression within live cells and has been indicated within the literature using more complex probe designs by Wu et al., 2015. (Wu et al., 2015). Here, they demonstrate a reporter system able to be transfected into live cells, with an amplified fluorescent signal output upon target mRNA transcription. Using confocal fluorescence microscopy, they were able to detect signal output within cells, as well as quantify this using flow cytometry. Although not presented within the main body of results, mainly due to this experiment being very rough and need of a great deal of

optimisation, data was able to be generated for intracellular detection of miR-221-3p within PC3 and LNCaP cells with results showing the previously described differences in expression [Figure 8.10]. This provides another potential application of the technology outside of the proposed detection assay looking to be developed within this chapter, allowing detection of miRNA within *in virto* experiments and could potentially be used to monitor expression changes in response to a given cellular treatment.

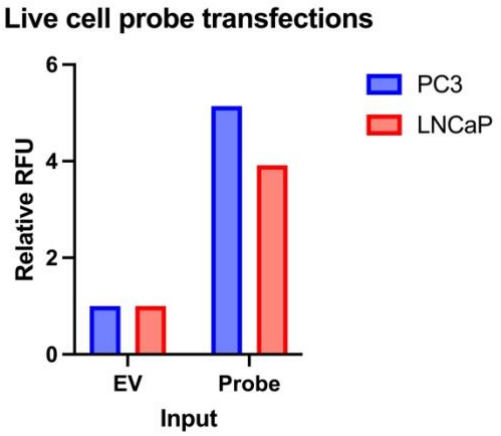


Figure 8.10 – Live cell transfection of 2xSEP probes targeting miR-221-3p show fluorescent output that aligns with differences seen from cell lysate detection. *PC3 and LNCaP cells transfected with either empty vector (EV) or miR-221-3p probes and RFU output quantified for each cell line relative to the EV controls. Data is representative of 3 independent experimental repeats.*

Aside from the comparisons of results between the options for miRNA detection there are many benefits of the probe-based detection system over qPCR that were observed from hands-on use of both. The key being the costs associated with running each system, with the total costs for the 2xSEP probes used within the project being around ~£500, and the cost for ordering the QIAGEN system for the miR-221-5p, miR-221-3p and miR-16-5p targets being ~£1,500. Aside from this, being able to adapt and order different probes with alternate conjugated fluorophores for any given probe design is a key benefit to the technology as opposed to only being able to purchase select targets from QIAGEN for the qPCR assay whilst being given no indication of the primer sequences used. The final main benefit is the time to results, as positive readouts could be observed within 60 minutes of adding the probes directly to the samples, whereas RT and qPCR steps with the QIAGEN are upwards of 120 minutes disregarding plate setup.

There are some key limitations to the current system, namely fluorophore availability and the use of miRNA as biomarkers. Currently when ordering in DNA oligonucleotides for probe design there is a limited repertoire of combinations of fluorophores that can be conjugated on both the 5' and 3' ends of the oligonucleotide strands. Without validating in-house oligonucleotide conjugation system this is currently a limiting factor on the number of probes able to be ran within a given experimental set-up. Further, there will ultimately be signal bleed through into alternate detection channels as fluorophore content increases, however the threshold for this is unknown and would need to be tested experimentally to see if this significantly affects probe sensitivity and standard curve calibration. Although they possess inherent key characteristics that make them suitable biomarkers there are some issues with using them in a diagnostic assay to detect the presence of suspected PC from patient samples. There is still conflict within the literature about the effects on certain miRNAs expression relative to disease state with no clear target seeing increased expression changes across multiple studies. Also still present in the literature is the issue of nomenclature which may, again, be adding to the confusion when collating expression changes across published studies as some papers are still generalising miRNA and not including the -5p/-3p suffixes. Together, these indicate it will take time before any specific miRNA can be validated as PC specific and the expression levels can be used to detected PC or direct patient treatment, however folding miRNA expression levels could still be used as part of a larger assay looking to combine multiple PC biomarkers to give a more accurate snapshot of a patients' disease state.

Ultimately work enclosed within this chapter provides a proof-of-concept study for a novel miRNA detection and quantification platform using DNA-based nanotechnology that is able to potentially provide results in line with RT-qPCR and NGS analysis. With further iteration on the detection system presented there is potential for a reliable, affordable, sensitive and selective assay to be developed that could be used to detect PC miRNA biomarkers form liquid biopsy samples.

Chapter 9 – Key conclusions and future work

9.1 Key conclusions

PC is the second most common cancer related death in the UK, seeing 47,500 diagnoses and 11,500 deaths annually (PCUK, 2021). Initial responses to treatment are generally well received however within 5 years disease recurrence occurs, and as this cancer progresses viable treatment options diminish. Biomarkers, such as PSA, have been utilised in diagnostic screening of PC and many genetic factors have been described to enhance PC progression and predict aggressiveness however these have displayed limited sensitivity and selectivity leading to false positives and misdiagnosis of PC. As a results identification and exploitation of genetic biomarkers and generation of new biomarker detection platforms are a key areas of research to better diagnose and direct patient PC treatment.

In attempts to identify new potential PC biomarkers, work within this project focusses on the SPOP protein, found mutated in 4-24% of PC patients. SPOP has been described to play diverse roles in cell homeostasis in a range of cellular processes, namely DNA repair. PC-associate *SPOP* mutations have been described to repress the error-free HR DDR mechanism, causing an over-reliance on the error-prone NHEJ DDR pathway and accumulation of generic aberrations. With *SPOP* mutations causing a *BRCAness*-like phenotype within cells, there are therapeutic vulnerabilities such as increased sensitivity to PARPi. Due to the overreliance on NHEJ-mediated DDR, inhibition of key NHEJ protein DNA-PKcs may cause similar therapeutic sensitivities. In order to investigate this potential synthetic lethal relationship, this project validated a range of SPOP targeting reagents, as well as generated multiple PC cell line model systems to allow for investigation into the effects of functional SPOP loss within PC cell lines. Ultimately, through extensive *in vitro* investigation, the sensitivity to DNA-PKcsi was validated and also shown to be enhanced through the introduction of exogenous DNA damage through IR exposure. Alongside this, using *in silico* analysis of a range of in-house and published datasets investigating SPOP loss of function, new pathways were identified that may reveal new drug susceptibilities conferred by SPOP mutations.

To bolster the work presented, utilising the methodologies and reagents validated within the project, creation of *SPOP* knock-out and knock-in models in alternate PC cell line backgrounds and assessment on DNA-PKcsi sensitivity is essential. With indications from *in silico* analysis

indicating the presence of the SPOP protein, regardless of mutation state, having effects on gene transcription, using the SPOP knock-out models and re-introducing SPOP wildtype or mutant protein through plasmid based-expression would allow for easier and insightful investigation into this previously undescribed role of SPOP.

Due to SPOP playing diverse roles in normal cell homeostasis, analysis into the wider transcriptomic effects of SPOP loss of function may reveal new pathways warranting further investigation. Using RNAseq datasets generated using the CWR22Rv1-AR-EK-SPOP$^{-/-}$ clones, TGCA patient data for SPOP^{MT} cases and published datasets using SPOP overexpression/knock-down, similarities in positively and negatively enriched hallmark pathways were observed. Importantly, across a majority of datasets analysed, the *TNFa signalling via NFKB* hallmark was observed as negatively enriched, contrasting with published literature (Geng et al., 2023). Further, experimental evidence of enhanced cellular EMT to confirm results from the *in silico* analysis, would be beneficial to validate findings from the RNAseq analysis.

The second focus of this project was to utilise DNA-based nanotechnology to develop new PC biomarker detection assays. The current repertoire of PC detection assays are limited in their sensitivity and selectivity whist only targeting certain panels of PC biomarkers. DNA-based nanotechnology has many benefits over traditional biomarker detection and quantification techniques currently, specifically the high degree of programmability using well defined reaction fundamentals. Within this project, early proof-of-concept experiments have generated encouraging results demonstrating the ability to detect miRNA, mRNA and point mutations within synthetic systems. Alongside this positive detection were seen in convoluted background media, such as human serum, as well as experimental results attaining the same, and in some cases out-performing, commercial or published biomarker detection platforms.

Importantly, from the work into validating AR-V7 detection using duplex probes, evidence of differences in total AR expression were observed between VCaP and control cells. Unfortunately, no significant detection of AR-V7 mRNA was observed indicating that current LODs are not sensitive enough to detect from whole cell line lysates. Alongside this, detection of SNVs within synthetic systems demonstrated sensitivity that outperformed that described by commercially available detection systems, using *SPOP* mutations as candidate biomarkers. Another key point from this work was the demonstration of detection capabilities in human serum, as all other experiments were run in optimum buffer conditions. Although many positive preliminary insights were generated from this work, drastic probe iteration would be

necessary to detect physiologically relevant levels of mRNA from patient liquid biopsy samples. Factoring in later 2xSEP probe designs as well as developing downstream signal amplification steps could boost sensitivity if this work was to be taken further. If these iterations on assay designs prove beneficial, future work could be conducted in more representative samples, such as spiked-in experiments in healthy blood to demonstrate the translatability of this technology as a diagnostic tool.

As well as mRNA detection using PC associated candidate biomarkers, investigation into differences in PC-associated miRNA detection from two PC cell lines using 2xSEP probe designs showed promising results. Importantly here, using commercially available miRNA detection RT-qPCR tools and published NGS data, results from the probe-based detection systems aligned partially with NGS data and indicated that the qPCR-based techniques are lacking in specificity. Aligning this project objective to the study of SPOP-mediated therapeutic vulnerability, further development of this biomarker detection assay has the potential to see translational benefit to assessing a patients *SPOP* mutation status and direct the use of DNA-PKcsi to bring about beneficial effects in aggressive late-stage PC.

9.2 Future work

9.2.1 Further work for the investigation of SPOP as a potential biomarker for PC

To fully validate the promising initial findings of the synthetic lethal relationship between SPOP mutation and DNA-PKcsi, similar assays in alternative PC cell backgrounds would provide more pre-clinical evidence of this therapeutic vulnerability. To support this work, deeper investigation into the mechanism of action causing the observed reductions in cell viability in response to DNA-PKci would be of great importance. Finally, to ultimately assess the translatability of this synthetic lethality, *in vivo* studies incorporating xenograft mouse models would greatly boost the impact of the presented findings. By grafting either the CWR22Rv1-AR-EK or CWR22Rv1-AR-EK-SPOP$^{-/-}$ clones into immune compromised mice and treating with NU7441 or vehicle controls around the calculated IC50, the effects on total tumour size and ultimately the tractability of the synthetic lethality in an *in vitro* setting can be further assessed in complex tumour environments, whilst also allowing the generation of material to undergo sectioning and subsequent immunohistochemical staining for a range of relevant PC targets (such as RAD51, γH2AX) to assess the resultant tumours formed within the mice. Ultimately

promising results from these *in vivo* experiments may potentially lead to assessing this potential treatment in an early-phase clinical trial.

Alongside this, work presented here, alongside the wealth of literature, indicates that *SPOP* mutations are a new key biomarker of interest and hold the potential to better direct patient treatment. As mentioned in Section 5.1, a substantial amount of work has been conducted into potentiating sensitivity to PARPi in PC, through exploitation of mutations or combination with other drugs. Using the models developed within this project, similar investigations into alternate DNA repair inhibitors could be conducted, either alone or combined with DNA-PKcsi treatments to enhance sensitivity. Expanding this further, high-throughput screening methods such as using an FDA approved compound library in *SPOP* knock-out model background may reveal novel drug susceptibilities and allow for repurposing of currently approved drugs to exploit *SPOP* mutations. Alongside the effects of loss of SPOP function within cells, published work has also investigated the structure of the SPOP protein and found it targetable with small compound inhibitors targeting the substrate binding cleft (Guo et al., 2016). *SPOP* mutations present in a large population of patients, however inhibiting SPOP may allow for benefits of combinations with DNA-PKcsi to be translated to all PC patients. Alongside assessing the translatability of the novel synthetic lethality described from this work in more patient representative *in* vivo models, investigation into combining SPOPi and DNA-PKcsi would be highly beneficial. Finally, outside of PC, SPOP has been found to mutated or downregulated in other cancers such as, liver, lung and endometrial (Clark et al., 2020). If results show translation of DNA-PKcsi having beneficial effects in pre-clinical PC models, potential future work to expanding this investigation into alternate cancer background could provide pan-cancer benefits.

To couple with the transcriptomic analysis presented here, and to give definitive lists of SPOP substrates, proteomic analysis using these SPOP knock-out models would be of great interest and potentially generate more useful information into alternate pathways worth investigating for new therapeutic sensitivities. Further, many negatively enriched hallmarks were observed when comparing the effects of SPOP knockout, SPOP mutations in patients and published literature overexpressing mutant SPOP. Consistent across all of these, repression was seen in the *TNFa signalling via NFKB* hallmark, which contradicts published literature. Running preliminary experiments into this signalling pathway within *in vitro* models may uncover the cause of this discrepancy with the literature and build the basis for further projects

investigating SPOP contributions to immune surveillance and potential contributions to inflammatory processes.

9.2.2 Future directions of DNA-based nanotechnology biomarker detection

A key point from the early experimental work conducted utilising DNA nanotechnology for the detection of mRNA was the demonstration of detection capabilities to a high sensitivity and selectivity in buffer and early indications of positive detection in human serum. Although these results yielded positive preliminary insights, drastic probe iteration would be necessary to detect physiologically relevant levels of mRNA from patient liquid biopsy samples. Factoring in later 2xSEP probe designs as well as developing downstream signal amplification steps could boost sensitivity if this work was to be taken further, potentially removing the need to extract material from patient samples if LODs were able to be substantially reduced. If these iterations on assay designs prove beneficial future work could be conducted in more representative samples, such as spiked-in experiments in healthy blood to demonstrate the translatability of this technology as a diagnostic tool.

Presented work into the detection of miRNAs within PC cell line lysates showed great promise and was able to compete to a similar level of commercial RT-qPCR based miRNA detection techniques. With the discrepancies seen between the probe-based detection, published NGS data and commercial RT-qPCR results, in order to complete this work in-house NGS analysis is necessary and would allow for conclusive evidence into the expression differences between the two cell lines tested. As with the mRNA work, if the probe detection system were to be fully validated and confirmed to outperform competitors, moving to assess detection within samples more representative of patient liquid biopsies would greatly increase the impact of this preliminary work.

