



Hartley, Andrew (2022) *Investigating ARID1A in prostate cancer*. PhD thesis.

<http://theses.gla.ac.uk/83300/>

Copyright and moral rights for this work are retained by the author

A copy can be downloaded for personal non-commercial research or study,
without prior permission or charge

This work cannot be reproduced or quoted extensively from without first
obtaining permission in writing from the author

The content must not be changed in any way or sold commercially in any
format or medium without the formal permission of the author

When referring to this work, full bibliographic details including the author,
title, awarding institution and date of the thesis must be given

Enlighten: Theses

<https://theses.gla.ac.uk/>
research-enlighten@glasgow.ac.uk

Investigating ARID1A in Prostate Cancer

Andrew Hartley

Thesis submitted in fulfilment of the requirements for
Degree of Doctor of Philosophy



University
of Glasgow

Beatson Institute for Cancer Research, Institute of Cancer
Sciences, College of Medical, Veterinary and Life Sciences,
University of Glasgow

Abstract

Prostate cancer (PC) is the most common cancer in men in the developed world, and the second leading cause of cancer death in the U.K. Most of these deaths are caused by the advanced and metastatic forms of this disease which have no curative options. Understanding and targeting metastatic prostate cancer remains one of the primary research goals in prostate cancer research.

The *Sleeping Beauty* screen used a forward-mutagenesis transposon-based system to try to identify novel drivers of prostate cancer which cooperated with loss of *Pten* *in vivo*. This screen showed that reduced *Arid1a* expression correlated with a poorer survival and elevated metastasis *in vivo*. ARID1A functions as part of the BAF chromatin remodelling complex, an epigenetic regulator of gene expression which controls chromatin accessibility.

Using a genetically engineered mouse (GEM) model, we were able to delete *Arid1a* and *Pten* from the mouse prostate genome to investigate the genetic interaction of these two genes. This showed that *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* mice developed clinical endpoint tumours very rapidly compared to *Pb-Cre+ve Pten^{fl/fl} Arid1a^{+/+}* (median 4 months vs 10 months respectively). These *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* tumours also have an invasive phenotype. *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/+}* tumours however showed no significant change in survival compared to *Pb-Cre+ve Pten^{fl/fl} Arid1a^{+/+}*.

RNA-Seq and ChIP-Seq were performed to understand how loss of *Arid1a* influenced tumorigenesis in our GEM model. This showed that following loss of *Arid1a*, tumours upregulated growth and cell cycle signalling pathways. Ki67, a marker of proliferation, confirmed that *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* tumours had

a higher rate of proliferation compared to *Pb-Cre+ve Pten^{fl/fl} Arid1a^{+/+}* tumours. ChIP-Seq findings suggested that *Arid1a* directly binds genes involved in migration and invasion which were subsequently upregulated following loss of *Arid1a*. This suggested that the increased proliferation may be due to a re-targeting of the BAF complex rather than direct repression of growth by *Arid1a* itself. Immunohistochemistry (IHC) and GSEA suggests that P53-dependent senescence signalling is still present in *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* though has likely been bypassed.

Generating ARID1A knockout clones with CRISPR-Cas9 of DU145, a human PC cell line, phenotypically increases their growth and survival following ARID1A loss, though decreases their invasion abilities. RNA-Seq on DU145 ARID1A^{KO} clones shows an elevation in growth and cell cycle signalling following ARID1A loss, though a reduction in migration signalling. RNA-Seq suggests however that these cells may be capable of local extracellular matrix remodelling through matrix metalloprotease upregulation. Comparisons between the GEM model and DU145 show overlapping upregulation in pathways relating to cell cycle and DNA-damage response. DNA damage response findings correlate with clinical data suggest ARID1A mutant PC has high mutational burden.

Data presented in this thesis therefore shows that *Arid1a* loss *in vivo* cooperated with *Pten* loss and produced aggressive and invasive prostate tumours. Loss of *Arid1a* upregulated growth and cell cycle signalling, though this is likely due to a re-targeting of the BAF complex itself as *Arid1a* was only shown to directly repress genes involved in migration and invasion. Using ARID1A^{KO} PC cell line DU145 also shows an elevation in cell cycle signalling. Understanding how loss of ARID1A causes BAF complex re-targeting warrants additional research.

Table of Contents

Investigating ARID1A in Prostate Cancer	1
Abstract	2
Table of Contents	4
List of Figures	8
List of Tables	10
Acknowledgment	11
Author's Declaration	13
Definitions/Abbreviations	14
1 Introduction.....	17
1.1 Introduction to Prostate Cancer.....	18
1.1.1 Prostate Cancer Epidemiology	18
1.1.2 Prostate Cancer Pathophysiology	20
1.1.3 Classification of Prostate Cancer.....	22
1.2 Prostate Cancer Signaling	26
1.2.1 PTEN and PI3K/AKT Signaling	26
1.3 Chromatin Remodelling.....	28
1.3.1 BAF complex.....	28
1.3.2 ARID1A and BAF Complex Dynamics.....	32
1.3.3 Genetically Engineered Mouse Models of Prostate Cancer	35
1.3.4 Mutagenesis to Study Prostate Cancer.....	38
2 Materials and Methods.....	41
2.1 <i>In vivo</i> materials and methods.....	42
2.1.1 Generation of Mice colonies	42
2.1.2 Mouse experiments for chapter 3	42
2.1.3 Tissue Isolation.....	43
2.1.4 Immunohistochemistry	43
2.1.5 Halo Scoring	46
2.2 Cell culture	48
2.2.1 Cell lines	48
2.2.2 Cryopreservation of cells	49
2.2.3 <i>In vitro</i> proliferation assay as measured by cell count.....	50
2.2.4 <i>In vitro</i> colony forming assay	50
2.2.5 <i>In vitro</i> scratch wound healing assay	50
2.2.6 <i>In vitro</i> transwell invasion assay	51
2.2.7 Reverse transfection of siRNA for transient gene knockdown	52

2.2.8	Generating stable ARID1A knockout clones in human cell lines	52
2.3	Protein Analysis	53
2.3.1	Protein Extraction	53
2.3.2	Protein Quantification	53
2.3.3	Western electrophoresis and immunoblotting.....	54
2.4	RNA Analysis.....	55
2.4.1	RNA extraction	55
2.4.2	Reverse transcription of RNA to produce cDNA	55
2.4.3	Quantitative real-time polymerase chain reaction (QRT-PCR).....	56
2.5	Omics techniques.....	57
2.5.1	RNA-Sequencing.....	57
2.5.2	ChIP-Seq	58
2.5.3	ATAC-Seq	60
2.6	Statistics	68
3	Loss of <i>Arid1a</i> Cooperates With Loss of <i>Pten</i> to Drive Prostate Tumour Growth in a Murine Model.....	70
3.1	Introduction	71
3.2	Hypothesis and Aims	74
3.3	Results	75
3.3.1	Sleeping Beauty - A Transposon-based Forward Mutagenesis Screen	75
3.3.2	Floxing <i>Arid1a</i> and <i>Pten</i> Leads to Loss of Protein in Murine Tumours	77
3.3.3	Loss of <i>Arid1a</i> Accelerates Tumorigenesis in <i>Pten</i> Null Mice.....	82
3.3.4	<i>Pten</i> Loss is a Prerequisite to <i>Arid1a</i> Loss Being Tumorigenic	86
3.3.5	Characterising Systemic Effects of <i>Pb-Cre +ve Pten^{fl/fl} Arid1a^{fl/fl}</i> mice in blood biochemistry.....	91
3.3.6	Characterising Systemic Effects of <i>Pb-Cre +ve Pten^{fl/fl} Arid1a^{fl/fl}</i> mice with tissue staining for Arid1a.....	102
3.4	Discussion.....	106
3.4.1	ARID1A and PTEN in PI3K-AKT Signalling	106
3.4.2	ARID1A in other murine models	107
3.4.3	Methodological changes and future directions	109
3.4.4	Conclusions	110
4	Loss of <i>Arid1a</i> upregulates cell cycle and migration signalling in <i>Pten</i> null mice	111
4.1	Introduction	112
4.1.1	ARID1A as a tumour suppressor in cancer	112
4.1.2	Senescence in <i>Pb-Cre+ve Pten^{fl/fl}</i> mice.....	113
4.2	Hypothesis and Aims	115

4.3	Results	116
4.3.1	RNA-Sequencing shows upregulated cell cycle signalling following <i>Arid1a</i> loss	116
4.3.2	Loss of one copy of <i>Arid1a</i> does not highly dysregulate <i>Pten^{fl/fl}</i> prostate tumour signalling	118
4.3.3	Loss of <i>Arid1a</i> upregulates cell cycle signalling in <i>Pten^{fl/fl}</i> prostate tumours.....	120
4.3.4	ChIP-Seq performing and optimisation	124
4.3.5	Quantifying ChIP DNA pulldown by QRT-PCR	125
4.3.6	ChIP-Seq results.....	127
4.3.7	ATAC-Sequencing	134
4.3.8	<i>Arid1a</i> loss may bypass P53-dependent senescence in <i>Pten</i> null mice 138	
4.4	Discussion	146
4.4.1	ARID1A as a regulator of cell cycle in cancer	146
4.4.2	Neuroendocrine Prostate Cancer	150
4.4.3	Methodological changes and future directions	151
4.4.4	Conclusions	152
5	ARID1A in Human Prostate Cancer Cell Lines	153
5.1	Introduction	154
5.1.1	ARID1A and Prostate Cancer	154
5.2	Hypothesis and Aims:	156
5.3	Results	157
5.3.1	ARID1A mutations in human prostate cancer - CBioPortal.....	157
5.3.2	Cell seeding and lysis buffer optimisation for western blotting	160
5.3.3	Initial characterising ARID1A knockdown in PC cell lines:	163
5.3.4	Phenotypic characterisation of ARID1A knockdown in PC cell lines: 167	
5.3.5	Generating ARID1A KO clones:.....	170
5.3.6	Characterising DU145 ARID1A KO clones by RNA-Seq.....	173
5.3.7	Phenotypic analysis of ARID1A KO DU145 clones.....	178
5.3.8	Comparing human and murine datasets following ARID1A loss	181
5.4	Discussion	183
5.4.1	LNCaP and PC3 dependency on ARID1A	183
5.4.2	DU145 ARID1A KO cells compared to the <i>Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}</i> mouse model.....	185
5.4.3	Methodological changes and future directions	186
5.4.4	Conclusions	187
6	Discussion and future directions	187
6.1	<i>Arid1a</i> in a murine GEM model	187

6.2	ARID1A in Human Prostate Cancer	188
6.3	Relevance for the Clinic	189
6.4	Future directions	190
6.5	Concluding Summary	191
	References.....	192

List of Figures

Figure 1.1 - The 10 most common causes of cancer death in males.....	18
Figure 1.2 - Prostate cancer mortality rate by age.....	19
Figure 1.3- PC incidence in different age groups over time	20
Figure 1.4 - Prostate Gland Structure	21
Figure 1.5 - Gleason score to grade prostate cancer.....	23
Figure 1.6 - TNM Diagram	24
Figure 1.7 - PC stage at diagnosis U.K.....	25
Figure 1.8 PI3K-AKT signalling pathway	27
Figure 1.9 - Schematic of BAF complex interchangeability	30
Figure 1.10 - BAF Complex in PC	31
Figure 1.11 ARID1A and ARID1B dynamically alter chromatin organisation.....	36
Figure 1.12 - Types of mutagenesis screens	40
Figure 2.1 - Halo Classification to Identify Cell Types	47
Figure 2.2 - ATAC-Seq flowthrough	61
Figure 3.1 - <i>Probasin Cre-recombinase</i> system in mouse prostate.....	72
Figure 3.2 - <i>Sleeping Beauty</i> screen breeding scheme.....	73
Figure 3.3 - <i>Sleeping Beauty</i> transposon insertion sites in <i>Arid1a</i> gene.....	75
Figure 3.4 - Reduced expression of <i>Arid1a</i> reduces survival and promotes metastasis	77
Figure 3.5 - ARID1A nuclear staining is reduced following Pb-Cre expression to delete the gene in a mouse model.....	78
Figure 3.6 - Recombination of <i>Pten</i> did not lower detectable levels of PTEN protein by IHC.....	79
Figure 3.7 - Loss of PTEN leads to pAKT phosphorylation which is unaltered by <i>Arid1a</i> status	81
Figure 3.8 - Tumour morphology in Pb-Cre+ <i>Pten</i> ^{fl/fl} mice with different <i>Arid1a</i> status.....	83
Figure 3.9 - Loss of <i>Arid1a</i> reduces survival in <i>Pten</i> null mice	83
Figure 3.10 - Loss of <i>Arid1a</i> alters tumour morphology in anterior and dorsolateral prostate lobes	84
Figure 3.11 - Loss of <i>Arid1a</i> causes invasive phenotype of dorsolateral prostate	84
Figure 3.12 - Tumour weight at clinical endpoint for each genotype	86
Figure 3.13- Complete loss of <i>Pten</i> is required for <i>Arid1a</i> loss to be tumorigenic	87
Figure 3.14 - <i>Pten</i> loss is a prerequisite to <i>Arid1a</i> loss being tumorigenic	89
Figure 3.15 - Spontaneous loss of PTEN in <i>Pb-Cre+ve Pten</i> ^{fl/+} <i>Arid1a</i> ^{fl/fl} dorsolateral tumour.....	90
Figure 3.16 - Tumour weights in <i>Pb-Cre+ve Pten</i> ^{fl/+} cohorts showing spontaneous loss of PTEN is required for tumour growth in <i>Arid1a</i> ^{fl/fl} prostates.....	91
Figure 3.17 - <i>Pb-Cre+ve Pten</i> ^{fl/fl} <i>Arid1a</i> ^{fl/fl} prostate morphology over time.....	93
Figure 3.18 - Blood biochemistry describing haematology of <i>Pb-Cre+ve Pten</i> ^{fl/fl} <i>Arid1a</i> ^{fl/fl} mice over time.....	95
Figure 3.19 - Blood biochemistry indicating liver function of <i>Pb-Cre+ve Pten</i> ^{fl/fl} <i>Arid1a</i> ^{fl/fl} mice over time.....	97
Figure 3.20 - Blood biochemistry indicating kidney function	99
Figure 3.21 - Blood biochemistry indicating white cells function	101
Figure 3.22 - ARID1A protein levels in testis, brain, and liver	103
Figure 3.23 - ARID1A protein levels in lymph node (L/N), heart, and kidney ...	104
Figure 3.24 - ARID1A protein levels in spleen, gut, and lung	105