With regards to the development of DNA nanotechnology-based PC detection, in total the presented work highlights the adaptability and potential of this as a biomarker detection platform. With downstream amplification elements added to the current 2xSEP designs, this system may be able to detect physiologically relevant levels of biomarkers from patient liquid biopsies. Although work presented here focussed on specific biomarkers, future probe systems can be generated for any mRNA/miRNA target due to the high adaptability of the technology. Thus, it may be beneficial to combine multiple biomarkers to potentially quantify certain

biomarkers such as PSA or AR-V7 whilst genotyping tumours through probes targeting common genetic aberrations common in PC such as *SPOP* mutations and *TMPRSS2:ERG* fusions. Together, with future iteration on the current assay designs, this technology could be used as a tool for not only diagnosis but to predict disease aggressiveness and ideally direct patient treatment.

Chapter 10 – References

AARON, L., FRANCO, O. E. & HAYWARD, S. W. 2016. Review of Prostate Anatomy and Embryology and the Etiology of Benign Prostatic Hyperplasia. *Urol Clin North Am,* 43**,** 279-88.

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.

ABRAMOVIC, I., ULAMEC, M., KATUSIC BOJANAC, A., BULIC-JAKUS, F., JEZEK, D. & SINCIC, N. 2020. miRNA in prostate cancer: challenges toward translation. *Epigenomics,* 12**,** 543-558.

ADAMAKI, M. & ZOUMPOURLIS, V. 2021. Prostate Cancer Biomarkers: From diagnosis to prognosis and precision-guided therapeutics. *Pharmacol Ther,* 228**,** 107932.

ADAMS, J. Y., LEAV, I., LAU, K. M., HO, S. M. & PFLUEGER, S. M. 2002. Expression of estrogen receptor beta in the fetal, neonatal, and prepubertal human prostate. *Prostate,* 52**,** 69-81.

AKBARI, M., PEÑA-DIAZ, J., ANDERSEN, S., LIABAKK, N. B., OTTERLEI, M. & KROKAN, H. E. 2009. Extracts of proliferating and non-proliferating human cells display different base excision pathways and repair fidelity. *DNA Repair (Amst),* 8**,** 834-43.

ALMEIDA, K. H. & SOBOL, R. W. 2007. A unified view of base excision repair: lesion-dependent protein complexes regulated by post-translational modification. *DNA Repair (Amst),* 6**,** 695-711.

ALVAREZ-GARCIA, V., BARTOS, C., KERAITE, I., TRIVEDI, U., BRENNAN, P. M., KERSAUDY-KERHOAS, M., GHARBI, K., OIKONOMIDOU, O. & LESLIE, N. R. 2018. A simple and robust real-time qPCR method for the detection of PIK3CA mutations. *Sci Rep,* 8**,** 4290.

AMBS, S., PRUEITT, R. L., YI, M., HUDSON, R. S., HOWE, T. M., PETROCCA, F., WALLACE, T. A., LIU, C. G., VOLINIA, S., CALIN, G. A., YFANTIS, H. G., STEPHENS, R. M. & CROCE, C. M. 2008. Genomic profiling of microRNA and messenger RNA reveals deregulated microRNA expression in prostate cancer. *Cancer Res,* 68**,** 6162-70.

AN, J., REN, S., MURPHY, S. J., DALANGOOD, S., CHANG, C., PANG, X., CUI, Y., WANG, L., PAN, Y., ZHANG, X., ZHU, Y., WANG, C., HALLING, G. C., CHENG, L., SUKOV, W. R., KARNES, R. J., VASMATZIS, G., ZHANG, Q., ZHANG, J., CHEVILLE, J. C., YAN, J., SUN, Y. & HUANG, H. 2015. Truncated ERG Oncoproteins from TMPRSS2-ERG Fusions Are Resistant to SPOP-Mediated Proteasome Degradation. *Mol Cell,* 59**,** 904-16.

AN, J., WANG, C., DENG, Y., YU, L. & HUANG, H. 2014. Destruction of full-length androgen receptor by wild-type SPOP, but not prostate-cancer-associated mutants. *Cell Rep,* 6**,** 657-69.

ANDRES, S. N., VERGNES, A., RISTIC, D., WYMAN, C., MODESTI, M. & JUNOP, M. 2012. A human XRCC4-XLF complex bridges DNA. *Nucleic Acids Res,* 40**,** 1868-78.

ANDREWS, S. 2010. *FastQC: A Quality Control Tool for High Throughput Sequence Data* [Online]. Available: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/ [Accessed].

ANTONARAKIS, E. S., LU, C., LUBER, B., WANG, H., CHEN, Y., NAKAZAWA, M., NADAL, R., PALLER, C. J., DENMEADE, S. R., CARDUCCI, M. A., EISENBERGER, M. A. & LUO, J. 2015. Androgen Receptor Splice Variant 7 and Efficacy of Taxane Chemotherapy in Patients With Metastatic Castration-Resistant Prostate Cancer. *JAMA Oncol,* 1**,** 582-91.

ANTONARAKIS, E. S., LU, C., WANG, H., LUBER, B., NAKAZAWA, M., ROESER, J. C., CHEN, Y., MOHAMMAD, T. A., CHEN, Y., FEDOR, H. L., LOTAN, T. L., ZHENG, Q., DE MARZO, A. M., ISAACS, J. T., ISAACS, W. B., NADAL, R., PALLER, C. J., DENMEADE, S. R., CARDUCCI, M. A., EISENBERGER, M. A. & LUO, J. 2014. AR-V7 and resistance to enzalutamide and abiraterone in prostate cancer. *N Engl J Med,* 371**,** 1028-38.

ATTARD, G., PARKER, C., EELES, R. A., SCHRÖDER, F., TOMLINS, S. A., TANNOCK, I., DRAKE, C. G. & DE BONO, J. S. 2016. Prostate cancer. *Lancet,* 387**,** 70-82.

ATTARD, G., REID, A. H., YAP, T. A., RAYNAUD, F., DOWSETT, M., SETTATREE, S., BARRETT, M., PARKER, C., MARTINS, V., FOLKERD, E., CLARK, J., COOPER, C. S., KAYE, S. B., DEARNALEY, D., LEE, G. & DE BONO, J. S. 2008. Phase I clinical trial of a selective inhibitor of CYP17, abiraterone acetate, confirms that castration-resistant prostate cancer commonly remains hormone driven. *J Clin Oncol,* 26**,** 4563-71.

AUGELLO, M. A., LIU, D., DEONARINE, L. D., ROBINSON, B. D., HUANG, D., STELLOO, S., BLATTNER, M., DOANE, A. S., WONG, E. W. P., CHEN, Y., RUBIN, M. A., BELTRAN, H., ELEMENTO, O., BERGMAN, A. M., ZWART, W., SBONER, A., DEPHOURE, N. & BARBIERI, C. E. 2019. CHD1 Loss Alters AR Binding at Lineage-Specific Enhancers and Modulates Distinct Transcriptional Programs to Drive Prostate Tumorigenesis. *Cancer Cell,* 35**,** 603-617.e8.

BAKER, M. E. 2002. Albumin, steroid hormones and the origin of vertebrates. *J Endocrinol,* 175**,** 121-7.

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.

BEATO, M., HERRLICH, P. & SCHÜTZ, G. 1995. Steroid hormone receptors: many actors in search of a plot. *Cell,* 83**,** 851-7.

BECKER, J. R., CLIFFORD, G., BONNET, C., GROTH, A., WILSON, M. D. & CHAPMAN, J. R. 2021. BARD1 reads H2A lysine 15 ubiquitination to direct homologous recombination. *Nature,* 596**,** 433-437.

BEDFORD, L., LOWE, J., DICK, L. R., MAYER, R. J. & BROWNELL, J. E. 2011. Ubiquitin-like protein conjugation and the ubiquitin-proteasome system as drug targets. *Nat Rev Drug Discov,* 10**,** 29-46.

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. 2008. The Risk of Prostate Cancer amongst Black Men in the United Kingdom: The PROCESS Cohort Study. *European Urology,* 53**,** 99-105.

BERGENGREN, O., PEKALA, K. R., MATSOUKAS, K., FAINBERG, J., MUNGOVAN, S. F., BRATT, O., BRAY, F., BRAWLEY, O., LUCKENBAUGH, A. N., MUCCI, L., MORGAN, T. M. & CARLSSON, S. V. 2023. 2022 Update on Prostate Cancer Epidemiology and Risk Factors-A Systematic Review. *Eur Urol,* 84**,** 191-206.

BERNASOCCHI, T., EL TEKLE, G., BOLIS, M., MUTTI, A., VALLERGA, A., BRANDT, L. P., SPRIANO, F., SVINKINA, T., ZOMA, M., CESERANI, V., RINALDI, A., JANOUSKOVA, H., BOSSI, D., CAVALLI, M., MOSOLE, S., GEIGER, R., DONG, Z., YANG, C. G., ALBINO, D., RINALDI, A., SCHRAML, P., LINDER, S., CARBONE, G. M., ALIMONTI, A., BERTONI, F., MOCH, H., CARR, S. A., ZWART, W., KRUITHOF-DE JULIO, M., RUBIN, M. A., UDESHI, N. D. & THEURILLAT, J. P. 2021. Dual functions of SPOP and ERG dictate androgen therapy responses in prostate cancer. *Nat Commun,* 12**,** 734.

BEUCHER, A., BIRRAUX, J., TCHOUANDONG, L., BARTON, O., SHIBATA, A., CONRAD, S., GOODARZI, A. A., KREMPLER, A., JEGGO, P. A. & LÖBRICH, M. 2009. ATM and Artemis promote homologous recombination of radiation-induced DNA double-strand breaks in G2. *Embo j,* 28**,** 3413-27.

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.

BLATTNER, M., LEE, D. J., O'REILLY, C., PARK, K., MACDONALD, T. Y., KHANI, F., TURNER, K. R., CHIU, Y. L., WILD, P. J., DOLGALEV, I., HEGUY, A., SBONER, A., RAMAZANGOLU, S., HIERONYMUS, H., SAWYERS, C., TEWARI, A. K., MOCH, H., YOON, G. S., KNOWN, Y. C., ANDRÉN, O., FALL, K., DEMICHELIS, F., MOSQUERA, J. M., ROBINSON, B. D., BARBIERI, C. E. & RUBIN, M. A. 2014. SPOP mutations in prostate cancer across demographically diverse patient cohorts. *Neoplasia,* 16**,** 14-20.

BLATTNER, M., LIU, D., ROBINSON, B. D., HUANG, D., POLIAKOV, A., GAO, D., NATARAJ, S., DEONARINE, L. D., AUGELLO, M. A., SAILER, V., PONNALA, L., ITTMANN, M., CHINNAIYAN, A. M., SBONER, A., CHEN, Y., RUBIN, M. A. & BARBIERI, C. E. 2017. SPOP Mutation Drives Prostate Tumorigenesis In Vivo through Coordinate Regulation of PI3K/mTOR and AR Signaling. *Cancer Cell,* 31**,** 436-451.

BLUME-JENSEN, P., BERMAN, D. M., RIMM, D. L., SHIPITSIN, M., PUTZI, M., NIFONG, T. P., SMALL, C., CHOUDHURY, S., CAPELA, T., COUPAL, L., ERNST, C., HURLEY, A., KAPRELYANTS, A., CHANG, H., GILADI, E., NARDONE, J., DUNYAK, J., LODA, M., KLEIN, E. A., MAGI-GALLUZZI, C., LATOUR, M., EPSTEIN, J. I., KANTOFF, P. & SAAD, F. 2015. Development and clinical validation of an in situ biopsy-based multimarker assay for risk stratification in prostate cancer. *Clin Cancer Res,* 21**,** 2591-600.

BOLL, K., REICHE, K., KASACK, K., MÖRBT, N., KRETZSCHMAR, A. K., TOMM, J. M., VERHAEGH, G., SCHALKEN, J., VON BERGEN, M., HORN, F. & HACKERMÜLLER, J. 2013. MiR-130a, miR-203 and miR-205 jointly repress key oncogenic pathways and are downregulated in prostate carcinoma. *Oncogene,* 32**,** 277-85.

BOOTH, L., CRUICKSHANKS, N., RIDDER, T., DAI, Y., GRANT, S. & DENT, P. 2013. PARP and CHK inhibitors interact to cause DNA damage and cell death in mammary carcinoma cells. *Cancer Biology & Therapy,* 14**,** 458-465.

BOTUYAN, M. V., LEE, J., WARD, I. M., KIM, J. E., THOMPSON, J. R., CHEN, J. & MER, G. 2006. Structural basis for the methylation state-specific recognition of histone H4-K20 by 53BP1 and Crb2 in DNA repair. *Cell,* 127**,** 1361-73.

BOUCHARD, J. J., OTERO, J. H., SCOTT, D. C., SZULC, E., MARTIN, E. W., SABRI, N., GRANATA, D., MARZAHN, M. R., LINDORFF-LARSEN, K., SALVATELLA, X., SCHULMAN, B. A. & MITTAG, T. 2018. Cancer Mutations of the Tumor Suppressor SPOP Disrupt the Formation of Active, Phase-Separated Compartments. *Mol Cell,* 72**,** 19-36.e8.

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.

BOYSEN, G., RODRIGUES, D. N., RESCIGNO, P., SEED, G., DOLLING, D., RIISNAES, R., CRESPO, M., ZAFEIRIOU, Z., SUMANASURIYA, S., BIANCHINI, D., HUNT, J., MOLONEY, D., PEREZ-LOPEZ, R., TUNARIU, N., MIRANDA, S., FIGUEIREDO, I., FERREIRA, A., CHRISTOVA, R., GIL, V., AZIZ, S., BERTAN, C., DE OLIVEIRA, F. M., ATKIN, M., CLARKE, M., GOODALL, J., SHARP, A., MACDONALD, T., RUBIN, M. A., YUAN, W., BARBIERI, C. E., CARREIRA, S., MATEO, J. & DE BONO, J. S. 2018. SPOP-Mutated/CHD1-Deleted Lethal Prostate Cancer and Abiraterone Sensitivity. *Clin Cancer Res,* 24**,** 5585-5593.

BROWN, C. J., GOSS, S. J., LUBAHN, D. B., JOSEPH, D. R., WILSON, E. M., FRENCH, F. S. & WILLARD, H. F. 1989. Androgen receptor locus on the human X chromosome: regional localization to Xq11-12 and description of a DNA polymorphism. *Am J Hum Genet,* 44**,** 264-9.

BRYANT, H. E. & HELLEDAY, T. 2006. Inhibition of poly (ADP-ribose) polymerase activates ATM which is required for subsequent homologous recombination repair. *Nucleic Acids Research,* 34**,** 1685-1691.

BRYANT, H. E., SCHULTZ, N., THOMAS, H. D., PARKER, K. M., FLOWER, D., LOPEZ, E., KYLE, S., MEUTH, M., CURTIN, N. J. & HELLEDAY, T. 2005. Specific killing of BRCA2-deficient tumours with inhibitors of poly(ADP-ribose) polymerase. *Nature,* 434**,** 913-917.

BUETOW, L. & HUANG, D. T. 2016. Structural insights into the catalysis and regulation of E3 ubiquitin ligases. *Nat Rev Mol Cell Biol,* 17**,** 626-42.

BUNCE, M. W., BORONENKOV, I. V. & ANDERSON, R. A. 2008. Coordinated activation of the nuclear ubiquitin ligase Cul3-SPOP by the generation of phosphatidylinositol 5-phosphate. *J Biol Chem,* 283**,** 8678-86.

BURLESON, M., DENG, J. J., QIN, T., DUONG, T. M., YAN, Y., GU, X., DAS, D., EASLEY, A., LISS, M. A., YEW, P. R., BEDOLLA, R., KUMAR, A. P., HUANG, T. H., ZOU, Y., CHEN, Y., CHEN, C. L., HUANG, H., SUN, L. Z. & BOYER, T. G. 2022. GLI3 Is Stabilized by SPOP Mutations and Promotes Castration Resistance via Functional Cooperation with Androgen Receptor in Prostate Cancer. *Mol Cancer Res,* 20**,** 62-76.

CALDECOTT, K. W. 2014. DNA single-strand break repair. *Experimental Cell Research,* 329**,** 2-8.

CALLEWAERT, L., VAN TILBORGH, N. & CLAESSENS, F. 2006. Interplay between two hormone-independent activation domains in the androgen receptor. *Cancer Res,* 66**,** 543-53.

CARLSSON, S., MASCHINO, A., SCHRÖDER, F., BANGMA, C., STEYERBERG, E. W., VAN DER KWAST, T., VAN LEENDERS, G., VICKERS, A., LILJA, H. & ROOBOL, M. J. 2013. Predictive value of four kallikrein markers for pathologically insignificant compared with aggressive prostate cancer in radical prostatectomy specimens: results from the European Randomized Study of Screening for Prostate Cancer section Rotterdam. *Eur Urol,* 64**,** 693-9.

CATALONA, W. J., PARTIN, A. W., SANDA, M. G., WEI, J. T., KLEE, G. G., BANGMA, C. H., SLAWIN, K. M., MARKS, L. S., LOEB, S., BROYLES, D. L., SHIN, S. S., CRUZ, A. B., CHAN, D. W., SOKOLL, L. J., ROBERTS, W. L., VAN SCHAIK, R. H. & MIZRAHI, I. A. 2011. A multicenter study of [-2]pro-prostate specific antigen combined with prostate specific antigen and

free prostate specific antigen for prostate cancer detection in the 2.0 to 10.0 ng/ml prostate specific antigen range. *J Urol,* 185**,** 1650-5.

CATALONA, W. J., SMITH, D. S., RATLIFF, T. L., DODDS, K. M., COPLEN, D. E., YUAN, J. J., PETROS, J. A. & ANDRIOLE, G. L. 1991. Measurement of prostate-specific antigen in serum as a screening test for prostate cancer. *N Engl J Med,* 324**,** 1156-61.

CHATTERJEE, N. & WALKER, G. C. 2017. Mechanisms of DNA damage, repair, and mutagenesis. *Environ Mol Mutagen,* 58**,** 235-263.

CHATTERJEE, P., SCHWEIZER, M. T., LUCAS, J. M., COLEMAN, I., NYQUIST, M. D., FRANK, S. B., THARAKAN, R., MOSTAGHEL, E., LUO, J., PRITCHARD, C. C., LAM, H. M., COREY, E., ANTONARAKIS, E. S., DENMEADE, S. R. & NELSON, P. S. 2019. Supraphysiological androgens suppress prostate cancer growth through androgen receptor-mediated DNA damage. *J Clin Invest,* 129**,** 4245-4260.

CHEN, H. Y. & CHEN, R. H. 2016. Cullin 3 Ubiquitin Ligases in Cancer Biology: Functions and Therapeutic Implications. *Front Oncol,* 6**,** 113.

CHEN, L., NIEVERA, C. J., LEE, A. Y. & WU, X. 2008a. Cell cycle-dependent complex formation of BRCA1.CtIP.MRN is important for DNA double-strand break repair. *J Biol Chem,* 283**,** 7713-20.

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.

CICCIA, A. & ELLEDGE, S. J. 2010. The DNA damage response: making it safe to play with knives. *Mol Cell,* 40**,** 179-204.

CISZEWSKI, W. M., TAVECCHIO, M., DASTYCH, 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.

CLARK, A. & BURLESON, M. 2020. SPOP and cancer: a systematic review. *Am J Cancer Res,* 10**,** 704-726.

CLAVIJO, R. I. & HSIAO, W. 2018. Update on male reproductive endocrinology. *Transl Androl Urol,* 7**,** S367-s372.

CLINCKEMALIE, L., VANDERSCHUEREN, D., BOONEN, S. & CLAESSENS, F. 2012. The hinge region in androgen receptor control. *Mol Cell Endocrinol,* 358**,** 1-8.

COFFEY, K. & ROBSON, C. N. 2012. Regulation of the androgen receptor by post-translational modifications. *J Endocrinol,* 215**,** 221-37.

CORNFORD, P., BELLMUNT, J., BOLLA, M., BRIERS, E., DE SANTIS, M., GROSS, T., HENRY, A. M., JONIAU, S., LAM, T. B., MASON, M. D., VAN DER POEL, H. G., VAN DER KWAST, T. H., ROUVIÈRE, O., WIEGEL, T. & MOTTET, N. 2017. EAU-ESTRO-SIOG Guidelines on Prostate Cancer. Part II: Treatment of Relapsing, Metastatic, and Castration-Resistant Prostate Cancer. *Eur Urol,* 71**,** 630-642.

CORTEZ, M. A., BUESO-RAMOS, C., FERDIN, J., LOPEZ-BERESTEIN, G., SOOD, A. K. & CALIN, G. A. 2011. MicroRNAs in body fluids--the mix of hormones and biomarkers. *Nat Rev Clin Oncol,* 8**,** 467-77.

CRAWFORD, E. D., SCHELLHAMMER, P. F., MCLEOD, D. G., MOUL, J. W., HIGANO, C. S., SHORE, N., DENIS, L., IVERSEN, P., EISENBERGER, M. A. & LABRIE, F. 2018. Androgen Receptor Targeted Treatments of Prostate Cancer: 35 Years of Progress with Antiandrogens. *J Urol,* 200**,** 956-966.

CULIG, Z. & SANTER, F. R. 2014. Androgen receptor signaling in prostate cancer. *Cancer Metastasis Rev,* 33**,** 413-27.

CULLEN, J., ROSNER, I. L., BRAND, T. C., ZHANG, N., TSIATIS, A. C., MONCUR, J., ALI, A., CHEN, Y., KNEZEVIC, D., MADDALA, T., LAWRENCE, H. J., FEBBO, P. G., SRIVASTAVA, S., SESTERHENN, I. A. & MCLEOD, D. G. 2015. A Biopsy-based 17-gene Genomic Prostate Score Predicts Recurrence After Radical Prostatectomy and Adverse Surgical Pathology

in a Racially Diverse Population of Men with Clinically Low- and Intermediate-risk Prostate Cancer. *Eur Urol,* 68**,** 123-31.

CUNEO, M. J. & MITTAG, T. 2019. The ubiquitin ligase adaptor SPOP in cancer. *Febs j,* 286**,** 3946-3958.

CUNEO, M. J., O'FLYNN, B. G., LO, Y. H., SABRI, N. & MITTAG, T. 2023. Higher-order SPOP assembly reveals a basis for cancer mutant dysregulation. *Mol Cell,* 83**,** 731-745.e4.

CUNHA, G. R., VEZINA, C. M., ISAACSON, D., RICKE, W. A., TIMMS, B. G., CAO, M., FRANCO, O. & BASKIN, L. S. 2018. Development of the human prostate. *Differentiation,* 103**,** 24-45.

CUTRESS, M. L., WHITAKER, H. C., MILLS, I. G., STEWART, M. & NEAL, D. E. 2008. Structural basis for the nuclear import of the human androgen receptor. *J Cell Sci,* 121**,** 957-68.

DART, D. A., KOUSHYAR, S., LANNING, B. E. & JIANG, W. 2019. MiR-221 Is Specifically Elevated in PC3 Cells and its Deletion Reduces Adhesion, Motility and Growth. *Anticancer Res,* 39**,** 5311-5327.

DE BONO, J. S., LOGOTHETIS, C. J., MOLINA, A., FIZAZI, K., NORTH, S., CHU, L., CHI, K. N., JONES, R. J., GOODMAN, O. B., JR., SAAD, F., STAFFURTH, J. N., MAINWARING, P., HARLAND, S., FLAIG, T. W., HUTSON, T. E., CHENG, T., PATTERSON, H., HAINSWORTH, J. D., RYAN, C. J., STERNBERG, C. N., ELLARD, S. L., FLÉCHON, A., SALEH, M., SCHOLZ, M., EFSTATHIOU, E., ZIVI, A., BIANCHINI, D., LORIOT, Y., CHIEFFO, N., KHEOH, T., HAQQ, C. M. & SCHER, H. I. 2011. Abiraterone and increased survival in metastatic prostate cancer. *N Engl J Med,* 364**,** 1995-2005.

DEHM, S. M. & TINDALL, D. J. 2006. Molecular regulation of androgen action in prostate cancer. *J Cell Biochem,* 99**,** 333-44.

DEHM, S. M. & TINDALL, D. J. 2011. Alternatively spliced androgen receptor variants. *Endocr Relat Cancer,* 18**,** R183-96.

DIANOV, G. L. & HÜBSCHER, U. 2013. Mammalian base excision repair: the forgotten archangel. *Nucleic Acids Res,* 41**,** 3483-90.

DILLON, M. T., BERGERHOFF, K. F., PEDERSEN, M., WHITTOCK, H., CRESPO-RODRIGUEZ, E., PATIN, E. C., PEARSON, A., SMITH, H. G., PAGET, J. T. E., PATEL, R. R., FOO, S., BOZHANOVA, G., RAGULAN, C., FONTANA, E., DESAI, K., WILKINS, A. C., SADANANDAM, A., MELCHER, A., MCLAUGHLIN, M. & HARRINGTON, K. J. 2019. ATR Inhibition Potentiates the Radiation-induced Inflammatory Tumor Microenvironment. *Clinical Cancer Research,* 25**,** 3392-3403.

DING, D., XU, L., MENON, M., REDDY, G. P. & BARRACK, E. R. 2004. Effect of a short CAG (glutamine) repeat on human androgen receptor function. *Prostate,* 58**,** 23-32.

DOBBS, T. A., TAINER, J. A. & LEES-MILLER, S. P. 2010. A structural model for regulation of NHEJ by DNA-PKcs autophosphorylation. *DNA Repair (Amst),* 9**,** 1307-14.

DOESBURG, P., KUIL, C. W., BERREVOETS, C. A., STEKETEE, K., FABER, P. W., MULDER, E., BRINKMANN, A. O. & TRAPMAN, J. 1997. Functional in vivo interaction between the amino-terminal, transactivation domain and the ligand binding domain of the androgen receptor. *Biochemistry,* 36**,** 1052-64.

DONG, Z., XUE, X., LIANG, H., GUAN, J. & CHANG, L. 2021. DNA Nanomachines for Identifying Cancer Biomarkers in Body Fluids and Cells. *Anal Chem,* 93**,** 1855-1865.

DONOVAN, M. J., NOERHOLM, M., BENTINK, S., BELZER, S., SKOG, J., O'NEILL, V., COCHRAN, J. S. & BROWN, G. A. 2015. A molecular signature of PCA3 and ERG exosomal RNA from non-DRE urine is predictive of initial prostate biopsy result. *Prostate Cancer Prostatic Dis,* 18**,** 370-5.

EL BEZAWY, R., TRIPARI, M., PERCIO, S., CICCHETTI, A., TORTORETO, M., STUCCHI, C., TINELLI, S., ZUCO, V., DOLDI, V., GANDELLINI, P., VALDAGNI, R. & ZAFFARONI, N. 2020. SPOP Deregulation Improves the Radiation Response of Prostate Cancer Models by Impairing DNA Damage Repair. *Cancers (Basel),* 12.

ETZIONI, R., TSODIKOV, A., MARIOTTO, A., SZABO, A., FALCON, S., WEGELIN, J., DITOMMASO, D., KARNOFSKI, K., GULATI, R., PENSON, D. F. & FEUER, E. 2008. Quantifying the role of PSA screening in the US prostate cancer mortality decline. *Cancer Causes Control,* 19**,** 175-81.

EWING, C. M., RAY, A. M., LANGE, E. M., ZUHLKE, K. A., ROBBINS, C. M., TEMBE, W. D., WILEY, K. E., ISAACS, S. D., JOHNG, D., WANG, Y., BIZON, C., YAN, G., GIELZAK, M., PARTIN, A. W., SHANMUGAM, V., IZATT, T., SINARI, S., CRAIG, D. W., ZHENG, S. L., WALSH, P. C., MONTIE, J. E., XU, J., CARPTEN, J. D., ISAACS, W. B. & COONEY, K. A. 2012. Germline mutations in HOXB13 and prostate-cancer risk. *N Engl J Med,* 366**,** 141-9.

FABRIS, L., CEDER, Y., CHINNAIYAN, A. M., JENSTER, G. W., SORENSEN, K. D., TOMLINS, S., VISAKORPI, T. & CALIN, G. A. 2016. The Potential of MicroRNAs as Prostate Cancer Biomarkers. *Eur Urol,* 70**,** 312-22.

FAGBEMI, A. F., ORELLI, B. & SCHÄRER, O. D. 2011. Regulation of endonuclease activity in human nucleotide excision repair. *DNA Repair (Amst),* 10**,** 722-9.