Figure 4.1 Quality assessment for RNA-Seq.....	118
Figure 4.2 - Network and Pathway analysis of <i>Pb-Cre+ve Pten^{fl/fl} Arid1a^{+/+}</i> vs <i>Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/+}</i> prostate tumours.....	120
Figure 4.3 - Network and Pathway analysis of <i>Pb-Cre+ve Pten^{fl/fl} Arid1a^{+/+}</i> vs <i>Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}</i> prostate tumours	122
Figure 4.4 - Loss of <i>Arid1a</i> and <i>Pten</i> upregulates pro-growth signalling	123
Figure 4.5 - Mnase digest optimisation for ChIP	125
Figure 4.6 - QRT-PCR for ChIP	126
Figure 4.7 - Library preparation for ChIP	127
Figure 4.8 - Proximity of ARID1A binding to genes	128
Figure 4.9 - Networks and pathway analysis of ARID1A-bound genes.....	130
Figure 4.10 - ARID1A binds and represses drivers of migration and invasion	131
Figure 4.11 - β -catenin signalling is downregulated following loss of <i>Arid1a</i> ...	133
Figure 4.12 - ATAC-Seq Sample Variance	135
Figure 4.13 - ATAC-Seq fragment alignment and peak-calling.....	138
Figure 4.14 - Loss of <i>Arid1a</i> increases cell proliferation in <i>Pten</i> null prostate tumours	139
Figure 4.15 - Loss of <i>Arid1a</i> upregulates P21 in <i>Pten</i> null mice	141
Figure 4.16 - Loss of <i>Arid1a</i> does not alter gamma-H2AX in <i>Pten</i> null mice	143
Figure 4.17 - Loss of <i>Arid1a</i> upregulates P53.....	145
Figure 4.18 - GSEA DNA repair and P53	146
Figure 4.19 - ARID1A can control cell cycle by altering accessibility to cell cycle regulators.....	149
Figure 5.1 - ARID1A mutations in Prostate Cancer Cohorts	157
Figure 5.2 - Hotspot mutations in ARID1A gene	158
Figure 5.3 - Lysis buffer optimisation	161
Figure 5.4 - Optimisation of seeding densities	162
Figure 5.5 - Characterising ARID protein levels in prostate cancer cell lines ...	164
Figure 5.6 - Validating ARID1A protein knockdown from transfection.....	165
Figure 5.7 - ARID1A expression in PC cell lines	167
Figure 5.8 - Transient knockdown of ARID1A reduces growth in PC cell lines	168
Figure 5.9 - Transient knockdown of ARID1A reduces colony forming assays in PC cell lines	169
Figure 5.10 - Optimisation of puromycin kill curve.....	171
Figure 5.11 - ARID1A CRISPR-Cas9 plasmid expresses in LNCaP and PC3	172
Figure 5.12 - Validating ARID1A knockout of DU145 clones	173
Figure 5.13 - RNA-Seq library preparation and characterisation	175
Figure 5.14 - Pathway analysis for DU145 KO clones	176
Figure 5.15 - ARID1A knockout increases growth, survival, and wound healing in DU145 cells	179
Figure 5.16 - Loss of ARID1A impairs chemotaxis	180
Figure 5.17 - Loss of ARID1A upregulates DDR and cell cycle pathways.....	182

List of Tables

Table 2.1 - Mouse genotypes and cohort sizes	43
Table 2.2 - Immunohistochemistry Reagents	46
Table 2.3 - Immunohistochemistry antibodies	46
Table 2.4 - Immunoblotting antibodies and dilutions	55
Table 2.5 - QRT-PCR primer sequences	56
Table 2.6 - Indexes for CHIP samples	60
Table 2.7 - ATAC-Seq Reagents.....	62
Table 2.8 - ATAC-Seq initial library construction.....	65
Table 2.9 - ATAC-Seq barcodes in each library	66
Table 2.10 - ATAC-Seq PCR steps for library amplification	67
Table 5.1 - Alteration frequency of oncogenic drivers of prostate cancer with or without ARID1A mutation	160
Table 5.2 - Molecular drivers and characteristics of LNCaP vs CWR22	185

Acknowledgment

Firstly, I would like to give my deepest gratitude to my supervisor Dr Imran Ahmad for all of his guidance and support during my PhD. Imran has helped me grow both personally and professionally and I am grateful for all he has done. I would like to also thank my second supervisor, Professor Karen Blyth for her endless encouragement and guidance throughout my PhD, as well as for allowing me to also be a part of her group.

I am grateful to all past and present members of M31 and R08 particularly Professor Hing Leung. These people taught me how to be a scientist, and in particular I would like to thank Dr Laura Galbraith for her support and advice when I first began my PhD. I would also like to thank the services of the Beatson including staff of the BSU and BRU facilities, histology, and our molecular services teams.

I would also like to thank Syed Umbreen of the University of Belfast, Glasgow Polyomics, and Robin Shaw for their help and expertise with all sequencing and informatics. This project involved more informatics than I would ever have hoped and without these people this thesis would be half the length and not nearly as interesting.

Finally, I would like to thank my partner, Beth, and all my family for their continuous support and encouragement over the last 4 years. Thank you all.

Author's Declaration

I declare that the work presented in this thesis has been carried out by myself, except where otherwise cited or acknowledged. It is entirely my own composition and has not, in whole or part, been submitted for any other higher degree.

Andrew Hartley

September, 2022

Definitions/Abbreviations

°C	Degrees Celsius
EMS	Ethyl Methanesulphonate
ADT	Androgen deprivation therapy
AKT	Protein kinase B
ANXA1	Annexin A1
APC	Adenomatous polyposis coli
AR	Androgen receptor
ARID	AT-Rich Interacting Domain
ARID1A	ARID containing protein 1A
ARID1B	ARID containing protein 1B
ARID2	ARID containing protein 2
ATCC	American type culture collection
ATP	Adenosine triphosphate
AXIN2	Axis Inhibition protein 2
BAF	BRG1-associated factors
BPH	Benign prostate hyperplasia
BRAF	v-Raf murine sarcoma viral oncogene homolog B
BRG1	BRM-related gene 1
BRM	Brahma
BSA	Bovine serum albumin
CASC3	Cancer susceptibility candidate gene 3
CCND1	Cyclin D1
ChIP-Seq	ChIP-Sequencing
ChIP	Chromatin immunoprecipitation
CLDN4	Claudin 4
CO2	Carbon Dioxide
CTNNB1	Beta Catenin
CXCL2	Chemokine ligand 2
CXCL3	Chemokine ligand 3
DHT	dihydrotestosterone
DMSO	Dimethyl sulphoxide
DNA	Deoxyribonucleic acid
DPX	Dibutylphthalate Polystyrene Xylene
ECL	Enhanced chemiluminescence
ENU	N-ethyl-N-nitrosourea
ER	Oestrogen Receptor
EZH2	Enhanced of zeste homology 2
FBS	Foetal Bovine Serum
FFPE	Formalin-fixed paraffin embedded
Fl	Flox
GEM	Genetically engineered mouse
H	Hours
H&E	Haematoxylin and eosin
H2AX	H2A histone family member X
HER2	Human epidermal growth factor 2
HIER	Heat-induced epitope retrieval
HDR	Homology-directed repair

HRP	Horseradish peroxidase
IHC	Immunohistochemistry
INT	Insertion
IRFP	Near Infra-red fluorescent protein
KRAS	Kirsten Rat Sarcoma Virus
LAMC2	Laminin Subunit Gamma 2
LN	Lymph node
MSI	Microsatellite instability
OCCC	Ovarian clear cell carcinoma
ORF	open-reading frame
P21	CIP1/WAF1
P53	Tumour protein 53
PanIN	Pancreatic intraepithelial neoplasm
PARD6	Partitioning defective 6 homolog alpha
PBac	<i>piggyBac</i>
PBAF	Polybromo-BAF
Pb-Cre	Probasin Cre-recombinase
PBS	Phosphate buffered saline
PC	Prostate cancer
PCA	Principle component analysis
PCR	Polymerase chain reaction
pH	Potential of hydrogen
PI	Phosphatidylinositol
PI3K	Phosphatidylinositol 3-kinase
PI3KIP1	PI3K interacting protein 1
PIP2	PI 3-4 biphosphate
PIP3	PI 3,4,5 triphosphate
PLAUR	Plasminogen activator urokinase receptor
PMSF	Phenylmethylsulfonyl fluoride
PSA	Prostate specific antigen
PTEN	Phosphatase and tensin homolog
PVDF	Polyvinylidene fluoride
QRT-PCR	Quantitative real-time PCR
rEGF	human recombinant epidermal growth factor
RNA	ribonucleic acid
RNA-I	ribonucleic acid interference
RNA-Seq	RNA-Sequencing
ROS	Reactive oxygen species
RT	Room temperature
RTK	Receptor tyrosine kinase
SB	<i>Sleeping Beauty</i>
SDS	sodium dodecyl sulfate
shRNA	short-hairpin RNA
SWI/SNF	Switch-induced/sucrose non-fermentable
TBST	Tris-buffered saline with tween
Tbt	Tween
TGF- β	Transforming growth factor-beta
	T-cell lymphoma invasion and metastasis-inducing protein 1
TIAM1	
TNM	Tumour, node, metastasis

TSS

Transcriptional start site

1 Introduction

1.1 Introduction to Prostate Cancer

1.1.1 Prostate Cancer Epidemiology

Prostate cancer (PC) is the most common cancer in men in the U.K. with 48,000 new cases diagnosed every year, accounting for 26% all new cancer cases in males (CRUK-Incidence [Accessed July 2022]). PC kills around 12,000 men per year making it the second most lethal cancer in men in the U.K after lung cancer (Figure 1.1)(CRUK-Mortality [Accessed July 2022]). Worldwide, PC accounts for 15% of all cancer cases with 1.25 million new cases per year and 360,000 deaths (Fund , Taitt 2018).

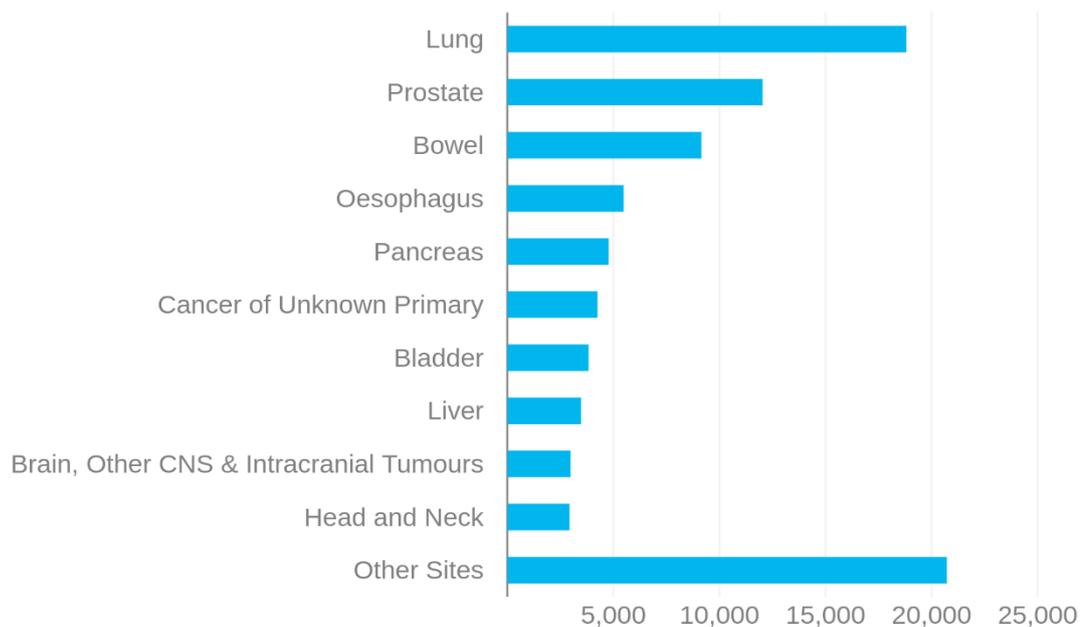


Figure 1.1 - The 10 most common causes of cancer death in males

Lung cancer is the most common cancer in men in the U.K., with prostate cancer being the second most common cancer. U.K. 2017. Source: cruk.org/cancerstats

PC is far more common in older men, with only 0.3% of U.K. incidence occurring in men under the age of 50. Both incidence and mortality increase exponentially

with age with 90% of PC deaths occurring in men over the age of 70 (Figure 1.2). However, overall mortality in PC patients have steadily declined since the early 1990s falling from around 60 per 100,000 people to around 48 per 100,000 in recent years (CRUK-Mortality [Accessed July 2022]).