FARMER, H., MCCABE, N., LORD, C. J., TUTT, A. N. J., JOHNSON, D. A., RICHARDSON, T. B., SANTAROSA, M., DILLON, K. J., HICKSON, I., KNIGHTS, C., MARTIN, N. M. B., JACKSON, S. P., SMITH, G. C. M. & ASHWORTH, A. 2005. Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. *Nature,* 434**,** 917-921.

FENG, D. C., ZHU, W. Z., WANG, J., LI, D. X., SHI, X., XIONG, Q., YOU, J., HAN, P., QIU, S., WEI, Q. & YANG, L. 2024. The implications of single-cell RNA-seq analysis in prostate cancer: unraveling tumor heterogeneity, therapeutic implications and pathways towards personalized therapy. *Mil Med Res,* 11**,** 21.

FENG, E., RYDZEWSKI, N. R., ZHANG, M., LUNDBERG, A., BOOTSMA, M., HELZER, K. T., LANG, J. M., AGGARWAL, R., SMALL, E. J., QUIGLEY, D. A., SJÖSTRÖM, M. & ZHAO, S. G. 2022. Intrinsic Molecular Subtypes of Metastatic Castration-Resistant Prostate Cancer. *Clin Cancer Res,* 28**,** 5396-5404.

FENG, X., TUBBS, A., ZHANG, C., TANG, M., SRIDHARAN, S., WANG, C., JIANG, D., SU, D., ZHANG, H., CHEN, Z., NIE, L., XIONG, Y., HUANG, M., NUSSENZWEIG, A. & CHEN, J. 2020. ATR inhibition potentiates ionizing radiation-induced interferon response via cytosolic nucleic acid-sensing pathways. *The EMBO Journal,* 39**,** e104036.

FERRO, P., CATALANO, M. G., DELL'EVA, R., FORTUNATI, N. & PFEFFER, U. 2002. The androgen receptor CAG repeat: a modifier of carcinogenesis? *Mol Cell Endocrinol,* 193**,** 109-20.

FILELLA, X. & FOJ, L. 2018. Novel Biomarkers for Prostate Cancer Detection and Prognosis. *Adv Exp Med Biol,* 1095**,** 15-39.

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., NIKKILÄ, 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. *Nature Communications,* 10**,** 5065.

FRADET-TURCOTTE, A., CANNY, M. D., ESCRIBANO-DÍAZ, C., ORTHWEIN, A., LEUNG, C. C. Y., HUANG, H., LANDRY, M.-C., KITEVSKI-LEBLANC, J., NOORDERMEER, S. M., SICHERI, F. & DUROCHER, D. 2013. 53BP1 is a reader of the DNA-damage-induced H2A Lys 15 ubiquitin mark. *Nature,* 499**,** 50-54.

FU, M., WANG, C., REUTENS, A. T., WANG, J., ANGELETTI, R. H., SICONOLFI-BAEZ, L., OGRYZKO, V., AVANTAGGIATI, M. L. & PESTELL, R. G. 2000. p300 and p300/cAMP-response element-binding protein-associated factor acetylate the androgen receptor at sites governing hormone-dependent transactivation. *J Biol Chem,* 275**,** 20853-60.

GALARDI, S., MERCATELLI, N., GIORDA, E., MASSALINI, S., FRAJESE, G. V., CIAFRÈ, S. A. & FARACE, M. G. 2007. miR-221 and miR-222 expression affects the proliferation potential of human prostate carcinoma cell lines by targeting p27Kip1. *J Biol Chem,* 282**,** 23716-24.

GAN, L., CHEN, S., WANG, Y., WATAHIKI, A., BOHRER, L., SUN, Z., WANG, Y. & HUANG, H. 2009. Inhibition of the androgen receptor as a novel mechanism of taxol chemotherapy in prostate cancer. *Cancer Res,* 69**,** 8386-94.

GAN, W., DAI, X., LUNARDI, A., LI, Z., INUZUKA, H., LIU, P., VARMEH, S., ZHANG, J., CHENG, L., SUN, Y., ASARA, J. M., BECK, A. H., HUANG, J., PANDOLFI, P. P. & WEI, W. 2015. SPOP Promotes Ubiquitination and Degradation of the ERG Oncoprotein to Suppress Prostate Cancer Progression. *Mol Cell,* 59**,** 917-30.

GAUGHAN, L., LOGAN, I. R., COOK, S., NEAL, D. E. & ROBSON, C. N. 2002. Tip60 and histone deacetylase 1 regulate androgen receptor activity through changes to the acetylation status of the receptor. *J Biol Chem,* 277**,** 25904-13.

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.

GELMANN, E. P. 2002. Molecular biology of the androgen receptor. *J Clin Oncol,* 20**,** 3001-15.

GENG, C., HE, B., XU, L., BARBIERI, C. E., EEDUNURI, V. K., CHEW, S. A., ZIMMERMANN, M., BOND, R., SHOU, J., LI, C., BLATTNER, M., LONARD, D. M., DEMICHELIS, F., COARFA, C., RUBIN, M. A., ZHOU, P., O'MALLEY, B. W. & MITSIADES, N. 2013. Prostate cancer-associated mutations in speckle-type POZ protein (SPOP) regulate steroid receptor coactivator 3 protein turnover. *Proc Natl Acad Sci U S A,* 110**,** 6997-7002.

GENG, C., KAOCHAR, S., LI, M., RAJAPAKSHE, K., FISKUS, W., DONG, J., FOLEY, C., DONG, B., ZHANG, L., KWON, O. J., SHAH, S. S., BOLAKI, M., XIN, L., ITTMANN, M., O'MALLEY, B. W., COARFA, C. & MITSIADES, N. 2017. SPOP regulates prostate epithelial cell proliferation and promotes ubiquitination and turnover of c-MYC oncoprotein. *Oncogene,* 36**,** 4767-4777.

GENG, C., RAJAPAKSHE, K., SHAH, S. S., SHOU, J., EEDUNURI, V. K., FOLEY, C., FISKUS, W., RAJENDRAN, M., CHEW, S. A., ZIMMERMANN, M., BOND, R., HE, B., COARFA, C. &

MITSIADES, N. 2014. Androgen receptor is the key transcriptional mediator of the tumor suppressor SPOP in prostate cancer. *Cancer Res,* 74**,** 5631-43.

GENG, C., ZHANG, M. C., MANYAM, G. C., VYKOUKAL, J. V., FAHRMANN, J. F., PENG, S., WU, C., PARK, S., KONDRAGANTI, S., WANG, D., ROBINSON, B. D., LODA, M., BARBIERI, C. E., YAP, T. A., CORN, P. G., HANASH, S., BROOM, B. M., PILIÉ, P. G. & THOMPSON, T. C. 2023. SPOP Mutations Target STING1 Signaling in Prostate Cancer and Create Therapeutic Vulnerabilities to PARP Inhibitor-Induced Growth Suppression. *Clin Cancer Res,* 29**,** 4464-4478.

GLEASON, D. F. & MELLINGER, G. T. 1974. Prediction of prognosis for prostatic adenocarcinoma by combined histological grading and clinical staging. *J Urol,* 111**,** 58-64.

GLINSKY, G. V., GLINSKII, A. B., STEPHENSON, A. J., HOFFMAN, R. M. & GERALD, W. L. 2004. Gene expression profiling predicts clinical outcome of prostate cancer. *J Clin Invest,* 113**,** 913-23.

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.

GOTO, Y., KOJIMA, S., NISHIKAWA, R., KUROZUMI, A., KATO, M., ENOKIDA, H., MATSUSHITA, R., YAMAZAKI, K., ISHIDA, Y., NAKAGAWA, M., NAYA, Y., ICHIKAWA, T. & SEKI, N. 2015. MicroRNA expression signature of castration-resistant prostate cancer: the microRNA-221/222 cluster functions as a tumour suppressor and disease progression marker. *British Journal of Cancer,* 113**,** 1055-1065.

GRBESA, I., AUGELLO, M. A., LIU, D., MCNALLY, D. R., GAFFNEY, C. D., HUANG, D., LIN, K., IVENITSKY, D., GOUELI, R., ROBINSON, B. D., KHANI, F., DEONARINE, L. D., BLATTNER, M., ELEMENTO, O., DAVICIONI, E., SBONER, A. & BARBIERI, C. E. 2021. Reshaping of the androgen-driven chromatin landscape in normal prostate cells by early cancer drivers and effect on therapeutic sensitivity. *Cell Rep,* 36**,** 109625.

GROELLY, F. J., FAWKES, M., DAGG, R. A., BLACKFORD, A. N. & TARSOUNAS, M. 2023. Targeting DNA damage response pathways in cancer. *Nat Rev Cancer,* 23**,** 78-94.

GRÖNBERG, H. 2003. Prostate cancer epidemiology. *The Lancet,* 361**,** 859-864.

GROSKOPF, J., AUBIN, S. M., DERAS, I. L., BLASE, A., BODRUG, S., CLARK, C., BRENTANO, S., MATHIS, J., PHAM, J., MEYER, T., CASS, M., HODGE, P., MACAIRAN, M. L., MARKS, L. S. & RITTENHOUSE, H. 2006. APTIMA PCA3 molecular urine test: development of a method to aid in the diagnosis of prostate cancer. *Clin Chem,* 52**,** 1089-95.

GRYPARI, I. M., TZELEPI, V. & GYFTOPOULOS, K. 2023. DNA Damage Repair Pathways in Prostate Cancer: A Narrative Review of Molecular Mechanisms, Emerging Biomarkers and Therapeutic Targets in Precision Oncology. *Int J Mol Sci,* 24.

GULHAN, D. C., LEE, J. J.-K., MELLONI, G. E. M., CORTÉS-CIRIANO, I. & PARK, P. J. 2019. Detecting the mutational signature of homologous recombination deficiency in clinical samples. *Nature Genetics,* 51**,** 912-919.

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.

GUO, Z. Q., ZHENG, T., CHEN, B., LUO, C., OUYANG, S., GONG, S., LI, J., MAO, L. L., LIAN, F., YANG, Y., HUANG, Y., LI, L., LU, J., ZHANG, B., ZHOU, L., DING, H., GAO, Z., ZHOU, L., LI, G., ZHOU, R., CHEN, K., LIU, J., WEN, Y., GONG, L., KE, Y., YANG, S. D., QIU, X. B., ZHANG, N., REN, J., ZHONG, D., YANG, C. G., LIU, J. & JIANG, H. 2016. Small-Molecule Targeting of E3 Ligase Adaptor SPOP in Kidney Cancer. *Cancer Cell,* 30**,** 474-484.

HAMMEL, M., REY, M., YU, Y., MANI, R. S., CLASSEN, S., LIU, M., PIQUE, M. E., FANG, S., MAHANEY, B. L., WEINFELD, M., SCHRIEMER, D. C., LEES-MILLER, S. P. & TAINER, J. A. 2011. XRCC4 protein interactions with XRCC4-like factor (XLF) create an extended grooved scaffold for DNA ligation and double strand break repair. *J Biol Chem,* 286**,** 32638-50.

HAMMEL, M., YU, Y., MAHANEY, B. L., CAI, B., YE, R., PHIPPS, B. M., RAMBO, R. P., HURA, G. L., PELIKAN, M., SO, S., ABOLFATH, R. M., CHEN, D. J., LEES-MILLER, S. P. & TAINER, J. A. 2010. Ku and DNA-dependent protein kinase dynamic conformations and assembly regulate DNA binding and the initial non-homologous end joining complex. *J Biol Chem,* 285**,** 1414-23.

HANAHAN, D. 2022. Hallmarks of Cancer: New Dimensions. *Cancer Discovery,* 12**,** 31-46.

HARPER, J. W. & ELLEDGE, S. J. 2007. The DNA damage response: ten years after. *Mol Cell,* 28**,** 739-45.

HARRIS, W. P., MOSTAGHEL, E. A., NELSON, P. S. & MONTGOMERY, B. 2009. Androgen deprivation therapy: progress in understanding mechanisms of resistance and optimizing androgen depletion. *Nat Clin Pract Urol,* 6**,** 76-85.

HE, B., KEMPPAINEN, J. A., VOEGEL, J. J., GRONEMEYER, H. & WILSON, E. M. 1999. Activation function 2 in the human androgen receptor ligand binding domain mediates interdomain communication with the NH(2)-terminal domain. *J Biol Chem,* 274**,** 37219-25.

HEGDE, M. L., HAZRA, T. K. & MITRA, S. 2008. Early steps in the DNA base excision/single-strand interruption repair pathway in mammalian cells. *Cell Res,* 18**,** 27-47.

HJORTH-JENSEN, K., MAYA-MENDOZA, A., DALGAARD, N., SIGURÐSSON, J. O., BARTEK, J., IGLESIAS-GATO, D., OLSEN, J. V. & FLORES-MORALES, A. 2018. SPOP promotes transcriptional expression of DNA repair and replication factors to prevent replication stress and genomic instability. *Nucleic Acids Res,* 46**,** 9484-9495.

HONG, M., XU, L., XUE, Q., LI, L. & TANG, B. 2016. Fluorescence Imaging of Intracellular Telomerase Activity Using Enzyme-Free Signal Amplification. *Anal Chem,* 88**,** 12177-12182.

HQO 2017. Prolaris Cell Cycle Progression Test for Localized Prostate Cancer: A Health Technology Assessment. *Ont Health Technol Assess Ser,* 17**,** 1-75.

HUANG, C. J., CHEN, H. Y., LIN, W. Y. & CHOO, K. B. 2014. Differential expression of speckled POZ protein, SPOP: putative regulation by miR-145. *J Biosci,* 39**,** 401-13.

HUERTAS, P. & JACKSON, S. P. 2009. Human CtIP mediates cell cycle control of DNA end resection and double strand break repair. *J Biol Chem,* 284**,** 9558-65.

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.

HWANG, D. J., HE, Y., PONNUSAMY, S., MOHLER, M. L., THIYAGARAJAN, T., MCEWAN, I. J., NARAYANAN, R. & MILLER, D. D. 2019. New Generation of Selective Androgen Receptor Degraders: Our Initial Design, Synthesis, and Biological Evaluation of New Compounds with Enzalutamide-Resistant Prostate Cancer Activity. *J Med Chem,* 62**,** 491-511.

JENSTER, G., TRAPMAN, J. & BRINKMANN, A. O. 1993. Nuclear import of the human androgen receptor. *Biochem J,* 293 (Pt 3)**,** 761-8.

JIANG, H., XI, H., JUHAS, M. & ZHANG, Y. 2021. Biosensors for Point Mutation Detection. *Front Bioeng Biotechnol,* 9**,** 797831.