There are racial differences associated with PC incidence (*Lloyd, Hounsome et al.* 2015). In the U.K., Asian males have a 6.3-10.5% lifetime risk of being diagnosed with PC, white males have a significantly higher risk at 13.2-15.0%, whereas black males have the highest risk at 23.5-37.2%.

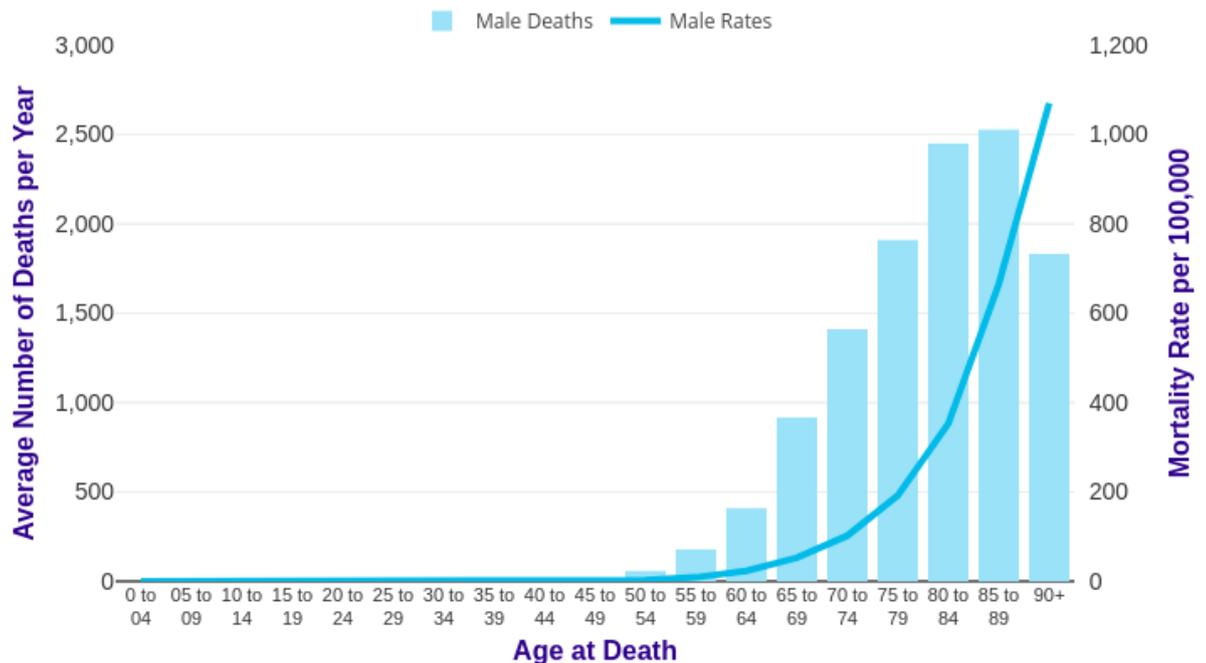


Figure 1.2 - Prostate cancer mortality rate by age

Prostate cancer mortality rates rise exponentially from the age of 50 peaking to over 1,000 deaths per 100,000 men in the 90-age group. Deaths per year to prostate cancer peak between 80 – 89 years old at around 2,500 deaths. Source: cruk.org/cancerstats

Since 1990, there was a 41% increase in PC incidence (CRUK-Incidence [Accessed July 2022]). This rise is likely due to the introduction of the prostate-specific antigen (PSA) test which was introduced in the early 1990s (Roobol 2013)(Figure

1.3). PSA is a biomarker secreted into the blood by both healthy and cancerous prostate cells so indicates the amount of prostate tissue present. Elevations in blood PSA can be an indicator of benign prostate hyperplasia (BPH), prostate infection or inflammation, or cancer and is not diagnostic alone without a biopsy.(Velonas, Woo *et al.* 2013). By 2035, there are expected to be 77,000 new PC cases diagnosed accounting for 18,000 deaths each year (CRUK-Incidence [Accessed July 2022]).

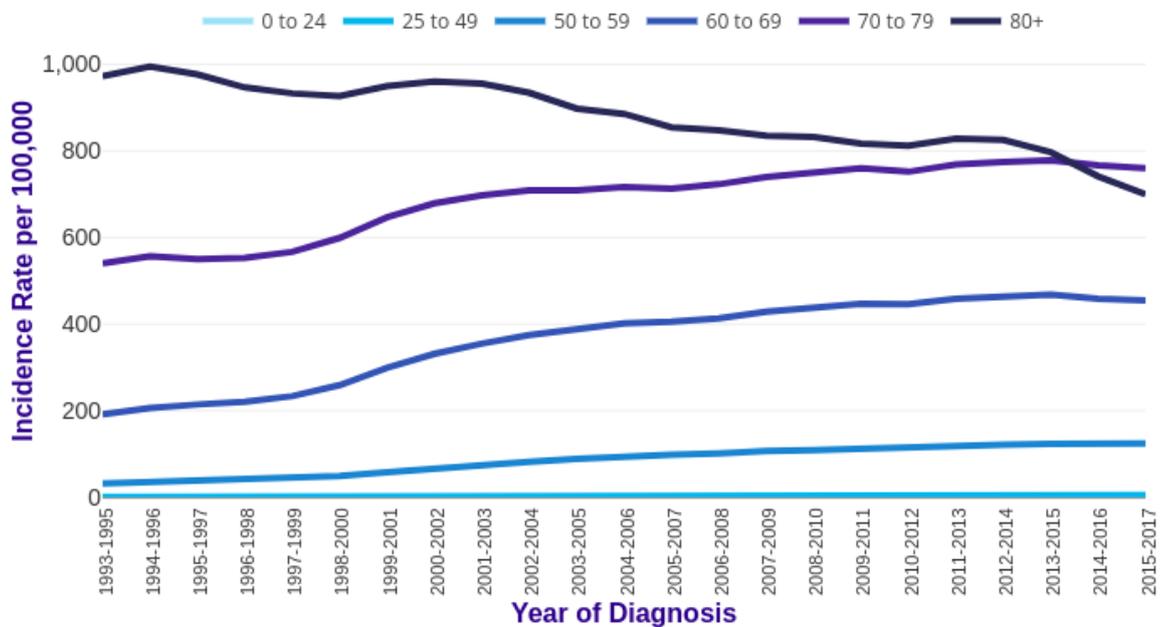


Figure 1.3- PC incidence in different age groups over time

Age 25-49 account for the largest increase in incidence with a 544% increase, although this group still makes up a small fraction of total PC cases. Age 70-79 incidence rose by 41% to become the largest age group of PC cases following a 28% decrease in age 80+ which was the largest age group of PC cases from 1993-2015 (U.K. 2020).Source: cruk.org/cancerstats

1.1.2 Prostate Cancer Pathophysiology

The prostate is a walnut sized gland located in the male pelvis located between the bladder and the penis. Multiple small glands in the prostate secrete fluid that enriches the sperm during ejaculation.

Prostate epithelium turnover is relatively slow; however, 50% of men aged 50-61 have developed BPH as a consequence of prostate growth (Lloyd, Hounsome *et al.* 2015). BPH is caused by a proliferation of prostatic cells which leads to an increase in prostate mass and can cause urinary obstruction (Skinder, Zacharia *et al.* 2016). The typical prostate weighs 20 ± 6 grams, whereas a prostate with BPH weighs 33 ± 16 grams (Skinder, Zacharia *et al.* 2016).

The glands have a tubuloalveolar structure with prostate epithelial lining the lumen to form secretory acini (Denmeade SR 2003). There are two main cell types involved in the gland: luminal and basal cells, both of which develop independently from urogenital sinus epithelium. Luminal cells line the lumen of the gland and secrete fluid which drains through ducts into the prostate urethra. Basal cells support duct integrity by maintaining the basement membrane. Additionally, basal cells have also been shown to maintain the proper differentiation of luminal cells (Kurita, Medina *et al.* 2004). These luminal and basal cells are surrounded by a stroma of fibroblasts, endothelial, neurons, and smooth muscle cells (Denmeade SR 2003)(Figure 1.4).

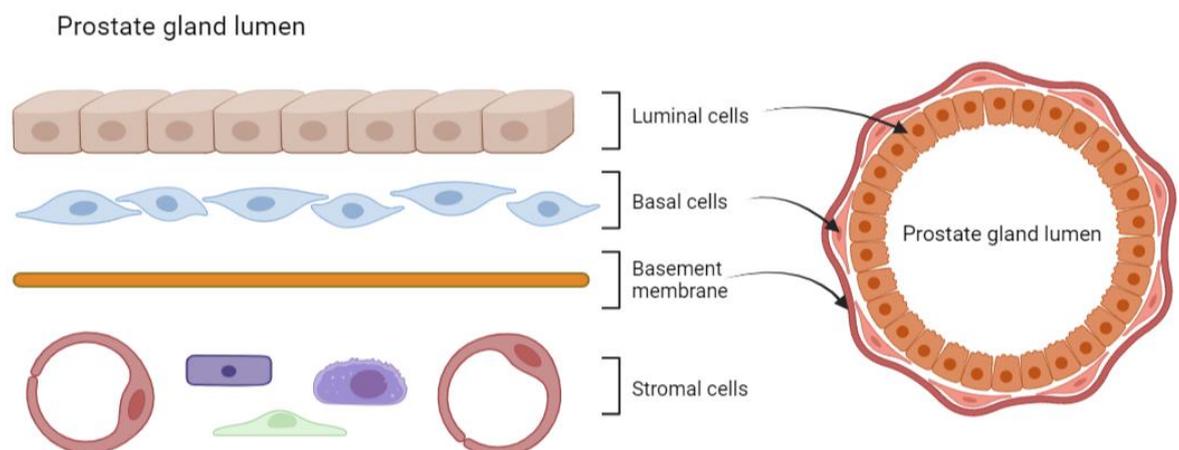


Figure 1.4 - Prostate Gland Structure

Prostate gland is made up of layers of distinct cells to support the gland. Luminal epithelial cells line the prostate gland lumen and are supported by basal epithelial cells. A basement membrane surrounds the gland itself separating it from a stroma of fibroblasts, endothelial cells, neurons, and smooth muscle cells.

Molecularly, the androgen receptor (AR) is required for normal prostate development and maintenance. AR signalling is driven by the androgen testosterone and its more potent derivative, dihydrotestosterone (DHT), which act as ligands to AR. After binding an androgen, AR dimerises and translocates to the nucleus where it activates transcription of AR-target genes. These genes provide survival and growth signals to prostate cells and are thus often exploited in cancer (Fujita and Nonomura 2019).

However, AR has been shown to be dispensable in differentiated luminal and basal cells although required for maintenance of stem cells of both types (Xie, Liu *et al.* 2017). AR activity can be ablated by removing androgen synthesis by preventing testosterone production. This is a mainstay treatment in PC via androgen deprivation therapy (ADT) (Perlmutter and Lepor 2007). As such castrated men such as eunuchs often rarely develop BPH and even more rarely develop PC (Stocking, Fiandalo *et al.* 2016).

1.1.3 Classification of Prostate Cancer

The most common subtype of PC is acinar adenocarcinoma, which constitutes nearly all of the PC cases (Baig, Hamid *et al.* 2015). These cancers can be classified in two ways, grading and staging.

PC can be graded by using the Gleason score (Figure 1.5). The Gleason score is a means to determine how similar the tissue biopsy looks similar to normal healthy tissue (Chen and Zhou 2016). The biopsy will be assigned two numbers which represent the most common and second most common cell scores for the cells

on a scale of 1 to 5. Healthy tissue is given a score of 1 on a Gleason score, whereas a complete absence of normal glandular structure would result in a score of 5. However, Gleason scores of 1 - 2 no longer define prostate cancer and only higher Gleason scores of 3 -5 denote cancer (Chen and Zhou 2016). The total score allows the cancer to be assigned a grade (Figure 1.5).

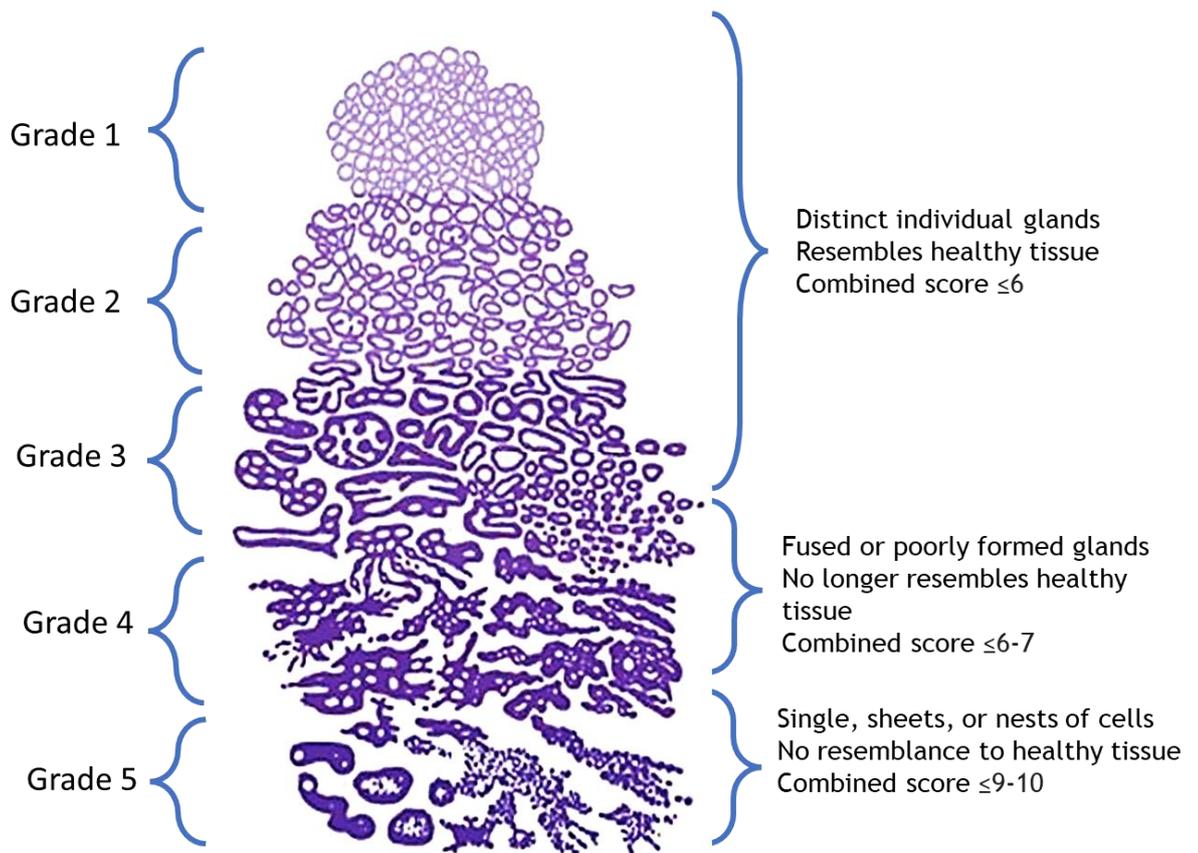


Figure 1.5 - Gleason score to grade prostate cancer

Gleason score is assigned to the two most common features of a biopsy. Low scores between 1-3 indicate clear and distinct glands and structures. Scores up to 5 show a complete lack of normal gland structure. Adding both Gleason scores together i.e., 1+2 or 4+4 gives a grading of the prostate. Adapted from Chen & Zhou (2016).

Staging of PC uses the tumour, node, metastasis (TNM) system (Pospihalj 2015).

Tumour stages 1 and 2 refer to the size of the tumour within the prostate gland

(Figure 1.6). Tumour stage 3 is representative of a cancer which has broken out of the prostate capsule. Tumour stage 3A indicates the cancer has not migrated to the seminal vesicles whereas tumours stage 3B indicates it has migrated to the seminal vesicles. In tumour stage 4, the cancer has spread to other neighbouring tissues such as the bladder.

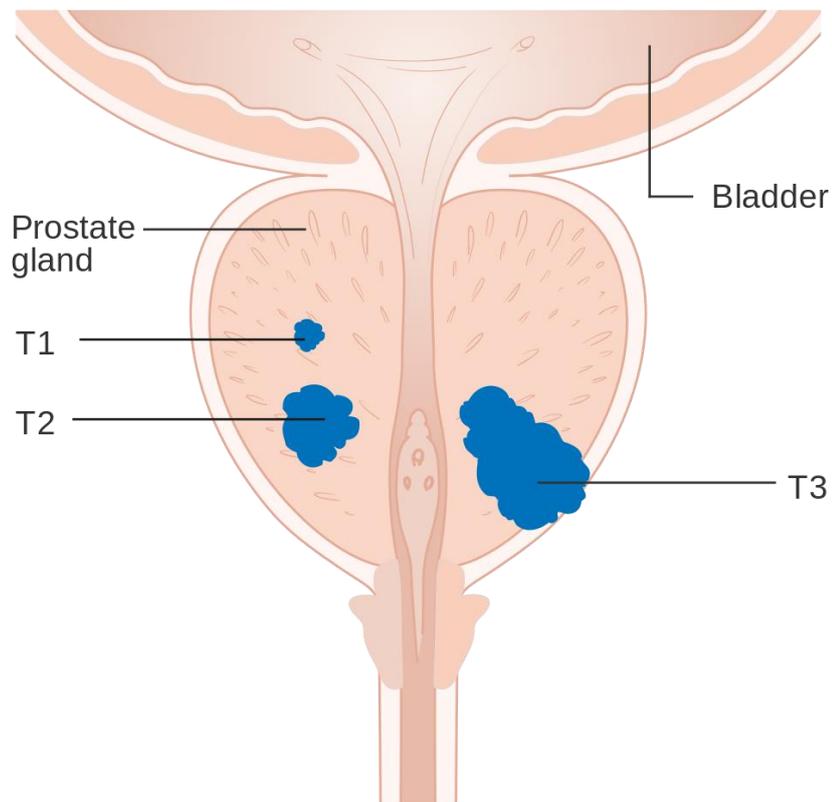


Figure 1.6 - TNM Diagram

Diagram showing different stages of prostate cancer. T1 denotes stage one prostate cancer with small localised cancer. T2 denotes stage two prostate cancer with larger localised cancer. T3 denotes stage three a cancer which has broken out of the prostate capsule and is locally invasive.

Node has two stages: N0 -there are no cancer cells found in the local lymph nodes, and N1 - cancer cells were found in local lymph nodes. Metastasis is similar with M0 indicating the cancer has not spread to other parts of the body, and M1 indicating the cancer has spread to other parts of the body. M1a

indicates the cancer metastasised to lymph nodes outside of the groin; M1b indicates the cancer has metastasised to the bones; and M1c indicates the cancer has spread to another part of the body with or without bone metastasis.

PC often also secretes PSA which typically rises as the disease progresses (Adhyam and Gupta 2012). Using the PSA reading, Gleason score, and the TMN system, the cancer can be assigned a stage. Stages 1-4 of PC are diagnosed at similar rates though regional variations exist in the U.K. with Scotland having the highest Stage IV diagnosis rate at 28% (Figure 1.7)

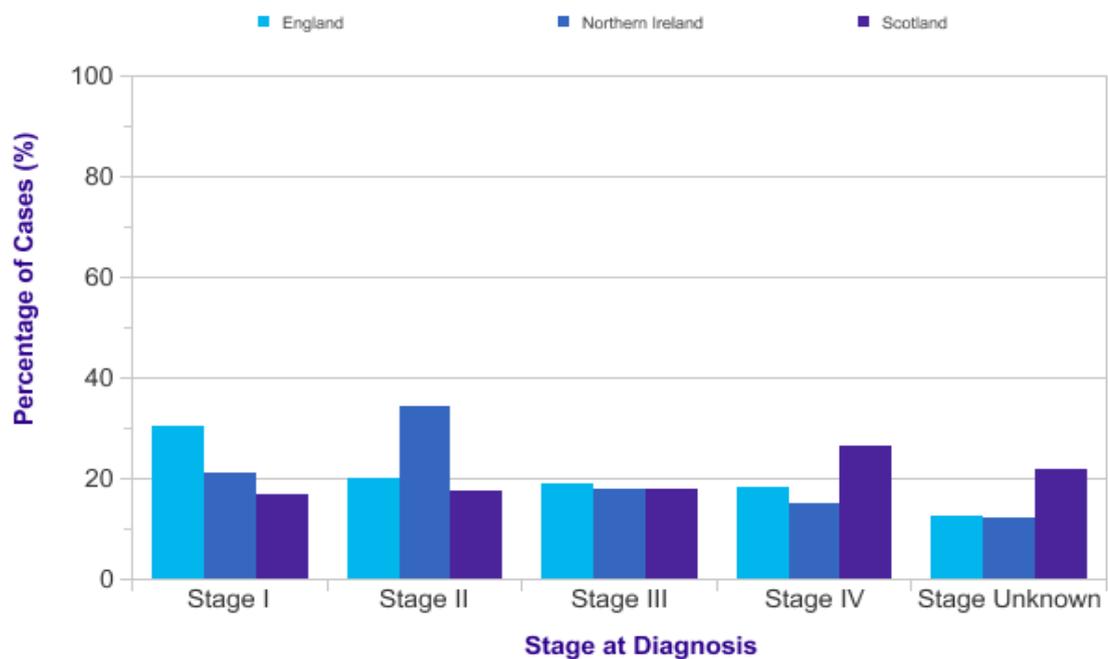


Figure 1.7 - PC stage at diagnosis U.K.

Most stages are diagnosed at similar rates in England, Northern Ireland, and Scotland. Stage IV is diagnosed at the highest rates in Scotland with Stage I being the lowest diagnosed indicating more advanced disease presentation in Scotland. Source: cruk.org/cancerstats

1.2 Prostate Cancer Signaling

1.2.1 PTEN and PI3K/AKT Signaling

Phosphatase and tensin homolog (PTEN) is a well characterised tumour suppressor in many different cancers. It functions as an inhibitor of the phosphatidylinositol 3-kinase (PI3K)/ protein kinase B (AKT) pathway which drives cell survival, growth, migration, and angiogenesis (Osaki, Oshimura *et al.* 2004, Castellano and Downward 2011) (Figure 1.8). The pathway is initially triggered by the activation of a receptor tyrosine kinase (RTK) which dimerises and cross-phosphorylates. A variety of RTK and their ligands have been shown to activate this pathway such as growth factors and metabolic hormones including insulin and melatonin (Auger, Serunian *et al.* 1989, Ruderman, Kapeller *et al.* 1990). Following ligand activation of the receptor, the regulatory subunit of the PI3K heterodimer (p85) binds either directly to the RTK or via a Grb2 adaptor (Holgado-Madruga, Moscatello *et al.* 1997). Binding of p85 induces a conformational change in the PI3K heterodimer whereby the regulatory subunit p85 stops inhibiting the catalytic subunit, p110 (Rordorf-Nikolic, Van Horn *et al.* 1995). Collectively this will form an activated PI3K which can activate a signalling cascade.

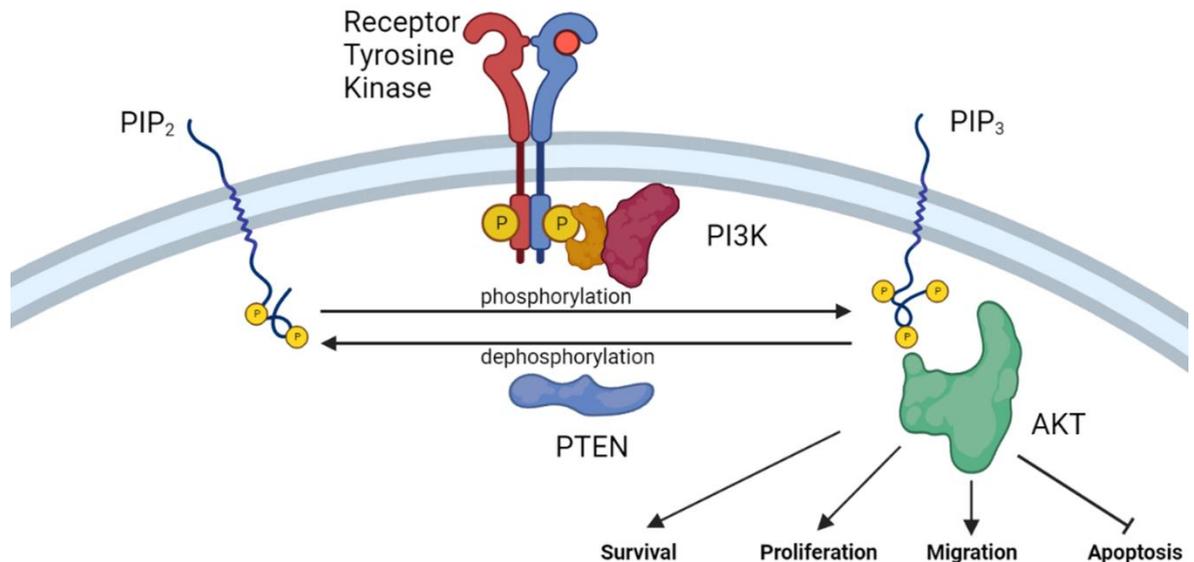


Figure 1.8 PI3K-AKT signalling pathway

Receptor tyrosine kinase binds ligand and dimerises. Following dimerisation PI3K binds intracellular domain of kinase and becomes activate. Active PI3K phosphorylates PIP₂ to PIP₃ which in turn activates AKT. AKT then activates a signalling cascade driving cell survival, proliferation, migration, and inhibition of apoptosis. Created with BioRender.