JOHNS, L. E. & HOULSTON, R. S. 2003. A systematic review and meta-analysis of familial prostate cancer risk. *BJU Int,* 91**,** 789-94.

JU, L.-G., ZHU, Y., LONG, Q.-Y., LI, X.-J., LIN, X., TANG, S.-B., YIN, L., XIAO, Y., WANG, X.-H., LI, L., ZHANG, L. & WU, M. 2019. SPOP suppresses prostate cancer through regulation of CYCLIN E1 stability. *Cell Death & Differentiation,* 26**,** 1156-1168.

KARANAM, K., KAFRI, R., LOEWER, A. & LAHAV, G. 2012. Quantitative live cell imaging reveals a gradual shift between DNA repair mechanisms and a maximal use of HR in mid S phase. *Mol Cell,* 47**,** 320-9.

KARATAS, O. F., YUCETURK, B., SUER, I., YILMAZ, M., CANSIZ, H., SOLAK, M., ITTMANN, M. & OZEN, M. 2016. Role of miR-145 in human laryngeal squamous cell carcinoma. *Head Neck,* 38**,** 260-6.

KARI, V., MANSOUR, W. Y., RAUL, S. K., BAUMGART, S. J., MUND, A., GRADE, M., SIRMA, H., SIMON, R., WILL, H., DOBBELSTEIN, M., DIKOMEY, E. & JOHNSEN, S. A. 2016. Loss of CHD1 causes DNA repair defects and enhances prostate cancer therapeutic responsiveness. *EMBO Rep,* 17**,** 1609-1623.

KARNES, R. J., BERGSTRALH, E. J., DAVICIONI, E., GHADESSI, M., BUERKI, C., MITRA, A. P., CRISAN, A., ERHO, N., VERGARA, I. A., LAM, L. L., CARLSON, R., THOMPSON, D. J., HADDAD, Z., ZIMMERMANN, B., SIEROCINSKI, T., TRICHE, T. J., KOLLMEYER, T., BALLMAN, K. V., BLACK, P. C., KLEE, G. G. & JENKINS, R. B. 2013. Validation of a genomic classifier that predicts metastasis following radical prostatectomy in an at risk patient population. *J Urol,* 190**,** 2047-53.

KASTAN, M. B. & BARTEK, J. 2004. Cell-cycle checkpoints and cancer. *Nature,* 432**,** 316-323.

KICIŃSKI, M., VANGRONSVELD, J. & NAWROT, T. S. 2011. An Epidemiological Reappraisal of the Familial Aggregation of Prostate Cancer: A Meta-Analysis. *PLOS ONE,* 6**,** e27130.

KIENER, M., CHEN, L., KREBS, M., GROSJEAN, J., KLIMA, I., KALOGIROU, C., RIEDMILLER, H., KNEITZ, B., THALMANN, G. N., SNAAR-JAGALSKA, E., SPAHN, M., KRUITHOF-DE JULIO, M. & ZONI, E. 2019. miR-221-5p regulates proliferation and migration in human prostate cancer cells and reduces tumor growth in vivo. *BMC Cancer,* 19**,** 627.

KIM, E. H. & BULLOCK, A. D. 2018. Surgical Management for Prostate Cancer. *Mo Med,* 115**,** 142-145.

KLEIN, E. A., HADDAD, Z., YOUSEFI, K., LAM, L. L., WANG, Q., CHOEURNG, V., PALMER-ARONSTEN, B., BUERKI, C., DAVICIONI, E., LI, J., KATTAN, M. W., STEPHENSON, A. J. & MAGI-GALLUZZI, C. 2016. Decipher Genomic Classifier Measured on Prostate Biopsy Predicts Metastasis Risk. *Urology,* 90**,** 148-52.

KLUTH, L. A., SHARIAT, S. F., KRATZIK, C., TAGAWA, S., SONPAVDE, G., RIEKEN, M., SCHERR, D. S. & PUMMER, K. 2014. The hypothalamic-pituitary-gonadal axis and prostate cancer: implications for androgen deprivation therapy. *World J Urol,* 32**,** 669-76.

KNEITZ, B., KREBS, M., KALOGIROU, C., SCHUBERT, M., JONIAU, S., VAN POPPEL, H., LERUT, E., KNEITZ, S., SCHOLZ, C. J., STRÖBEL, P., GESSLER, M., RIEDMILLER, H. & SPAHN, M. 2014. Survival in patients with high-risk prostate cancer is predicted by miR-221, which regulates proliferation, apoptosis, and invasion of prostate cancer cells by inhibiting IRF2 and SOCS3. *Cancer Res,* 74**,** 2591-603.

KOHAAR, I., PETROVICS, G. & SRIVASTAVA, S. 2019. A Rich Array of Prostate Cancer Molecular Biomarkers: Opportunities and Challenges. *Int J Mol Sci,* 20.

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., 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.

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,* 47**,** 5634-5647.

KOZOMARA, A. & GRIFFITHS-JONES, S. 2014. miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res,* 42**,** D68-73.

KREBS, M., SOLIMANDO, A. G., KALOGIROU, C., MARQUARDT, A., FRANK, T., SOKOLAKIS, I., HATZICHRISTODOULOU, G., KNEITZ, S., BARGOU, R., KÜBLER, H., SCHILLING, B., SPAHN, M. & KNEITZ, B. 2020. miR-221-3p Regulates VEGFR2 Expression in High-Risk Prostate Cancer and Represents an Escape Mechanism from Sunitinib In Vitro. *J Clin Med,* 9.

KROKAN, H. E. & BJØRÅS, M. 2013. Base excision repair. *Cold Spring Harb Perspect Biol,* 5**,** a012583.

KUNKEL, T. A. 2009. Evolving views of DNA replication (in)fidelity. *Cold Spring Harb Symp Quant Biol,* 74**,** 91-101.

LAN, L., NAKAJIMA, S., OOHATA, Y., TAKAO, M., OKANO, S., MASUTANI, M., WILSON, S. H. & YASUI, A. 2004. In situ analysis of repair processes for oxidative DNA damage in mammalian cells. *Proc Natl Acad Sci U S A,* 101**,** 13738-43.

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.

LANGLEY, E., KEMPPAINEN, J. A. & WILSON, E. M. 1998. Intermolecular NH2-/carboxyl-terminal interactions in androgen receptor dimerization revealed by mutations that cause androgen insensitivity. *J Biol Chem,* 273**,** 92-101.

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.

LARNE, O., HAGMAN, Z., LILJA, H., BJARTELL, A., EDSJÖ, A. & CEDER, Y. 2015. miR-145 suppress the androgen receptor in prostate cancer cells and correlates to prostate cancer prognosis. *Carcinogenesis,* 36**,** 858-66.

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.

LEIDINGER, P., HART, M., BACKES, C., RHEINHEIMER, S., KECK, B., WULLICH, B., KELLER, A. & MEESE, E. 2016. Differential blood-based diagnosis between benign prostatic hyperplasia and prostate cancer: miRNA as source for biomarkers independent of PSA level, Gleason score, or TNM status. *Tumour Biol,* 37**,** 10177-85.

LEIMBACHER, P. A., JONES, S. E., SHORROCKS, A. K., DE MARCO ZOMPIT, M., DAY, M., BLAAUWENDRAAD, J., BUNDSCHUH, D., BONHAM, S., FISCHER, R., FINK, D., KESSLER, B. M., OLIVER, A. W., PEARL, L. H., BLACKFORD, A. N. & STUCKI, M. 2019. MDC1 Interacts with TOPBP1 to Maintain Chromosomal Stability during Mitosis. *Mol Cell,* 74**,** 571-583.e8.

LI, H., GIGI, L. & ZHAO, D. 2023a. CHD1, a multifaceted epigenetic remodeler in prostate cancer. *Front Oncol,* 13**,** 1123362.

LI, Q., QIAN, W., ZHANG, Y., HU, L., CHEN, S. & XIA, Y. 2023b. A new wave of innovations within the DNA damage response. *Signal Transduction and Targeted Therapy,* 8**,** 338.

LIN, C., YANG, L., TANASA, B., HUTT, K., JU, B. G., OHGI, K., ZHANG, J., ROSE, D. W., FU, X. D., GLASS, C. K. & ROSENFELD, M. G. 2009. Nuclear receptor-induced chromosomal proximity and DNA breaks underlie specific translocations in cancer. *Cell,* 139**,** 1069-83.

LIN, H. K., WANG, L., HU, Y. C., ALTUWAIJRI, S. & CHANG, C. 2002. Phosphorylation-dependent ubiquitylation and degradation of androgen receptor by Akt require Mdm2 E3 ligase. *Embo j,* 21**,** 4037-48.

LIU, J., ZHANG, Y., XIE, H., ZHAO, L., ZHENG, L. & YE, H. 2019. Applications of Catalytic Hairpin Assembly Reaction in Biosensing. *Small,* 15**,** e1902989.

LOEB, S., VELLEKOOP, A., AHMED, H. U., CATTO, J., EMBERTON, M., NAM, R., ROSARIO, D. J., SCATTONI, V. & LOTAN, Y. 2013. Systematic review of complications of prostate biopsy. *Eur Urol,* 64**,** 876-92.

LONERGAN, P. E. & TINDALL, D. J. 2011. Androgen receptor signaling in prostate cancer development and progression. *J Carcinog,* 10**,** 20.

LONGO, D. L. 2010. New therapies for castration-resistant prostate cancer. *N Engl J Med,* 363**,** 479-81.

LORD, C. J. & ASHWORTH, A. 2017. PARP inhibitors: Synthetic lethality in the clinic. *Science,* 355**,** 1152-1158.

LOWDER, D., RIZWAN, K., MCCOLL, C., PAPARELLA, A., ITTMANN, M., MITSIADES, N. & KAOCHAR, S. 2022. Racial disparities in prostate cancer: A complex interplay between socioeconomic inequities and genomics. *Cancer Lett,* 531**,** 71-82.

LOZANO, R., CASTRO, E., ARAGÓN, I. M., CENDÓN, Y., CATTRINI, C., LÓPEZ-CASAS, P. P. & OLMOS, D. 2021. Genetic aberrations in DNA repair pathways: a cornerstone of precision oncology in prostate cancer. *Br J Cancer,* 124**,** 552-563.

LU, C. & LUO, J. 2013. Decoding the androgen receptor splice variants. *Transl Androl Urol,* 2**,** 178-186.

MA, J., CAI, M., MO, Y., FRIED, J. S., TAN, X., MA, Y., CHEN, J., HAN, S. & XU, B. 2021. The SPOP-ITCH Signaling Axis Protects Against Prostate Cancer Metastasis. *Front Oncol,* 11**,** 658230.

MAEKAWA, M. & HIGASHIYAMA, S. 2020. The Roles of SPOP in DNA Damage Response and DNA Replication. *Int J Mol Sci,* 21.

MAHAJAN, K., MALLA, P., LAWRENCE, H. R., CHEN, Z., KUMAR-SINHA, C., MALIK, R., SHUKLA, S., KIM, J., COPPOLA, D., LAWRENCE, N. J. & MAHAJAN, N. P. 2017. ACK1/TNK2

Regulates Histone H4 Tyr88-phosphorylation and AR Gene Expression in Castration-Resistant Prostate Cancer. *Cancer Cell,* 31**,** 790-803.e8.

MAHAJAN, N. P., LIU, Y., MAJUMDER, S., WARREN, M. R., PARKER, C. E., MOHLER, J. L., EARP, H. S. & WHANG, Y. E. 2007. Activated Cdc42-associated kinase Ack1 promotes prostate cancer progression via androgen receptor tyrosine phosphorylation. *Proc Natl Acad Sci U S A,* 104**,** 8438-43.

MAJUMDER, P. K. & KUMAR, V. L. 1997. Androgen receptor mRNA detection in the human foetal prostate. *Int Urol Nephrol,* 29**,** 633-5.

MALL, C., ROCKE, D. M., DURBIN-JOHNSON, B. & WEISS, R. H. 2013. Stability of miRNA in human urine supports its biomarker potential. *Biomark Med,* 7**,** 623-31.

MANGELSDORF, D. J., THUMMEL, C., BEATO, M., HERRLICH, P., SCHÜTZ, G., UMESONO, K., BLUMBERG, B., KASTNER, P., MARK, M., CHAMBON, P. & EVANS, R. M. 1995. The nuclear receptor superfamily: the second decade. *Cell,* 83**,** 835-9.

MANI, R. S., TOMLINS, S. A., CALLAHAN, K., GHOSH, A., NYATI, M. K., VARAMBALLY, S., PALANISAMY, N. & CHINNAIYAN, A. M. 2009. Induced chromosomal proximity and gene fusions in prostate cancer. *Science,* 326**,** 1230.

MARKOWSKI, M. C., KACHHAP, S., DE MARZO, A. M., SENA, L. A., LUO, J., DENMEADE, S. R. & ANTONARAKIS, E. S. 2022. Molecular and Clinical Characterization of Patients With Metastatic Castration Resistant Prostate Cancer Achieving Deep Responses to Bipolar Androgen Therapy. *Clin Genitourin Cancer,* 20**,** 97-101.

MASUTANI, C., SUGASAWA, K., YANAGISAWA, J., SONOYAMA, T., UI, M., ENOMOTO, T., TAKIO, K., TANAKA, K., VAN DER SPEK, P. J., BOOTSMA, D. & ET AL. 1994. Purification and cloning of a nucleotide excision repair complex involving the xeroderma pigmentosum group C protein and a human homologue of yeast RAD23. *Embo j,* 13**,** 1831-43.

MATEO, J., BOYSEN, G., BARBIERI, C. E., BRYANT, H. E., CASTRO, E., NELSON, P. S., OLMOS, D., PRITCHARD, C. C., RUBIN, M. A. & DE BONO, J. S. 2017. DNA Repair in Prostate Cancer: Biology and Clinical Implications. *Eur Urol,* 71**,** 417-425.

MATIAS, P. M., DONNER, P., COELHO, R., THOMAZ, M., PEIXOTO, C., MACEDO, S., OTTO, N., JOSCHKO, S., SCHOLZ, P., WEGG, A., BÄSLER, S., SCHÄFER, M., EGNER, U. & CARRONDO, M. A. 2000. Structural evidence for ligand specificity in the binding domain of the human androgen receptor. Implications for pathogenic gene mutations. *J Biol Chem,* 275**,** 26164-71.