PI3K can then phosphorylate phosphatidylinositol (PI) 3-4 biphosphate (PIP₂) to form PI 3,4,5 triphosphate (PIP₃) (Rameh, Tolia *et al.* 1997). Inactivate AKT in the cytosol can then bind PIP₃ via its pleckstrin homology domain triggering a conformational change in AKT (Alessi, Deak *et al.* 1997). The resultant conformation changes expose serine 473, which can be phosphorylated from numerous other kinases to trigger the full activation of AKT (Alessi, Andjelkovic *et al.* 1996, Chang, Lee *et al.* 2003, Hanada, Feng *et al.* 2004). AKT has many downstream targets ultimately driving cell survival, cell cycle, migration, and inhibiting apoptosis (Hanada, Feng *et al.* 2004).

PTEN is a well-characterised tumour suppressor as it inhibits this PI3K/AKT pathway by dephosphorylating PIP₃ back to PIP₂ and thus is directly antagonistic to PI3K in preventing AKT activation. Loss of PTEN in PC leads to numerous signalling cascades including a multitude of epigenetic changes which further alter cell growth (Wang, Feng *et al.* 2020, Yang and Yin 2020). Among these

changes, signalling cascades involved in chromatin remodelling have been demonstrated to be capable of driving advanced PC (Ding, Li *et al.* 2019).

1.3 Chromatin Remodelling

1.3.1 BAF complex

Chromatin is made up of nucleosomes of DNA wrapped around a histone octamer. Regulating the organisation of the chromatin allows for epigenetic control of gene expression by regulating DNA accessibility. This can be done by modifying the chromatin itself by post-translational modifications to the histones such as methylation or acetylation of the lysine tails of the histones. Alternatively, chromatin can be physically remodelled by moving the nucleosomes themselves to control access to the DNA. Chromatin remodelling is essential for regulating cellular responses as it allows for dynamic control of gene expression ensuring genes are only expressed when they need to be expressed.

The Brahma (BRM)-related gene 1 (BRG1)-associated factors (BAF) complex is part of the switch-induced/sucrose non-fermentable (SWI/SNF) subfamily which is responsible for chromatin remodelling. This subfamily includes a variety of different chromatin remodellers that use an adenosine triphosphate(ATP)-dependent chromatin remodeller to move nucleosome positions adjacent to DNA to open or close chromatin configuration. The complex is recruited to DNA by protein-protein interactions with a wide range of proteins including transcription factors, chromatin markers, or coactivators and repressors.

Like many SWI/SNF, the BAF complex has many interchangeable subunits, many of which are mutually exclusive and can have antagonistic functions. As such, the various combinations of subunits make a variety of BAF complex each with potentially different and overlapping functions.

The BAF complex requires the integration of an AT-Rich Interacting Domain (ARID) protein for the complex to function (Helming, Wang *et al.* 2014). There are two ARID1 proteins found in eukaryotes, ARID1A and ARID1B. Both share approximately 50% total homology, and 80% homology in their ARID domain (Wu and Roberts 2013). The BAF complex also needs ARID1A/B proteins to allow DNA binding via the ARID domains. Interestingly, the ARID domain has been shown to have no target sequence specificity (Patsialou, Wilsker *et al.* 2005). This would imply that ARID proteins, and the subsequent BAF complex, are localised to different loci via interactions with specific transcription factors which do show DNA binding specificity.

The BAF complex also requires an ATP-dependent chromatin remodelling subunit, which can be BRG1 or BRM (Wang, Côté *et al.* 1996, Wang, Xue *et al.* 1996). These catalytic units utilise ATP hydrolysis to disrupt the DNA-nucleosome interaction to slide the nucleosome and expose parts of the DNA (Strohner, Wachsmuth *et al.* 2005). Both BRG1 and BRM have been reported to be mutated in cancer and can function either as a tumour suppressor or oncogene in a context-dependent manner (Alfert, Moreno *et al.* 2019). Importantly, either BRG1 or BRM must be part of the BAF complex otherwise it has no catalytic activity. As such the different combinations of the BAF complex components creates many different subtypes of BAF complex, each with unique chromatin

remodelling roles and interactions giving rise to further complexity in remodelling dynamics (Figure 1.9).

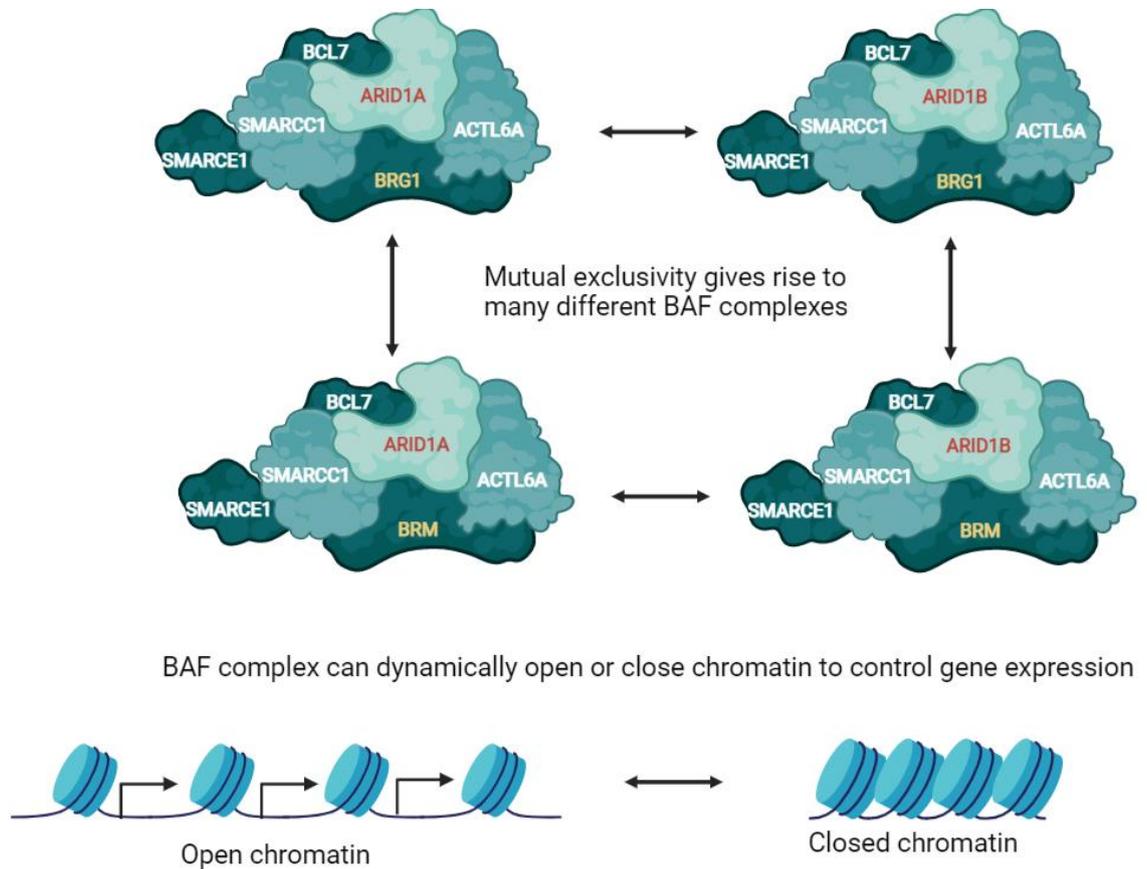


Figure 1.9 - Schematic of BAF complex interchangeability

Due to mutual exclusivity of ARID proteins (ARID1A and ARID1B) and catalytic proteins (BRM and BRG1), these subunits can be combined to create unique BAF complexes. Each of these complexes may have separate functions and targets to open or close chromatin to control gene expression. Created with BioRender.

The BAF complex is highly mutated in cancer as dysregulating chromatin organisation can facilitate oncogenic signalling to enable tumour growth. It is also well reported that the BAF complex is also highly mutated in metastatic PC with 60% of cases having some alteration in at least one component of the complex (Figure 1.10). Out of all the complex components ARID1A is the most mutated with the highest rate of missense and truncation mutations compared

to any other subunit. Consistent with this, ARID1A is also the most mutated epigenetic regulator in cancer (Wu and Roberts 2013).

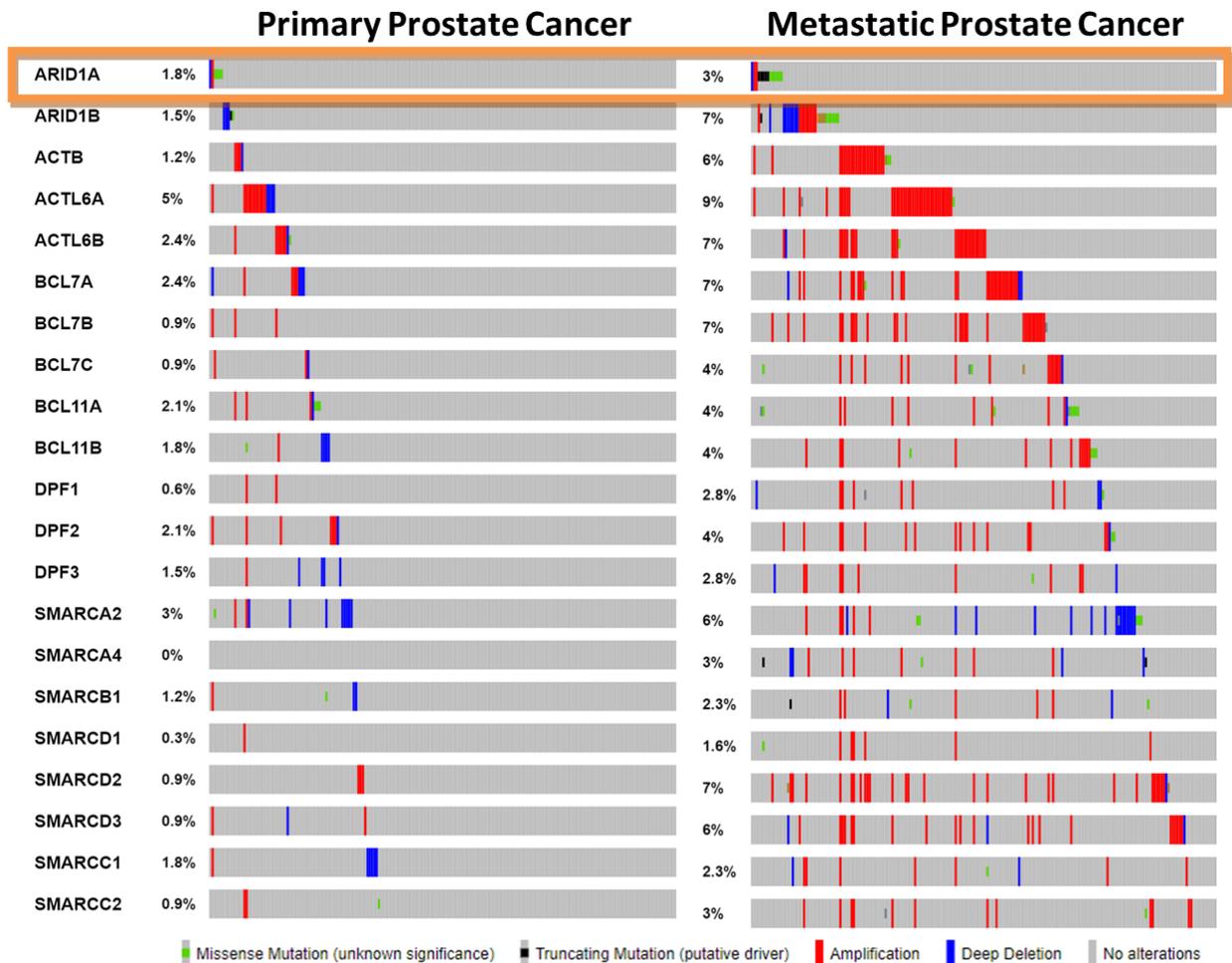


Figure 1.10 – BAF Complex in PC

Data shows every subunit of BAF complex from primary prostate cancer cohort TCGA, Cell 2015 (TCGA network, 2015)(N=334) and metastatic prostate cancer cohort SUC2/PCF Dream Team, PNAS 2019 cohort (Abida, Cyrta *et al.* 2019)(N=444) and visualised on CBioPortal. Key shows genetic alteration and indicated colour. ARID1A highlighted in orange as BAF complex component with most missense or truncation mutations. BRM is referred to as SMARCA2 and BRG1 referred to as SMARCA4.

1.3.2 ARID1A and BAF Complex Dynamics

While ARID1A is frequently mutated and lost in cancer, though primarily gynecological cancers such as in 57% of cases of ovarian clear cell carcinoma (Jones, Wang et al. 2010). It is also crucial to understand the remaining chromatin remodelling complexes when discussing the role of ARID1A in cancer. When a cancer is deficient in ARID1A the remaining BAF complex must incorporate ARID1B (Figure 1.11A). Therefore, it is important to consider the resultant phenotype as not only the loss of ARID1A, but potential gain of ARID1B. This is particularly relevant if when considering that ARID1A and ARID1B are antagonistic in many respects. While many genes are only regulated by ARID1A, many are differentially regulated by both ARID proteins. Therefore, the loss of ARID1A may produce an ARID1B oriented chromatin remodelling phenotype alongside losing regulation of the ARID1A-specific genes (Figure 1.11B).

Research into genome-wide transcription regulation by SWI/SNF are particularly important as they allow us to characterise what loss of a single component may do in a disease setting. One of the most interesting papers on this topic to date investigated the transcriptional interaction of ARID1A, ARID1B and ARID2 (Raab, Resnick *et al.* 2015). ARID1A and ARID1B are BAF components, while ARID2 is the only ARID component of the polybromo-BAF (PBAF) complex. ARID proteins primarily localise to actively transcribed regions of chromatin with specific transcription factors and coregulators. This finding reaffirms that it is the transcription factors that mediate the localisation of the BAF complex. This also leads to an understanding of why certain ARIDs open chromatin at one loci and close at another. It is likely the transcription factor and the coregulators bound to the BAF determine how chromatin structure at a specific loci will change. In

this way, the BAF integrates signals from multiple factors to determine chromatin organisation.

Cancers often have disrupted and aberrant activity of transcription factor activity when compared to healthy tissues. As the BAF complex's localisation is determined by transcription factor interactions, the BAF complex will be localised to different loci in a cancer compared to a healthy tissue. In this way, alterations in the BAF complex subunit which facilitate this interaction may be beneficial to the cancer. For example, if the oestrogen receptor (ER) is upregulated 3-fold in breast cancer tissue, it would be expected that the BAF complex would be mediating open chromatin structures at ER binding sites. The BAF subunits which facilitate this ER-BAF interaction may therefore be upregulated in ER-driven breast cancer.

It has also been shown that ARID proteins can all localise to the same locus but have different functions at these sites (Raab, Resnick et al. 2015). A consistent observation showed that ARID1B and ARID2 appeared to cooperate to repress transcription at specific loci, while ARID1A was responsible to open chromatin at these loci (Raab, Resnick et al. 2015). It could therefore be hypothesised that if a cancer cell were to remove ARID1B/ARID2 we might see upregulation of a gene at that locus due to a change in distribution of SWI/SNF complexes. Conversely, loss of ARID1A would result in a locus that remains shut. If a gene in this locus suppressed tumour growth, then loss of ARID1A could support tumour development. Of course, this observation likely plays out in different ways for many different genes depending on the context of the ARID protein and the available transcription factors and coregulators.

This is likely a crucial reason why mutations in SWI/SNF are not universal across all cancers. As different cancers are driven by different subsets of transcription factors, they bind different ARIDs. While one cancer may find it beneficial to lose ARID1A, another cancer may develop to be dependent on ARID1A as it also mediates the activity of a specific transcription factor. In colorectal cancer, there is a dependency on ARID1A in facilitating both Kirsten rat sarcoma virus (KRAS) signalling and beta-catenin signalling (Vasileiou, Ekici *et al.* 2015, Hiramatsu, Fukuda *et al.* 2019, Sen, Wang *et al.* 2019). However, loss of ARID1A can also facilitate microsatellite instability (Ye, Zhou *et al.* 2014) and dysregulate gene expression sufficiently to drive colorectal cancer in separate models (Mathur, Alver *et al.* 2017).

As the cancer evolves loss ARID1A may give a selective advantage, such as in response to therapy. In breast cancer, ARID1A has been shown to be important in mediating transcription of ER target genes by chromatin organisation at ER binding sites (Xu, Chhangawala *et al.* 2020). However, in response to fulvestrant, an ER degrader, loss of ARID1A gives a survival advantage and independency of ER (Figure 1.11C).

If loss of ARID1A is advantageous from onset, it would explain the high mutation rate in early disease. If ARID1A is beneficial during initial development, it may be lost only in advanced disease.

As loss of ARID1A forces the BAF complex to use ARID1B, the complex is therefore dependent on ARID1B expression. As such ARID1A and ARID1B are synthetically lethal (Helming, Wang *et al.* 2014). A proposed mechanisms for this suggest that the BAF complex forms around the ARID protein and loss of ARID

causes the BAF complex to fall apart. Alternatively, the complex might be able to form, but is unable to interact with transcription factors due to lack of ARID protein. However, the result is the same, with the loss of functional BAF there is a loss in enhancer architecture and likely global transcription dysregulation (Kelso, Porter *et al.* 2017). As such, targeting ARID1B could be seen as a synthetically lethal target in cancers lacking ARID1A.

1.3.3 Genetically Engineered Mouse Models of Prostate Cancer

GEM models of PC have been used extensively over the last three decades to improve our understanding of PC development and advanced disease (Parisotto and Metzger 2013). This is largely done by introducing targeted somatic mutations with the P1 bacteriophage Cre-Recombinase (Parisotto and Metzger 2013). This Cre-recombinase induces recombination of DNA which is flanked by DNA-elements called LoxP sites. By using the transcriptional regulatory elements

of genes which are expressed in the prostate epithelium, Cre-recombinase

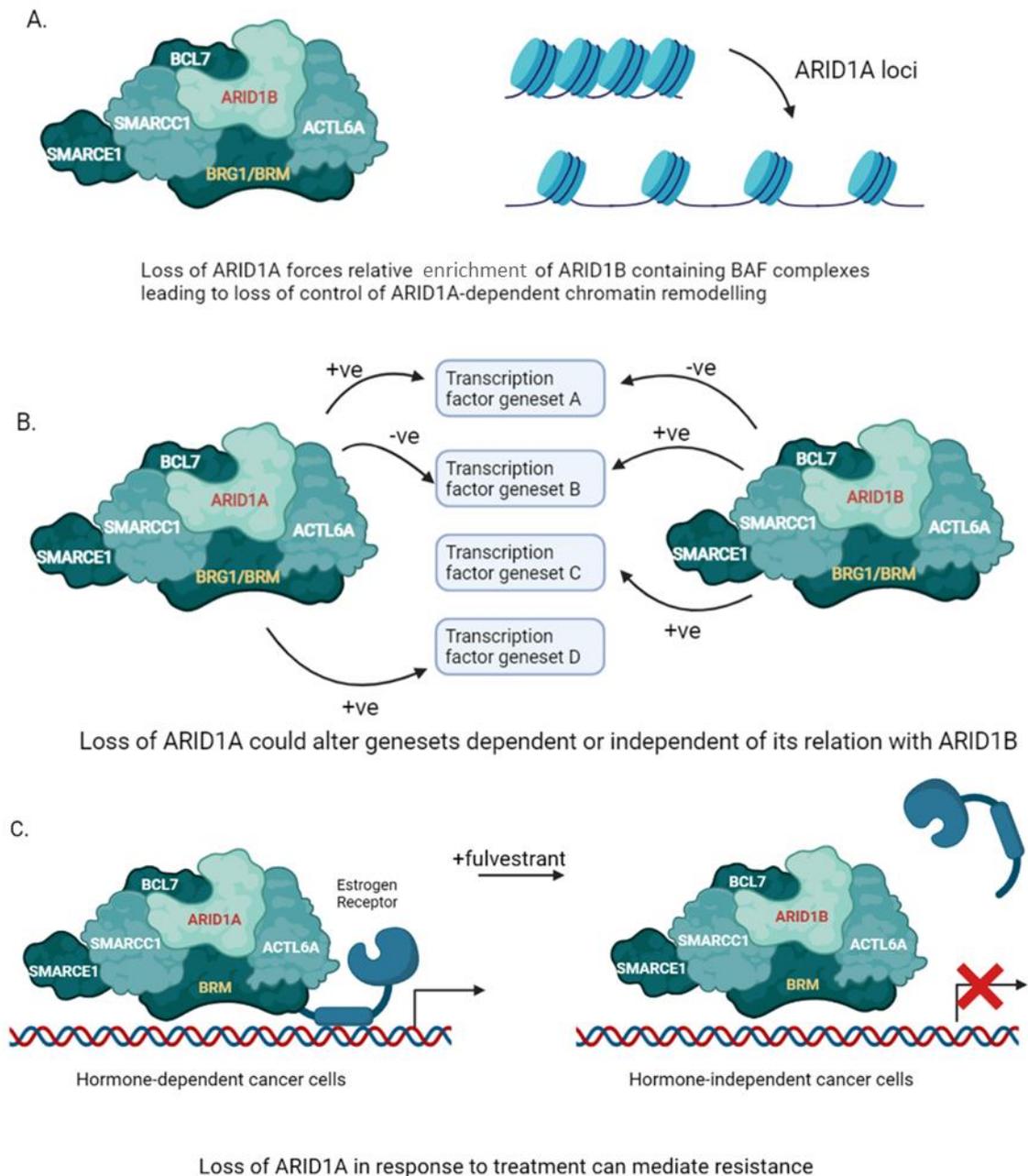


Figure 1.11 ARID1A and ARID1B dynamically alter chromatin organisation

A. Loss of ARID1A leads to loss of regulation of ARID1A dependent foci. **B.** ARID1A and ARID1B can antagonistically or independently regulated many different pathways and genesets. **C.** ARID1A can drive oncogenes such as estrogen receptor, but its loss can be beneficial in treatment resistance. Created with BioRender.

expression can be controlled to allow for tissue specific recombination. One of the first promoters used, was probasin (Pb) which included two androgen response elements which when bound by the androgen receptor (AR) can drive Pb expression (Wu, Wu et al. 2001). Pb is expressed specifically in the mouse

prostate and seminal vesicles which can ensure recombination events are specific to these tissues. Other promoters such as PSA are expressed at different times compared to *Pb* and in subtly different epithelial cell populations (Abdulkadir, Magee et al. 2002). Alternatively, promoters can include inducible elements, such as *NKX3.1^{CreERT2}* which can allow for prostate specific Cre-recombinase expression from the *NKX3.1* promoter following induction by tamoxifen.

There are however morphological differences in the mouse prostate compared to the human prostate. The mouse prostate is made up of 4 major lobes, anterior, dorsal, lateral, and ventral. Despite this, the glands and their organisation are said to highly resemble that of the human prostate, and the development of cancers within these glands follow similar patterns as seen in PC (Cunha, Donjacour et al. 1987, Marker, Donjacour et al. 2003).

Pb-Cre+ve Pten^{fl/fl} is a well characterised PC GEM model. Loss of *Pten* causes the development of large and cystic adenocarcinoma of the anterior prostate. These tumours typically reach a clinical endpoint size within 9 - 12 months, and can also develop lymph node metastasis, though distal metastasis are rare (Wang, Gao et al. 2003). Due to this long latency to endpoint, many PC GEM models use *Pten* loss as a driving mutation to identify other cooperating mutations.

By use of these tools, genes can be knocked out of the mouse prostate genome to allow investigation into their role in PC development. By changing the promoters or adding inducible elements, the cell type and timing of gene knockout can be controlled allowing a spatio-temporal investigation in PC development.

With this system, human epidermal growth factor receptor 2 (HER2), Mothers against decapentaplegic homolog 4 (SMAD4), beta-catenin (CTNNB1) among others have been identified as drivers of PC in the *Pb-Cre+ve Pten^{fl/fl}* model (Ahmad, Patel et al. 2011, Ding, Wu et al. 2011, Patel, Brzezinska et al. 2020). Each of these models shortens the latency and increases metastatic potential of the *Pb-Cre+ve Pten^{fl/fl}* model in different ways. This variability is representative of PC as it closely models a heterogeneous disease with different outcomes based on the genetic aberrations.

1.3.4 Mutagenesis to Study Prostate Cancer

Mutagenesis is an important experimental tool used to introduce mutations to a cell to characterise the resultant phenotype and identify drivers of disease. Currently three major classes of mutagenesis screens are utilised: chemical, forwards genetic, and transposon-mediated.

Chemical mutagenesis is historically the most frequently used type of mutagenesis (Auerbach 1949). This mutagenesis involves exposing cells to a chemical mutagen which will have DNA-damaging properties, such as N-ethyl-N-nitrosourea (ENU) which alkylates DNA or ethyl methanesulphonate (EMS) which reacts and binds to guanine inducing transition mutations (Sega 1984, Brammeld, Petljak *et al.* 2017). These types of mutagens can result in point mutations as well as causing base insertions or deletions which can in turn affect an entire gene by inducing frame-shift mutations (Figure 1.12A).

Forward genetic screens involve inducing gain-of-function or loss-of-function alterations in cells. This is the largest class of mutagenesis screen which encompasses RNA-interference (RNA-I), short-hairpin RNA (shRNA), CRISPR-Cas9,

and open-reading frame (ORF) screens which can mediated alterations (Hu and Zhang 2016). RNA-I, shRNA, and CRISPR-Cas9 can all mediate a downregulation of gene expression. RNA-I and shRNA downregulate RNA by interference leading to degradation of the mRNA transcript. CRISPR-Cas9 downregulates mRNA by knockout of the gene itself by introducing indels (Figure 1.12B). ORF screens can lead to gain-of-function by overexpressing a typically curated set of key genes involved in cell signalling.

Transposon mediated mutagenesis involves the use of a transposon and transposase to randomly insert a transposon into a gene (DeNicola, Karreth *et al.* 2015). Depending on the orientation and location of the transposon inserted, it can either upregulate or downregulated gene expression (Figure 1.12C). This has the advantage over forward genetic screens as it allows the screening of genes being up and downregulated in the same cell. The two most common types of transposons used are *Sleeping Beauty* (SB) and *piggyBac* (PBac) which have differences in insertional bias and integration, with SB preferentially inserting into the gene-body and PBac preferentially inserting into the transcriptional-start-site (TSS). In this study, the SB transposon was used in combination with *Pten* deletion in a murine model to identify drivers of advanced disease.