MCEWAN, I. J. 2004. Molecular mechanisms of androgen receptor-mediated gene regulation: structure-function analysis of the AF-1 domain. *Endocr Relat Cancer,* 11**,** 281-93.

MCKIERNAN, J., DONOVAN, M. J., O'NEILL, V., BENTINK, S., NOERHOLM, M., BELZER, S., SKOG, J., KATTAN, M. W., PARTIN, A., ANDRIOLE, G., BROWN, G., WEI, J. T., THOMPSON, I. M., JR. & CARROLL, P. 2016. A Novel Urine Exosome Gene Expression Assay to Predict High-grade Prostate Cancer at Initial Biopsy. *JAMA Oncol,* 2**,** 882-9.

MCKINNON, P. J. & CALDECOTT, K. W. 2007. DNA strand break repair and human genetic disease. *Annu Rev Genomics Hum Genet,* 8**,** 37-55.

MERCATELLI, N., COPPOLA, V., BONCI, D., MIELE, F., COSTANTINI, A., GUADAGNOLI, M., BONANNO, E., MUTO, G., FRAJESE, G. V., DE MARIA, R., SPAGNOLI, L. G., FARACE, M. G. & CIAFRÈ, S. A. 2008. The inhibition of the highly expressed miR-221 and miR-222 impairs the growth of prostate carcinoma xenografts in mice. *PLoS One,* 3**,** e4029.

MERRIEL, S. W. D., POCOCK, L., GILBERT, E., CREAVIN, S., WALTER, F. M., SPENCER, A. & HAMILTON, W. 2022. Systematic review and meta-analysis of the diagnostic accuracy of prostate-specific antigen (PSA) for the detection of prostate cancer in symptomatic patients. *BMC Medicine,* 20**,** 54.

MESSNER, E. A., STEELE, T. M., TSAMOURI, M. M., HEJAZI, N., GAO, A. C., MUDRYJ, M. & GHOSH, P. M. 2020. The Androgen Receptor in Prostate Cancer: Effect of Structure, Ligands and Spliced Variants on Therapy. *Biomedicines,* 8.

MITRA, R., LIN, C. C., EISCHEN, C. M., BANDYOPADHYAY, S. & ZHAO, Z. 2015. Concordant dysregulation of miR-5p and miR-3p arms of the same precursor microRNA may be a mechanism in inducing cell proliferation and tumorigenesis: a lung cancer study. *Rna,* 21**,** 1055-65.

MORTUSEWICZ, O., ROTHBAUER, U., CARDOSO, M. C. & LEONHARDT, H. 2006. Differential recruitment of DNA Ligase I and III to DNA repair sites. *Nucleic Acids Res,* 34**,** 3523-32.

MOSTAGHEL, E. A., MARCK, B. T., PLYMATE, S. R., VESSELLA, R. L., BALK, S., MATSUMOTO, A. M., NELSON, P. S. & MONTGOMERY, R. B. 2011. Resistance to CYP17A1 inhibition with abiraterone in castration-resistant prostate cancer: induction of steroidogenesis and androgen receptor splice variants. *Clin Cancer Res,* 17**,** 5913-25.

MOTTET, N., BELLMUNT, J., BOLLA, M., BRIERS, E., CUMBERBATCH, M. G., DE SANTIS, M., FOSSATI, N., GROSS, T., HENRY, A. M., JONIAU, S., LAM, T. B., MASON, M. D., MATVEEV, V. B., MOLDOVAN, P. C., VAN DEN BERGH, R. C. N., VAN DEN BROECK, T., VAN DER POEL, H. G., VAN DER KWAST, T. H., ROUVIÈRE, O., SCHOOTS, I. G., WIEGEL, T. & CORNFORD, P. 2017. EAU-ESTRO-SIOG Guidelines on Prostate Cancer. Part 1: Screening, Diagnosis, and Local Treatment with Curative Intent. *Eur Urol,* 71**,** 618-629.

MUCCI, L. A., HJELMBORG, J. B., HARRIS, J. R., CZENE, K., HAVELICK, D. J., SCHEIKE, T., GRAFF, R. E., HOLST, K., MÖLLER, S., UNGER, R. H., MCINTOSH, C., NUTTALL, E., BRANDT, I., PENNEY, K. L., HARTMAN, M., KRAFT, P., PARMIGIANI, G., CHRISTENSEN, K., KOSKENVUO, M., HOLM, N. V., HEIKKILÄ, K., PUKKALA, E., SKYTTHE, A., ADAMI, H. O. & KAPRIO, J. 2016. Familial Risk and Heritability of Cancer Among Twins in Nordic Countries. *Jama,* 315**,** 68-76.

MUKHOPADHYAY, C., YANG, C., XU, L., LIU, D., WANG, Y., HUANG, D., DEONARINE, L. D., CYRTA, J., DAVICIONI, E., SBONER, A., ROBINSON, B. D., CHINNAIYAN, A. M., RUBIN, M. A., BARBIERI, C. E. & ZHOU, P. 2021. G3BP1 inhibits Cul3(SPOP) to amplify AR signaling and promote prostate cancer. *Nat Commun,* 12**,** 6662.

NADIMINTY, N., TUMMALA, R., LIU, C., YANG, J., LOU, W., EVANS, C. P. & GAO, A. C. 2013. NF-κB2/p52 induces resistance to enzalutamide in prostate cancer: role of androgen receptor and its variants. *Mol Cancer Ther,* 12**,** 1629-37.

NAGAI, Y., KOJIMA, T., MURO, Y., HACHIYA, T., NISHIZAWA, Y., WAKABAYASHI, T. & HAGIWARA, M. 1997. Identification of a novel nuclear speckle-type protein, SPOP. *FEBS Lett,* 418**,** 23-6.

NAKAMURA, K., SAREDI, G., BECKER, J. R., FOSTER, B. M., NGUYEN, N. V., BEYER, T. E., CESA, L. C., FAULL, P. A., LUKAUSKAS, S., FRIMURER, T., CHAPMAN, J. R., BARTKE, T. & GROTH, A. 2019. H4K20me0 recognition by BRCA1–BARD1 directs homologous recombination to sister chromatids. *Nature Cell Biology,* 21**,** 311-318.

NI, L., LLEWELLYN, R., KESLER, C. T., KELLEY, J. B., SPENCER, A., SNOW, C. J., SHANK, L. & PASCHAL, B. M. 2013. Androgen induces a switch from cytoplasmic retention to nuclear import of the androgen receptor. *Mol Cell Biol,* 33**,** 4766-78.

NIKHIL, K., HAYMOUR, H. S., KAMRA, M. & SHAH, K. 2021. Phosphorylation-dependent regulation of SPOP by LIMK2 promotes castration-resistant prostate cancer. *Br J Cancer,* 124**,** 995-1008.

NIKHIL, K., KAMRA, M., RAZA, A., HAYMOUR, H. S. & SHAH, K. 2020. Molecular Interplay between AURKA and SPOP Dictates CRPC Pathogenesis via Androgen Receptor. *Cancers (Basel),* 12.

NISHI, R., OKUDA, Y., WATANABE, E., MORI, T., IWAI, S., MASUTANI, C., SUGASAWA, K. & HANAOKA, F. 2005. Centrin 2 stimulates nucleotide excision repair by interacting with xeroderma pigmentosum group C protein. *Mol Cell Biol,* 25**,** 5664-74.

NYBERG, T., GOVINDASAMI, K., LESLIE, G., DADAEV, T., BANCROFT, E., NI RAGHALLAIGH, H., BROOK, M. N., HUSSAIN, N., KEATING, D., LEE, A., MCMAHON, R., MORGAN, A., MULLEN, A., OSBORNE, A., RAGEEVAKUMAR, R., KOTE-JARAI, Z., EELES, R. & ANTONIOU, A. C. 2019. Homeobox B13 G84E Mutation and Prostate Cancer Risk. *Eur Urol,* 75**,** 834-845.

NYBERG, T., TISCHKOWITZ, M. & ANTONIOU, A. C. 2022. BRCA1 and BRCA2 pathogenic variants and prostate cancer risk: systematic review and meta-analysis. *Br J Cancer,* 126**,** 1067-1081.

NYQUIST, M. D., ANG, L. S., CORELLA, A., COLEMAN, I. M., MEERS, M. P., CHRISTIANI, A. J., PIERCE, C., JANSSENS, D. H., MEADE, H. E., BOSE, A., BRADY, L., HOWARD, T., DE SARKAR, N., FRANK, S. B., DUMPIT, R. F., DALTON, J. T., COREY, E., PLYMATE, S. R., HAFFNER, M. C., MOSTAGHEL, E. A. & NELSON, P. S. 2021. Selective androgen receptor modulators activate the canonical prostate cancer androgen receptor program and repress cancer growth. *J Clin Invest,* 131.

O'DONNELL, A., JUDSON, I., DOWSETT, M., RAYNAUD, F., DEARNALEY, D., MASON, M., HARLAND, S., ROBBINS, A., HALBERT, G., NUTLEY, B. & JARMAN, M. 2004. Hormonal impact of the 17alpha-hydroxylase/C(17,20)-lyase inhibitor abiraterone acetate (CB7630) in patients with prostate cancer. *Br J Cancer,* 90**,** 2317-25.

OZEN, M., KARATAS, O. F., GULLUOGLU, S., BAYRAK, O. F., SEVLI, S., GUZEL, E., EKICI, I. D., CASKURLU, T., SOLAK, M., CREIGHTON, C. J. & ITTMANN, M. 2015. Overexpression of miR-145-5p inhibits proliferation of prostate cancer cells and reduces SOX2 expression. *Cancer Invest,* 33**,** 251-8.

PANIER, S. & BOULTON, S. J. 2014. Double-strand break repair: 53BP1 comes into focus. *Nat Rev Mol Cell Biol,* 15**,** 7-18.

PCUK. 2021. Available: https://prostatecanceruk.org/prostate-information/about-prostate-cancer [Accessed 2021].

POLKINGHORN, W. R., PARKER, J. S., LEE, M. X., KASS, E. M., SPRATT, D. E., IAQUINTA, 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. 2013. Androgen receptor signaling regulates DNA repair in prostate cancers. *Cancer Discov,* 3**,** 1245-53.

PONNUSAMY, S., HE, Y., HWANG, D. J., THIYAGARAJAN, T., HOUTMAN, R., BOCHAROVA, V., SUMPTER, B. G., FERNANDEZ, E., JOHNSON, D., DU, Z., PFEFFER, L. M., GETZENBERG, R. H., MCEWAN, I. J., MILLER, D. D. & NARAYANAN, R. 2019. Orally Bioavailable Androgen Receptor Degrader, Potential Next-Generation Therapeutic for Enzalutamide-Resistant Prostate Cancer. *Clin Cancer Res,* 25**,** 6764-6780.

PRATT, W. B. & TOFT, D. O. 1997. Steroid receptor interactions with heat shock protein and immunophilin chaperones. *Endocr Rev,* 18**,** 306-60.

PRITCHARD, C. C., MATEO, J., WALSH, M. F., DE SARKAR, N., ABIDA, W., BELTRAN, H., GAROFALO, A., GULATI, R., CARREIRA, S., EELES, R., ELEMENTO, O., RUBIN, M. A., ROBINSON, D., LONIGRO, R., HUSSAIN, M., CHINNAIYAN, A., VINSON, J., FILIPENKO, J., GARRAWAY, L., TAPLIN, M. E., ALDUBAYAN, S., HAN, G. C., BEIGHTOL, M., MORRISSEY, C., NGHIEM, B., CHENG, H. H., MONTGOMERY, B., WALSH, T., CASADEI, S., BERGER, M., ZHANG, L., ZEHIR, A., VIJAI, J., SCHER, H. I., SAWYERS, C., SCHULTZ, N., KANTOFF, P. W., SOLIT, D., ROBSON, M., VAN ALLEN, E. M., OFFIT, K., DE BONO, J. & NELSON, P. S. 2016. Inherited DNA-Repair Gene Mutations in Men with Metastatic Prostate Cancer. *N Engl J Med,* 375**,** 443-53.

QIU, X., BOUFAIED, N., HALLAL, T., FEIT, A., DE POLO, A., LUOMA, A. M., ALAHMADI, W., LAROCQUE, J., ZADRA, G., XIE, Y., GU, S., TANG, Q., ZHANG, Y., SYAMALA, S., SEO, J.-H., BELL, C., O'CONNOR, E., LIU, Y., SCHAEFFER, E. M., JEFFREY KARNES, R., WEINMANN, S., DAVICIONI, E., MORRISSEY, C., CEJAS, P., ELLIS, L., LODA, M., WUCHERPFENNIG, K. W., POMERANTZ, M. M., SPRATT, D. E., COREY, E., FREEDMAN, M. L., SHIRLEY LIU, X., BROWN, M., LONG, H. W. & LABBÉ, D. P. 2022. MYC drives aggressive prostate cancer by disrupting transcriptional pause release at androgen receptor targets. *Nature Communications,* 13**,** 2559.

RAMIREZ-GARRASTACHO, M., BAJO-SANTOS, C., LINE, A., MARTENS-UZUNOVA, E. S., DE LA FUENTE, J. M., MOROS, M., SOEKMADJI, C., TASKEN, K. A. & LLORENTE, A. 2022. Extracellular vesicles as a source of prostate cancer biomarkers in liquid biopsies: a decade of research. *British Journal of Cancer,* 126**,** 331-350.

RAWLA, P. 2019. Epidemiology of Prostate Cancer. *World J Oncol,* 10**,** 63-89.

RIABINSKA, A., DAHEIM, M., HERTER-SPRIE, G. S., WINKLER, J., FRITZ, C., HALLEK, M., THOMAS, R. K., KREUZER, K.-A., FRENZEL, L. P., MONFARED, P., MARTINS-BOUCAS, J., CHEN, S. & REINHARDT, H. C. 2013. Therapeutic Targeting of a Robust Non-Oncogene Addiction to PRKDC in ATM-Defective Tumors. *Science Translational Medicine,* 5**,** 189ra78-189ra78.

RICE, K. R., CHEN, Y., ALI, A., WHITMAN, E. J., BLASE, A., IBRAHIM, M., ELSAMANOUDI, S., BRASSELL, S., FURUSATO, B., STINGLE, N., SESTERHENN, I. A., PETROVICS, G., MIICK, S., RITTENHOUSE, H., GROSKOPF, J., MCLEOD, D. G. & SRIVASTAVA, S. 2010. Evaluation of the ETS-related gene mRNA in urine for the detection of prostate cancer. *Clin Cancer Res,* 16**,** 1572-6.