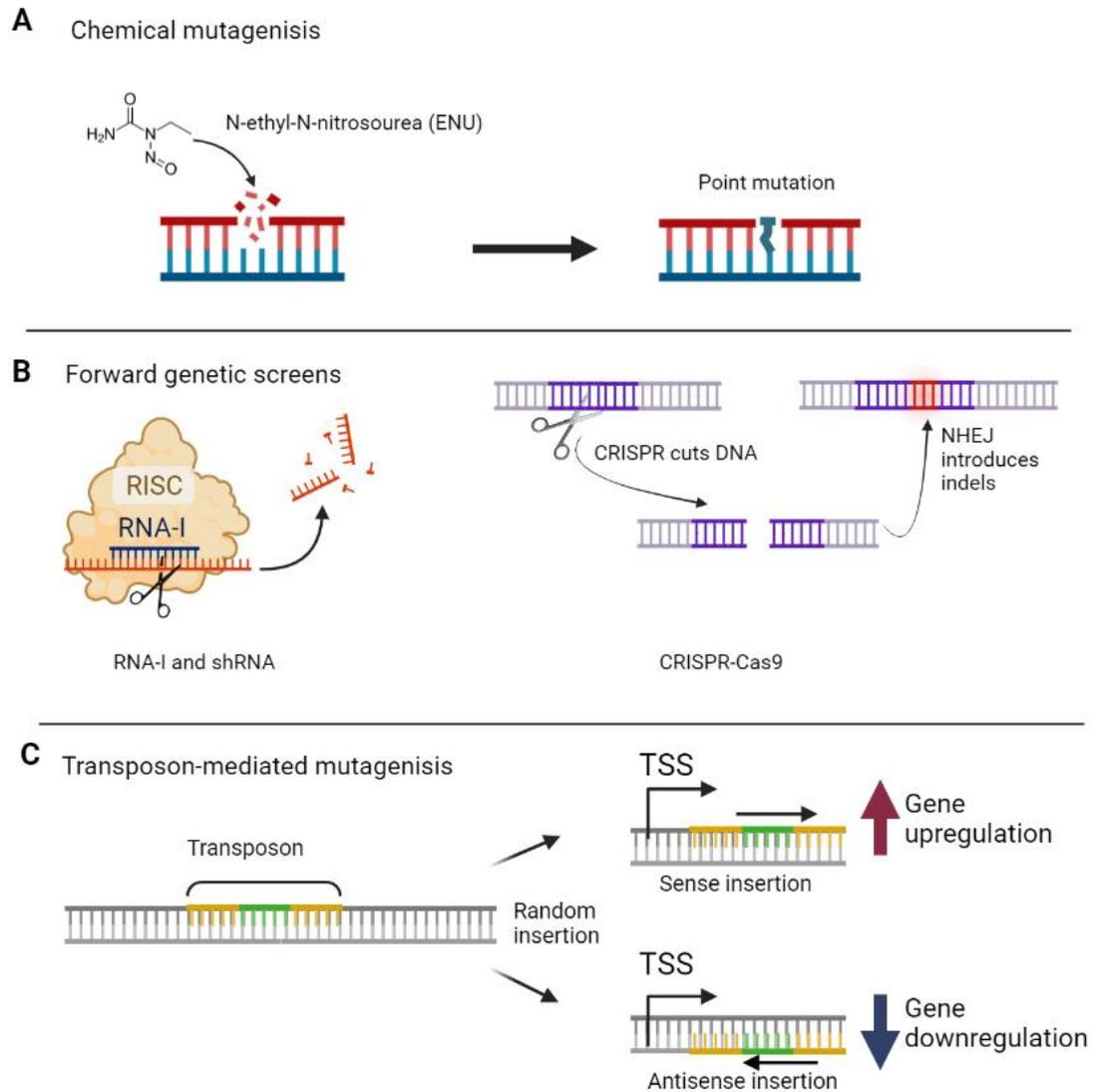


Figure 1.12 - Types of mutagenesis screens

A. Chemical mutagenesis with N-ethyl-N-nitrosourea (ENU) alkylates DNA and can introduce point mutations. **B.** Forward genetic screens alter gene expression, either by RNA-interference leading to mRNA degradation, or by CRISPR-Cas9 cutting DNA leading to indel introduction. **C.** Transposon mediated mutagenesis involves a transposon being inserted randomly into genes. Sense insertions can upregulate gene expression and antisense insertions downregulate gene expression. Created with BioRender.

2 **Materials and Methods**

2.1 *In vivo* materials and methods

2.1.1 Generation of Mice colonies

All experiments were performed under the U.K. Home Office and Animal Welfare and Ethical Review Body (AWERB) guidelines. All mice were maintained under barrier conditions though is not specific pathogen free and given an irradiated standard diet (Harlan) and water *ad libitum*. Mice were routinely checked two times per week and were of a mixed strain. All work performed under project license P5EE22AEE, and personal license ID1C33645.

Both *Pten*^{fl/fl} and *Arid1a*^{fl/fl} were bred onto a Probasin-Cre (Pb-Cre) mouse(Wu, Wu et al. 2001, Wang, Gao et al. 2003, Gao, Tate et al. 2008).

2.1.2 Mouse experiments for chapter 3

To investigate how *Pten* and *Arid1a* genetically interact, *Arid1a*^{fl/fl} was bred onto *Pten*^{fl/+} mice that are well established in the group. Mice were bred to genotypes and cohort sizes as seen in Table 2.1. All genotyping was performed by PCR by Transnetyx (<https://www.transnetyx.com/>). For breeding, only *Pb-Cre* -ve females were used to ensure no unintended recombination events occurred in these mice. *Pb-Cre*+ve *Pten*^{fl/fl} mice were also not used for breeding.

Mice were aged until clinical endpoint; tumour >1.2cm, blood in urine, rapid weight loss, and bladder distention. Mice which did not develop tumours or where tumours did not reach clinical endpoint by the 18-month aging endpoint month were humanely culled. Upon reaching clinical or aging endpoint, mice were humanely culled by CO₂ inhalation. Cervical dislocation was then used as secondary method to confirm the cull.

<i>Pten</i> Status	<i>Arid1a</i> Status	Cohort Size
+/+	<i>fl/fl</i>	10
<i>fl/+</i>	<i>fl/+</i>	20
<i>fl/+</i>	<i>fl/fl</i>	20
<i>fl/fl</i>	<i>fl/+</i>	20
<i>fl/fl</i>	<i>fl/fl</i>	20

Table 2.1 - Mouse genotypes and cohort sizes

All genotypes would be generated to be *Probasin-Cre recombinase* positive to allow for recombination of genes.

2.1.3 Tissue Isolation

Following culling of cohort mice after clinical or aging endpoint; prostate, seminal vesicle, epifat (abdominal fat attached to testis which associates with prostate) and testis, kidneys, spleen, pelvic and mesenteric lymph nodes, spleen, liver, lung, and heart were all removed and fixed in 4% formalin for 48 hours at room temperature (RT). Following formalin fixation samples were processed and embedded in paraffin blocks.

2.1.4 Immunohistochemistry

All Haematoxylin & Eosin (H&E) and immunohistochemistry (IHC) staining was performed on 4µm formalin fixed paraffin embedded sections (FFPE) which had previously been ovened at 60°C for 2 hours.

The following antibodies were stained on an Agilent AutostainerLink48, pAKT (4060, Cell Signaling) and p53 (NCL-L-p53CM5p, Leica). Sections were loaded

into an Agilent pre-treatment module to be dewaxed and undergo heat induced epitope retrieval (HIER) with high pH target retrieval solution (TRS) (K8004, Agilent). All sections were heated to 97°C for 20 minutes in the high TRS. After HIER all sections were rinsed in flex wash buffer (K8007, Agilent) prior to being loaded onto the autostainer. The sections underwent peroxidase blocking (S2024, Agilent) for 5 minutes and washed with flex buffer. The specific primary antibody was applied at a previously optimised dilution for 35 minutes (pAKT, 1/45; p53, 1/750). The sections were washed with flex wash buffer before application of rabbit secondary antibody (K4003, Agilent) for 30 minutes. Sections were rinsed with flex wash buffer. All sections had Liquid DAB (K3468, Agilent) for 10 minutes. The sections were then washed in water and counterstained with haematoxylin z (RBA-4201-00A, CellPath).

The following antibodies were stained on a Leica Bond Rx autostainer, ARID1a (12354, Cell Signaling), Col1a1 (93668, Cell Signaling), Ki67 (12202, Cell Signaling), p21 (ab107099, Abcam), PH2AX (9718, Cell Signaling) and PTEN (9559, Cell Signaling). All FFPE sections underwent on-board dewaxing (AR9222, Leica) and antigen retrieval using appropriate retrieval solution. Sections for Col1a1 were retrieved using ER1 retrieval method (AR9661, Leica) for 40 minutes at 100°C. Sections for ARID1A, Ki67, p21, PHa2x and PTEN were retrieved using ER2 (AR9640, Leica) for 20 minutes at 100°C except PHa2X which was with ER2 for 10 minutes. Sections were rinsed with Leica wash buffer (AR9590, Leica) before peroxidase block was performed using an Intense R kit (DS9263, Leica). Sections for p21 had blocking solution applied from the Rat ImmPRESS kit (MP-7404, Vector Labs) for 20 minutes. Sections were rinsed with wash buffer and then the primary antibody applied at the optimal dilution (ARID1A, 1/200; Col1a1, 1/200;

Ki67, 1/1000; p21, 1/150; PHa2x, 1/120, PTEN, 1/70). The sections were rinsed with wash buffer and appropriate secondary antibody applied for 30 minutes. ARID1A, Col1a1, Ki67, PHa2x and PTEN had Rabbit EnVision applied. Sections for p21 had Rat ImmPRESS secondary solution applied. The sections were rinsed with wash buffer, visualised using DAB and then counterstained with haematoxylin in the Intense R kit.

H&E staining was performed on a Leica autostainer (ST5020). Sections were dewaxed, taken through graded alcohols and then stained with Haem Z (RBA-4201-00A, CellPath) for 13 mins. Sections were washed in water, differentiated in 1% acid alcohol, washed and the nuclei blu'd in Scott's tap water substitute (in-house). After washing with tap water sections were placed in Putt's Eosin (in-house) for 3 minutes.

To complete H&E, IHC, and ISH staining sections were rinsed in tap water, dehydrated through graded ethanol's and placed in xylene. The stained sections were coverslipped in xylene using DPX mountant (SEA-1300-00A, CellPath). Reagents and antibody information is shown in Table 2.2 and 2.3 respectively.

Reagent	Company	Code
Antibody diluent	Agilent	S2023
Flex Wash Buffer	Agilent	K8007
High pH Target Retrieval Solution (TRS)	Agilent	K8004
Liquid DAB	Agilent	K3468
Peroxidase Block	Agilent	S2023
Rabbit EnVision	Agilent	K4003
DPX mountant for microscopy	CellPath	SEA-1300-00A
Haematoxylin Z	CellPath	RBA-4201-00A
Bond Diluent	Leica	AR9352
Bond Wash	Leica	AR9590
Dewax Solution	Leica	AR9222

Epitope Retrieval solution 1 (ER1)	Leica	AR9661
Epitope Retrieval solution 2 (ER2)	Leica	AR9640
Intense R Kit	Leica	DS9263
Rat ImmPRESS kit	Vector Labs	MP-7404

Table 2.2 - Immunohistochemistry Reagents

Antibody	Clone	Company	Autostainer	Retrieval	Dilution	Secondary Ab
ARID1A	D2A8U	Cell Signaling	Leica Bond Rx	ER2 20mins	1/200	Rabbit EnVision
COL1A1	E8F4L	Cell Signaling	Leica Bond Rx	ER1 40 mins	1/200	Rabbit EnVision
Ki67	D3B5	Cell Signaling	Leica Bond Rx	ER2 20 mins	1/1000	Rabbit EnVision
p21	Hugo291	Abcam	Leica Bond Rx	ER2 20 mins	1/150	Rat ImmPRESS
p53	CM5	Leica	Dako Link48	High TRS	1/750	Rabbit EnVision
pAKT	Ser473	Cell Signaling	Dako Link48	High TRS	1/45	Rabbit EnVision
PH2AX	20E3	Cell Signaling	Leica Bond Rx	ER2 10 mins	1/120	Rabbit EnVision
PTEN	138G6	Cell Signaling	Leica Bond Rx	ER2 20 mins	1/70	Rabbit EnVision

Table 2.3 - Immunohistochemistry antibodies

2.1.5 Halo Scoring

Slides of stained tissue were scanned, and digital images were analysed using HALO Image analysis Platform to quantify stain intensity and percentage of cells positive for the stain. The software was trained to classify tissue as epithelium, stroma, or background based on morphological characteristics (Figure 2.1). Only epithelium compartments were analysed as these constitute the tumour. The software could then allocate a score to each cell. Histoscore determined by formula: (% cells low intensity)+2(%cells medium intensity)+3(%cells high intensity)=Histoscore.