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.

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.

RUBIN, M. A., GIRELLI, G. & DEMICHELIS, F. 2016. Genomic Correlates to the Newly Proposed Grading Prognostic Groups for Prostate Cancer. *Eur Urol,* 69**,** 557-560.

SALIMINEJAD, K., KHORRAM KHORSHID, H. R., SOLEYMANI FARD, S. & GHAFFARI, S. H. 2019. An overview of microRNAs: Biology, functions, therapeutics, and analysis methods. *J Cell Physiol,* 234**,** 5451-5465.

SANTOS, J. I., TEIXEIRA, A. L., DIAS, F., MAURÍCIO, J., LOBO, F., MORAIS, A. & MEDEIROS, R. 2014. Influence of peripheral whole-blood microRNA-7 and microRNA-221 high expression levels on the acquisition of castration-resistant prostate cancer: evidences from in vitro and in vivo studies. *Tumour Biol,* 35**,** 7105-13.

SCHÄRER, O. D. 2013. Nucleotide excision repair in eukaryotes. *Cold Spring Harb Perspect Biol,* 5**,** a012609.

SCHATTEN, H. 2018. Brief Overview of Prostate Cancer Statistics, Grading, Diagnosis and Treatment Strategies. *Adv Exp Med Biol,* 1095**,** 1-14.

SCHEINBERG, T., LIN, H. M., FITZPATRICK, M., AZAD, A. A., BONNITCHA, P., DAVIES, A., HELLER, G., HUYNH, K., MAK, B., MAHON, K., SULLIVAN, D., MEIKLE, P. J. & HORVATH, L. G. 2024. PCPro: a clinically accessible, circulating lipid biomarker signature for poor-prognosis metastatic prostate cancer. *Prostate Cancer Prostatic Dis,* 27**,** 136-143.

SCHER, H. I., BEER, T. M., HIGANO, C. S., ANAND, A., TAPLIN, M. E., EFSTATHIOU, E., RATHKOPF, D., SHELKEY, J., YU, E. Y., ALUMKAL, J., HUNG, D., HIRMAND, M., SEELY, L., MORRIS, M. J., DANILA, D. C., HUMM, J., LARSON, S., FLEISHER, M. & SAWYERS, C. L. 2010. Antitumour activity of MDV3100 in castration-resistant prostate cancer: a phase 1-2 study. *Lancet,* 375**,** 1437-46.

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., FLÉCHON, A., MAINWARING, P., FLEMING, M., HAINSWORTH, J. D., HIRMAND, M., SELBY, B., SEELY, L. & DE BONO, J. S. 2012. Increased survival with enzalutamide in prostate cancer after chemotherapy. *N Engl J Med,* 367**,** 1187-97.

SCHWEIZER, M. T., GULATI, R., YEZEFSKI, T., CHENG, H. H., MOSTAGHEL, E., HAFFNER, M. C., PATEL, R. A., DE SARKAR, N., HA, G., DUMPIT, R., WOO, B., LIN, A., PANLASIGUI, P.,

MCDONALD, N., LAI, M., NEGA, K., HAMMOND, J., GRIVAS, P., HSIEH, A., MONTGOMERY, B., NELSON, P. S. & YU, E. Y. 2023. Bipolar androgen therapy plus olaparib in men with metastatic castration-resistant prostate cancer. *Prostate Cancer Prostatic Dis,* 26**,** 194-200.

SEN, T., RODRIGUEZ, B. L., CHEN, L., CORTE, C. M. D., MORIKAWA, N., FUJIMOTO, J., CRISTEA, S., NGUYEN, T., DIAO, L., LI, L., FAN, Y., YANG, Y., WANG, J., GLISSON, B. S., WISTUBA, I. I., SAGE, J., HEYMACH, J. V., GIBBONS, D. L. & BYERS, L. A. 2019. Targeting DNA Damage Response Promotes Antitumor Immunity through STING-Mediated T-cell Activation in Small Cell Lung Cancer. *Cancer Discovery,* 9**,** 646-661.

SEVEN, M., KARATAS, O. F., DUZ, M. B. & OZEN, M. 2014. The role of miRNAs in cancer: from pathogenesis to therapeutic implications. *Future Oncol,* 10**,** 1027-48.

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.

SHAHEEN, F. S., ZNOJEK, P., FISHER, A., WEBSTER, M., PLUMMER, R., GAUGHAN, L., SMITH, G. C., LEUNG, H. Y., CURTIN, N. J. & ROBSON, C. N. 2011. Targeting the DNA double strand break repair machinery in prostate cancer. *PLoS One,* 6**,** e20311.

SHANG, Y., MYERS, M. & BROWN, M. 2002. Formation of the androgen receptor transcription complex. *Mol Cell,* 9**,** 601-10.

SHAO, N., MA, G., ZHANG, J. & ZHU, W. 2018. miR-221-5p enhances cell proliferation and metastasis through post-transcriptional regulation of SOCS1 in human prostate cancer. *BMC Urol,* 18**,** 14.

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.

SHIPITSIN, M., SMALL, C., CHOUDHURY, S., GILADI, E., FRIEDLANDER, S., NARDONE, J., HUSSAIN, S., HURLEY, A. D., ERNST, C., HUANG, Y. E., CHANG, H., NIFONG, T. P., RIMM, D. L., DUNYAK, J., LODA, M., BERMAN, D. M. & BLUME-JENSEN, P. 2014. Identification of proteomic biomarkers predicting prostate cancer aggressiveness and lethality despite biopsy-sampling error. *Br J Cancer,* 111**,** 1201-12.

SINGH, M., JHA, R., MELAMED, J., SHAPIRO, E., HAYWARD, S. W. & LEE, P. 2014. Stromal androgen receptor in prostate development and cancer. *Am J Pathol,* 184**,** 2598-607.

SMITH, D. F. & TOFT, D. O. 2008. Minireview: the intersection of steroid receptors with molecular chaperones: observations and questions. *Mol Endocrinol,* 22**,** 2229-40.

SONG, C., CHEN, H., WANG, T., ZHANG, W., RU, G. & LANG, J. 2015. Expression profile analysis of microRNAs in prostate cancer by next-generation sequencing. *Prostate,* 75**,** 500-16.

SONG, J., OUYANG, Y., CHE, J., LI, X., ZHAO, Y., YANG, K., ZHAO, X., CHEN, Y., FAN, C. & YUAN, W. 2017. Potential Value of miR-221/222 as Diagnostic, Prognostic, and Therapeutic Biomarkers for Diseases. *Front Immunol,* 8**,** 56.

SONG, Y., XU, Y., PAN, C., YAN, L., WANG, Z. W. & ZHU, X. 2020. The emerging role of SPOP protein in tumorigenesis and cancer therapy. *Mol Cancer,* 19**,** 2.

SPAHN, M., KNEITZ, S., SCHOLZ, C. J., STENGER, N., RÜDIGER, T., STRÖBEL, P., RIEDMILLER, H. & KNEITZ, B. 2010. Expression of microRNA-221 is progressively reduced in aggressive prostate cancer and metastasis and predicts clinical recurrence. *Int J Cancer,* 127**,** 394-403.

SPIVAK, G. 2015. Nucleotide excision repair in humans. *DNA Repair (Amst),* 36**,** 13-18.

SRINIVAS, N., PARKIN, J., SEELIG, G., WINFREE, E. & SOLOVEICHIK, D. 2017. Enzyme-free nucleic acid dynamical systems. *Science,* 358.

SRINIVAS-SHANKAR, U. & WU, F. C. 2006. Drug insight: testosterone preparations. *Nat Clin Pract Urol,* 3**,** 653-65.

SRIVASTAVA, S. K., BHARDWAJ, A., LEAVESLEY, S. J., GRIZZLE, W. E., SINGH, S. & SINGH, A. P. 2013. MicroRNAs as potential clinical biomarkers: emerging approaches for their detection. *Biotech Histochem,* 88**,** 373-87.

STOVER, E. H., OH, C., KESKULA, P., CHOUDHURY, A. D., TSENG, Y. Y., ADALSTEINSSON, V. A., LOHR, J. G., THORNER, A. R., DUCAR, M., KRYUKOV, G. V., HA, G., ROSENBERG, M., FREEMAN, S. S., ZHANG, Z., WU, X., VAN ALLEN, E. M., TAKEDA, D. Y., LODA, M., WU, C. L., TAPLIN, M. E., GARRAWAY, L. A., BOEHM, J. S. & HUANG, F. W. 2022. Implementation of a prostate cancer-specific targeted sequencing panel for credentialing of patient-derived cell lines and genomic characterization of patient samples. *Prostate,* 82**,** 584-597.

STRACKER, T. H. & PETRINI, J. H. 2011. The MRE11 complex: starting from the ends. *Nat Rev Mol Cell Biol,* 12**,** 90-103.

SUN, T., YANG, M., CHEN, S., BALK, S., POMERANTZ, M., HSIEH, C. L., BROWN, M., LEE, G. M. & KANTOFF, P. W. 2012. The altered expression of MiR-221/-222 and MiR-23b/-27b is associated with the development of human castration resistant prostate cancer. *Prostate,* 72**,** 1093-103.

SUN, Y., JIANG, X., CHEN, S., FERNANDES, N. & PRICE, B. D. 2005. A role for the Tip60 histone acetyltransferase in the acetylation and activation of ATM. *Proc Natl Acad Sci U S A,* 102**,** 13182-7.

SVILAR, D., GOELLNER, E. M., ALMEIDA, K. H. & SOBOL, R. W. 2011. Base excision repair and lesion-dependent subpathways for repair of oxidative DNA damage. *Antioxid Redox Signal,* 14**,** 2491-507.

SWAMI, U., GRAF, R. P., NUSSENZVEIG, R. H., FISHER, V., TUKACHINSKY, H., SCHROCK, A. B., LI, G., ROSS, J. S., SAYEGH, N., TRIPATHI, N., MATHEW THOMAS, V., OXNARD, G. R., ANTONARAKIS, E. S. & AGARWAL, N. 2022. SPOP Mutations as a Predictive Biomarker for Androgen Receptor Axis–Targeted Therapy in De Novo Metastatic Castration-Sensitive Prostate Cancer. *Clinical Cancer Research,* 28**,** 4917-4925.

TAN, M. H., LI, J., XU, H. E., MELCHER, K. & YONG, E. L. 2015. Androgen receptor: structure, role in prostate cancer and drug discovery. *Acta Pharmacol Sin,* 36**,** 3-23.

TANG, W., ZHONG, W., TAN, Y., WANG, G. A., LI, F. & LIU, Y. 2020. DNA Strand Displacement Reaction: A Powerful Tool for Discriminating Single Nucleotide Variants. *Top Curr Chem (Cham),* 378**,** 10.

TANG, Z., PILIÉ, P. G., GENG, C., MANYAM, G. C., YANG, G., PARK, S., WANG, D., PENG, S., WU, C., PENG, G., YAP, T. A., CORN, P. G., BROOM, B. M. & THOMPSON, T. C. 2021. ATR Inhibition Induces CDK1–SPOP Signaling and Enhances Anti–PD-L1 Cytotoxicity in Prostate Cancer. *Clinical Cancer Research,* 27**,** 4898-4909.

TCGA 2015. The Molecular Taxonomy of Primary Prostate Cancer. *Cell,* 163**,** 1011-25.

THADANI-MULERO, M., NANUS, D. M. & GIANNAKAKOU, P. 2012. Androgen receptor on the move: boarding the microtubule expressway to the nucleus. *Cancer Res,* 72**,** 4611-5.

THADANI-MULERO, M., PORTELLA, L., SUN, S., SUNG, M., MATOV, A., VESSELLA, R. L., COREY, E., NANUS, D. M., PLYMATE, S. R. & GIANNAKAKOU, P. 2014. Androgen receptor splice variants determine taxane sensitivity in prostate cancer. *Cancer Res,* 74**,** 2270-2282.

THEURILLAT, J. P., UDESHI, N. D., ERRINGTON, W. J., SVINKINA, T., BACA, S. C., POP, M., WILD, P. J., BLATTNER, M., GRONER, A. C., RUBIN, M. A., MOCH, H., PRIVE, G. G., CARR, S. A. & GARRAWAY, L. A. 2014. Prostate cancer. Ubiquitylome analysis identifies dysregulation of effector substrates in SPOP-mutant prostate cancer. *Science,* 346**,** 85-89.

THOMPSON, D., EASTON, D. F. & CONSORTIUM, T. B. C. L. 2002. Cancer Incidence in BRCA1 Mutation Carriers. *JNCI: Journal of the National Cancer Institute,* 94**,** 1358-1365.

THOMPSON, I. M., PAULER, D. K., GOODMAN, P. J., TANGEN, C. M., LUCIA, M. S., PARNES, H. L., MINASIAN, L. M., FORD, L. G., LIPPMAN, S. M., CRAWFORD, E. D., CROWLEY, J. J. & COLTMAN, C. A., JR. 2004. Prevalence of prostate cancer among men with a prostate-specific antigen level < or =4.0 ng per milliliter. *N Engl J Med,* 350**,** 2239-46.

TOHIDAST, M., MEMARI, N., AMINI, M., HOSSEINI, S. S., JEBELLI, A., DOUSTVANDI, M. A., BARADARAN, B. & MOKHTARZADEH, A. 2023. MiR-145 inhibits cell migration and increases paclitaxel chemosensitivity in prostate cancer cells. *Iran J Basic Med Sci,* 26**,** 1350-1359.

TOLEDO, L. I., MURGA, M., ZUR, R., SORIA, R., RODRIGUEZ, A., MARTINEZ, S., OYARZABAL, J., PASTOR, J., BISCHOFF, J. R. & FERNANDEZ-CAPETILLO, O. 2011. A cell-based screen identifies ATR inhibitors with synthetic lethal properties for cancer-associated mutations. *Nature Structural & Molecular Biology,* 18**,** 721-727.

TSELIS, N., HOSKIN, P., BALTAS, D., STRNAD, V., ZAMBOGLOU, N., RÖDEL, C. & CHATZIKONSTANTINOU, G. 2017. High Dose Rate Brachytherapy as Monotherapy for Localised Prostate Cancer: Review of the Current Status. *Clin Oncol (R Coll Radiol),* 29**,** 401-411.

TU, X., KAHILA, M. M., ZHOU, Q., YU, J., KALARI, K. R., WANG, L., HARMSEN, W. S., YUAN, J., BOUGHEY, J. C., GOETZ, M. P., SARKARIA, J. N., LOU, Z. & MUTTER, R. W. 2018. ATR Inhibition Is a Promising Radiosensitizing Strategy for Triple-Negative Breast Cancer. *Molecular Cancer Therapeutics,* 17**,** 2462-2472.

VALLÉE, A., LE LOUPP, A. G. & DENIS, M. G. 2014. Efficiency of the Therascreen® RGQ PCR kit for the detection of EGFR mutations in non-small cell lung carcinomas. *Clin Chim Acta,* 429**,** 8-11.

VAN BUSSEL, M. T. J., AWADA, A., DE JONGE, M. J. A., MAU-SØRENSEN, M., NIELSEN, D., SCHÖFFSKI, 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 DEN EEDEN, S. K., LU, R., ZHANG, N., QUESENBERRY, C. P., JR., SHAN, J., HAN, J. S., TSIATIS, A. C., LEIMPETER, A. D., LAWRENCE, H. J., FEBBO, P. G. & PRESTI, J. C. 2018. A Biopsy-based 17-gene Genomic Prostate Score as a Predictor of Metastases and Prostate

Cancer Death in Surgically Treated Men with Clinically Localized Disease. *Eur Urol,* 73**,** 129-138.

VAN NESTE, L., HENDRIKS, R. J., DIJKSTRA, S., TROOSKENS, G., CORNEL, E. B., JANNINK, S. A., DE JONG, H., HESSELS, D., SMIT, F. P., MELCHERS, W. J., LEYTEN, G. H., DE REIJKE, T. M., VERGUNST, H., KIL, P., KNIPSCHEER, B. C., HULSBERGEN-VAN DE KAA, C. A., MULDERS, P. F., VAN OORT, I. M., VAN CRIEKINGE, W. & SCHALKEN, J. A. 2016a. Detection of High-grade Prostate Cancer Using a Urinary Molecular Biomarker-Based Risk Score. *Eur Urol,* 70**,** 740-748.

VAN NESTE, L., PARTIN, A. W., STEWART, G. D., EPSTEIN, J. I., HARRISON, D. J. & VAN CRIEKINGE, W. 2016b. Risk score predicts high-grade prostate cancer in DNA-methylation positive, histopathologically negative biopsies. *Prostate,* 76**,** 1078-87.

VICKERS, A. J., GUPTA, A., SAVAGE, C. J., PETTERSSON, K., DAHLIN, A., BJARTELL, A., MANJER, J., SCARDINO, P. T., ULMERT, D. & LILJA, H. 2011. A panel of kallikrein marker predicts prostate cancer in a large, population-based cohort followed for 15 years without screening. *Cancer Epidemiol Biomarkers Prev,* 20**,** 255-61.

VOIGT, J. D., ZAPPALA, S. M., VAUGHAN, E. D. & WEIN, A. J. 2014. The Kallikrein Panel for prostate cancer screening: its economic impact. *Prostate,* 74**,** 250-9.

VOLINIA, S., CALIN, G. A., LIU, C. G., AMBS, S., CIMMINO, A., PETROCCA, F., VISONE, R., IORIO, M., ROLDO, C., FERRACIN, M., PRUEITT, R. L., YANAIHARA, N., LANZA, G., SCARPA, A., VECCHIONE, A., NEGRINI, M., HARRIS, C. C. & CROCE, C. M. 2006. A microRNA expression signature of human solid tumors defines cancer gene targets. *Proc Natl Acad Sci U S A,* 103**,** 2257-61.

WALLIS, C. J. & NAM, R. K. 2015. Prostate Cancer Genetics: A Review. *Ejifcc,* 26**,** 79-91.

WANG, F., MAO, A., TANG, J., ZHANG, Q., YAN, J., WANG, Y., DI, C., GAN, L., SUN, C. & ZHANG, H. 2019. microRNA-16-5p enhances radiosensitivity through modulating Cyclin D1/E1-pRb-E2F1 pathway in prostate cancer cells. *J Cell Physiol,* 234**,** 13182-13190.

WANG, F., WANG, W., LU, L., XIE, Y., YAN, J., CHEN, Y., DI, C., GAN, L., SI, J., ZHANG, H. & MAO, A. 2020. MicroRNA-16-5p regulates cell survival, cell cycle and apoptosis by targeting AKT3 in prostate cancer cells. *Oncol Rep,* 44**,** 1282-1292.

WANG, G., ZHAO, D., SPRING, D. J. & DEPINHO, R. A. 2018. Genetics and biology of prostate cancer. *Genes Dev,* 32**,** 1105-1140.

WANG, W., WANG, M., WANG, L., ADAMS, T. S., TIAN, Y. & XU, J. 2014. Diagnostic ability of %p2PSA and prostate health index for aggressive prostate cancer: a meta-analysis. *Sci Rep,* 4**,** 5012.

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. 2011. Combined androgen deprivation therapy and radiation therapy for locally advanced prostate cancer: a randomised, phase 3 trial. *Lancet,* 378**,** 2104-11.

WEI, C., LIU, Y., LIU, X., CHENG, J., FU, J., XIAO, X., MOSES, R. E., LI, X. & FU, J. 2022. The speckle-type POZ protein (SPOP) inhibits breast cancer malignancy by destabilizing TWIST1. *Cell Death Discovery,* 8**,** 389.

WEN, S., NIU, Y. & HUANG, H. 2020. Posttranslational regulation of androgen dependent and independent androgen receptor activities in prostate cancer. *Asian J Urol,* 7**,** 203-218.

WIDMARK, A., KLEPP, O., SOLBERG, A., DAMBER, J. E., ANGELSEN, A., FRANSSON, P., LUND, J. A., TASDEMIR, I., HOYER, M., WIKLUND, F. & FOSSÅ, S. D. 2009. Endocrine treatment, with or without radiotherapy, in locally advanced prostate cancer (SPCG-7/SFUO-3): an open randomised phase III trial. *Lancet,* 373**,** 301-8.

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.

WU, F., DAI, X., GAN, W., WAN, L., LI, M., MITSIADES, N., WEI, W., DING, Q. & ZHANG, J. 2017. Prostate cancer-associated mutation in SPOP impairs its ability to target Cdc20 for poly-ubiquitination and degradation. *Cancer Lett,* 385**,** 207-214.

WU, M., ZHANG, R., ZHANG, Z., ZHANG, N., LI, C., XIE, Y., XIA, H., HUANG, F., ZHANG, R., LIU, M., LI, X., CEN, S. & ZHOU, J. 2023. Selective androgen receptor degrader (SARD) to overcome antiandrogen resistance in castration-resistant prostate cancer. *Elife,* 12.

WU, Z., LIU, G. Q., YANG, X. L. & JIANG, J. H. 2015. Electrostatic nucleic acid nanoassembly enables hybridization chain reaction in living cells for ultrasensitive mRNA imaging. *J Am Chem Soc,* 137**,** 6829-36.

WÜSTMANN, N., SEIFERT, K. E., HUMBERG, V., VIELER, J., GRUNDMANN, N., STEINESTEL, J., TIEDJE, D., DUENSING, S., KRABBE, L. M., BÖGEMANN, M., SCHRADER, A. J., BERNEMANN, C. & SCHLACK, K. 2023. Co-expression and clinical utility of AR-FL and AR splice variants AR-V3, AR-V7 and AR-V9 in prostate cancer. *Biomark Res,* 11**,** 37.

XIAO, M., FRIED, J. S., MA, J., SU, Y., BOOHAKER, R. J., ZENG, Q., MO, Y., MENG, F., XIANG, R. & XU, B. 2021. A disease-relevant mutation of SPOP highlights functional significance of ATM-mediated DNA damage response. *Signal Transduct Target Ther,* 6**,** 17.

XU, K., SHIMELIS, H., LINN, D. E., JIANG, R., YANG, X., SUN, F., GUO, Z., CHEN, H., LI, W., CHEN, H., KONG, X., MELAMED, J., FANG, S., XIAO, Z., VEENSTRA, T. D. & QIU, Y. 2009. Regulation of androgen receptor transcriptional activity and specificity by RNF6-induced ubiquitination. *Cancer Cell,* 15**,** 270-82.

XUE, Z., YOU, M., PENG, P., TONG, H., HE, W., LI, A., MAO, P., XU, T., XU, F. & YAO, C. 2021. Taqman-MGB nanoPCR for Highly Specific Detection of Single-Base Mutations. *Int J Nanomedicine,* 16**,** 3695-3705.

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, Y., MENG, Y., ZHANG, H., SHEN, X., LI, R., YU, L., LIU, B. & WANG, L. 2018. Detection of EGFR and BRAF mutations by competitive allele-specific TaqMan polymerase chain reaction in lung adenocarcinoma. *Oncol Lett,* 15**,** 3295-3304.

YE, D., SHEN, Z. & ZHOU, S. 2019. Function of microRNA-145 and mechanisms underlying its role in malignant tumor diagnosis and treatment. *Cancer Manag Res,* 11**,** 969-979.

YURKE, B., TURBERFIELD, A. J., MILLS, A. P., JR., SIMMEL, F. C. & NEUMANN, J. L. 2000. A DNA-fuelled molecular machine made of DNA. *Nature,* 406**,** 605-8.

ZENG, H., HUANG, Y., LIU, Q., LIU, H., LONG, T., ZHU, C. & WU, X. 2021. MiR-145 suppresses the motility of prostate cancer cells by targeting cadherin-2. *Mol Cell Biochem,* 476**,** 3635-3646.

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, C. Z., ZHANG, J. X., ZHANG, A. L., SHI, Z. D., HAN, L., JIA, Z. F., YANG, W. D., WANG, G. X., JIANG, T., YOU, Y. P., PU, P. Y., CHENG, J. Q. & KANG, C. S. 2010. MiR-221 and miR-222 target PUMA to induce cell survival in glioblastoma. *Mol Cancer,* 9**,** 229.

ZHANG, D., WANG, H., SUN, M., YANG, J., ZHANG, W., HAN, S. & XU, B. 2014a. Speckle-type POZ protein, SPOP, is involved in the DNA damage response. *Carcinogenesis,* 35**,** 1691-7.

ZHANG, D. Y., TURBERFIELD, A. J., YURKE, B. & WINFREE, E. 2007. Engineering entropy-driven reactions and networks catalyzed by DNA. *Science,* 318**,** 1121-5.

ZHANG, H., JIN, X. & HUANG, H. 2023. Deregulation of SPOP in Cancer. *Cancer Res,* 83**,** 489-499.

ZHANG, J., BU, X., WANG, H., ZHU, Y., GENG, Y., NIHIRA, N. T., TAN, Y., CI, Y., WU, F., DAI, X., GUO, J., HUANG, Y. H., FAN, C., REN, S., SUN, Y., FREEMAN, G. J., SICINSKI, P. & WEI, W. 2018a. Cyclin D-CDK4 kinase destabilizes PD-L1 via cullin 3-SPOP to control cancer immune surveillance. *Nature,* 553**,** 91-95.

ZHANG, J., GAO, K., XIE, H., WANG, D., ZHANG, P., WEI, T., YAN, Y., PAN, Y., YE, W., CHEN, H., SHI, Q., LI, Y., ZHAO, S. M., HOU, X., WEROHA, S. J., WANG, Y., ZHANG, J., KARNES, R. J., HE, H. H., WANG, L., WANG, C. & HUANG, H. 2021. SPOP mutation induces DNA methylation via stabilizing GLP/G9a. *Nat Commun,* 12**,** 5716.

ZHANG, P., GAO, K., TANG, Y., JIN, X., AN, J., YU, H., WANG, H., ZHANG, Y., WANG, D., HUANG, H., YU, L. & WANG, C. 2014b. Destruction of DDIT3/CHOP protein by wild-type SPOP but not prostate cancer-associated mutants. *Hum Mutat,* 35**,** 1142-51.

ZHANG, P., WANG, D., ZHAO, Y., REN, S., GAO, K., YE, Z., WANG, S., PAN, C. W., ZHU, Y., YAN, Y., YANG, Y., WU, D., HE, Y., ZHANG, J., LU, D., LIU, X., YU, L., ZHAO, S., LI, Y., LIN, D., WANG, Y., WANG, L., CHEN, Y., SUN, Y., WANG, C. & HUANG, H. 2017. Intrinsic BET inhibitor resistance in SPOP-mutated prostate cancer is mediated by BET protein stabilization and AKT-mTORC1 activation. *Nat Med,* 23**,** 1055-1062.

ZHANG, P., ZHANG, M., HAN, R., ZHANG, K., DING, H., LIANG, C. & ZHANG, L. 2018b. The correlation between microRNA-221/222 cluster overexpression and malignancy: an updated meta-analysis including 2693 patients. *Cancer Manag Res,* 10**,** 3371-3381.

ZHANG, T., ZHAO, F., LIN, Y., LIU, M., ZHOU, H., CUI, F., JIN, Y., CHEN, L. & SHENG, X. 2024. Integrated analysis of single-cell and bulk transcriptomics develops a robust neuroendocrine cell-intrinsic signature to predict prostate cancer progression. *Theranostics,* 14**,** 1065-1080.

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 NH2-terminal and carboxyl-terminal sequences. *J Biol Chem,* 269**,** 13115-23.

ZHU, H., REN, S., BITLER, B. G., AIRD, K. M., TU, Z., SKORDALAKES, E., ZHU, Y., YAN, J., SUN, Y. & ZHANG, R. 2015. SPOP E3 Ubiquitin Ligase Adaptor Promotes Cellular Senescence by Degrading the SENP7 deSUMOylase. *Cell Rep,* 13**,** 1183-1193.

ZHU, Y., WEN, J., HUANG, G., MITTLESTEADT, J., WEN, X. & LU, X. 2021. CHD1 and SPOP synergistically protect prostate epithelial cells from DNA damage. *Prostate,* 81**,** 81-88.

ZHUANG, M., CALABRESE, M. F., LIU, J., WADDELL, M. B., NOURSE, A., HAMMEL, M., MILLER, D. J., WALDEN, H., DUDA, D. M., SEYEDIN, S. N., HOGGARD, T., HARPER, J. W., WHITE, K. P. & SCHULMAN, B. A. 2009. Structures of SPOP-substrate complexes: insights into molecular architectures of BTB-Cul3 ubiquitin ligases. *Mol Cell,* 36**,** 39-50.