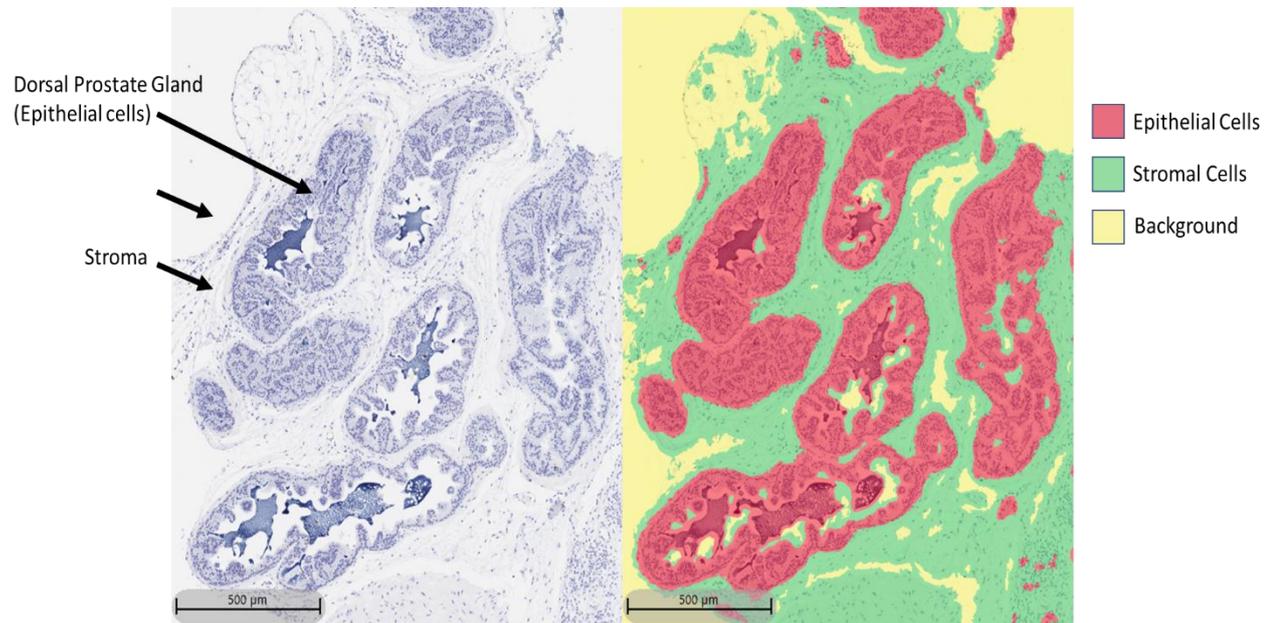


Figure 2.1 - Halo Classification to Identify Cell Types

Halo image analysis software is trained to identify epithelial cells, stromal cells, and separate those from each other and background.

2.2 Sleeping Beauty Screen

2.2.1 Generating Sleeping Beauty Mouse

T2/Onc3 transposon and $Rosa^{26Lox66SBLox71/+}$ transposase made up the Sleeping Beauty system. When both of these components are present in a cell, the transposon can be moved by the transposase and randomly inserted into the genome and gene expression can be randomly disrupted depending on location and orientation of insertion. This SB system was bred onto the Pb-Cre+ve $Pten^{fl/fl}$ mouse line to produce Pten deficient mice who express the SB system in their prostate which is capable of transposon-mediated mutagenesis. These mice were confirmed to contain the relevant genotypes by PCR by Transnetyx. Mice were all of a mixed background and aged to clinical endpoint as discussed in

section 2.1.2. Prostates were divided during dissection and placed into either formalin for overnight fixation prior to paraffin embedding or snap frozen at -80C prior to sequencing.

2.2.2 Common Insertion Site (CIS) analysis

Transposon integration was identified from genomic DNA using two rounds of linker-mediated PCR with unique indexes included for each individual tumour to allow for multiplex sequencing. Redundant sequences, as well as insertions in the *En2* gene and in the T2/Onc3 donor itself were removed from analysis and remaining CIS were identified using a Gaussian-Kernel correlation framework. Reads from sequenced tumours were then mapped to mouse genome to identify insertion sites. CISs predicted across multiple scales and overlapping in their genomic locations were clustered together such that the CIS with the smallest genomic footprint was reported as the representative CIS. For highly significant CISs with narrow spatial distribution of insertions sites, 15K kernel is typically the scale on which CISs are identified.

2.3 Cell culture

2.3.1 Cell lines

Prostate cancer cell lines were purchased from American Type Culture Collection (ATCC). These cells were maintained in a tissue culture class 1 incubator at 37° C and 5% CO₂. All cell lines were routinely tested every 6-months for mycoplasma using MycoAlert Mycoplasma Detection Kit (Lonza) according to manufacturer's instructions.

DU145, PC3, LNCaP, CWR22, were grown in RPMI-1640 (Sigma Aldrich) and supplemented with 1% L-Glutamine (Gibco) and 10% fetal bovine serum (FBS) (Sigma Aldrich). This media was used for most cell lines in most instances so will be referred to as standard culture media (SCM). 22Rv1 cell line was grown in RPMI-1640 (Sigma Aldrich) and supplemented with 1% L-Glutamine (Gibco) and with charcoal-stripped FBS (Thermo Fisher) to remove lipophilic materials such as androgen. RWPE was grown in Keratinocyte serum free media (Thermo Fisher) supplemented with human recombinant epidermal growth factor (rEGF) and bovine pituitary extract (BPE) as supplied, 10% FBS (Gibco), and 1% L-Glutamine (Gibco).

All cell lines were regularly passaged between 3 - 4 days at around 70% confluency as appropriate to the growth rate of the cell line. To passage, media was aspirated, cells were then washed with phosphate buffered saline (PBS) and PBS was aspirated, 1X trypsin (Gibco) was added for 3 minutes as cells were returned to incubator for cells to detach. Fresh medium was then added to neutralise trypsin and generate a cell suspension. Cells count could then be counted using a CASY counter (Innovatis) or CellDrop (DeNovix) and seeded as required.

2.3.2 Cryopreservation of cells

Cell lines were suspended as described in Cell Culture 2.2.1. Cells were counted in CASY counter (Innovatis) and appropriate number ($1-3 \times 10^6$ depending on cell line) were centrifuged at 300G for 5 minutes. Cell pellet was then resuspended in 90% FBS (Gibco) and 10% dimethyl sulphoxide (DMSO)(Fisher). 1ml of suspension was transferred to cryotubes (Nunc) and frozen overnight at -80C.

The next day, cells were then moved on dry ice to liquid nitrogen storage for long-term storage.

2.3.3 *In vitro* proliferation assay as measured by cell count

DU145, PC3, LNCaP, DU145 EV1, DU145 EV2, DU145 KO2, DU145 KO4, DU145 KOPOOL1 were seeded in 6-well plates (Corning) with 5×10^4 cells per well.

Where applicable, cells were transfected to knockdown gene expression using reverse transfection protocol detailed in section 2.2.7 After 72 hours the cells were counted using CellDrop (DeNovix) counter.

2.3.4 *In vitro* colony forming assay

Cells were seeded into 6-well plate (Corning) with SCM at low number to allow individual cells to form colonies. For PC3, DU145, and DU145 clonal derivatives, 200 cells were seeded. For LNCaP 600 cells were seeded. Cells were seeded in duplicate for each condition. Cells were left and allowed to grow for between 2-3 weeks until colonies were formed. Media was removed and cells were washed once with PBS. Cells were fixed in the plate with 1ml methanol for 30 minutes followed by staining with 0.5ml 0.25% crystal violet (Thermo Fisher) for 10 minutes. Cells were then washed with water until water ran clear. Plates were then left to air-dry overnight. Plates could then be imaged on LI-COR (LI-COR Biosciences) with 700nm wavelength.

2.3.5 *In vitro* scratch wound healing assay

1×10^5 cells per well were seeded with SCM into Essen Bioscience 96 well ImageLock Microplates and left overnight. After cell were observed to be between 90-100% confluency, the WoundMaker (Essen Bioscience) was used to

scratch 700-800 μ M wounds. Cells were then washed once with PBS and SCM was then added. Cells were then placed into Incucyte Zoom. Images were taken of cells hourly and recorded until wound closure.

2.3.6 *In vitro* transwell invasion assay

Cells were seeded in 6-well plate for 24h in RPMI-1640 (Sigma Aldrich) supplemented with 1% L-Glutamine (Gibco) without any FBS. 0.5ml and 1X Trypsin (Gibco) was then added and cells were then resuspended in 5ml SCM to deactivate trypsin. Cells were centrifuged at 300G for 5 minutes and resuspended in 3ml serum-free media. Cells were counted using CellDrop (DeNovix) counter. 750 μ l of SCM was added to bottom of 24-well plate and Boyden Chamber (Merck) added to each well. 5×10^4 cells suspended in 500 μ l of serum-free media was then added to each chamber. This creates a gradient of nutrients acting as a chemoattractant making cells migrate down and adhere to the lower surface of the Boyden chamber membrane.

Cells were incubated for 48 hours, media was then aspirated before fixing each chamber at -20° C in 0.5ml 100% methanol for 30 minutes. Methanol was then removed, and chambers were fixed in filtered with 0.5ml haematoxylin for 30 minutes. Cells were then washed three times with 0.5ml distilled water and inner membrane was scrubbed with a cotton bud to remove cells attached to upper membrane. Chambers were inverted and air dried for 15 minutes. Once dried, membrane was cut out with a scalpel and mounted onto a slide with DPX mountant (SEA-1300-00A, CellPath). These were left to dry overnight, and 5 images taken at 20X on the Zeiss AX10. Images were then quantified using ImageJ.

2.3.7 Reverse transfection of siRNA for transient gene knockdown

Lipofectamine RNAiMAX (Thermo Fisher) was used as per manufacturer's instructions.

9 μ l of lipofectamine RNAiMAX was combined with 150 μ l of Opti-MEM (Gibco). 3 μ l of siRNAs at 10 μ M were also combined with 150 μ l of Opti-MEM. These solutions were then combined and incubated at room temperature for 5 minutes. 250 μ l of this solution was then added to freshly seeded cells. After 24h, the media on these cells was aspirated and replaced with fresh media.

2.3.8 Generating stable ARID1A knockout clones in human cell lines

Stable ARID1A knock out (KO) clones were generated in DU145 using CRISPR/Cas9 plasmid with a specific guide RNA to the ARID1A sequence (Santa Cruz, sc-400469) and a homology directed repair (HDR) plasmid (Santa Cruz, sc-400469-HDR). Amaxa Cell Line Nucleofector Kit (Lonza) was used for electroporating the cells to allow the plasmids into the cell using electroporation. Setting A023 and nucleofector kit L was used for DU145. Setting T009 and nucleofector kit R was used for LNCaP. Setting T013 and nucleofector kit V was used for PC3 though no stable clones could be generated for these lines. A CRISPR/Cas9 control plasmid with a non-specific guide RNA (Santa Cruz, sc-418922) and an in-house Infra-Red Fluorescent Protein (IRFP) plasmid were used as a control with puromycin as a selectable marker on the IRFP plasmid. Following electroporation, cells were seeded in 2cm dishes and media was changed after 24h. The next day, cells were split into 10cm dishes in media containing puromycin (Sigma Aldrich) at 2 μ g/ μ l. The media was regularly changed every 2-3 days over approximately 14

days with continuous selection by puromycin until individual colonies could be picked. Picked clones were then screened by western blotting and QRT-PCR to confirm KO of ARID1A. DU145 KO2 and DU145 KO4 were selected for having the lowest levels of ARID1A. Clones with a reduction but not complete loss of ARID1A were pooled to form the DU145 KO POOL line. DU145 EV1 and EV2 were similarly derived from individual control clones which retained wildtype ARID1A. DU145 EV1, EV2, KO2, KO4, and KOPOOL were then used in subsequent experiments.

2.4 Protein Analysis

2.4.1 Protein Extraction

Protein was extracted from *in vitro* samples using standard method using 1% SDS lysis buffer (1% SDS made up with Milli-Q water) containing protease inhibitor (1mM phenylmethylsulfonyl fluoride (PMSF)) and phosphatase inhibitor (phosphoSTOP (Roche)). Samples were then sonicated for 20-30 seconds to reduce viscosity prior to quantification.

2.4.2 Protein Quantification

To estimate protein concentration, the Pierce BCA Protein Assay Kit (Thermo Fisher) was used in 96-well microplate (Corning) setup. Bovine serum albumin (BSA)(Sigma Aldrich) standard also found in kit is used as the standard control for quantification with a range of 0-2000ug/mL. BSA standard is diluted in 1% SDS lysis buffer.

10ul of sample or standard was combined in microplate (Corning) with 200ul of BCA working reagent (50:1 solutions A:B from kit). Each of the samples and standards were plated in triplicate and incubated at 37°C for 30 minutes prior to

measuring absorbance at 562nm on a SpectraMax ABS Plus microplate reader (Molecular Devices). BSA standard can then be plotted to determine protein concentration in sample.

2.4.3 Western electrophoresis and immunoblotting

Prior to Western electrophoresis, protein samples quantified from 2.3.2 were diluted to reach same concentration typically around 1µg/µl with 4X NuPAGE LDS sample loading buffer (Thermo Fisher) and 10 X NuPAGE Sample Reducing Agent (Thermo Fisher). Samples made up to total volume with Milli-Q water. Following dilution, samples were boiled at 100°C for 5 minutes. Samples were then centrifuged and loaded onto 3-8% NuPAGE Tris-Acetate Protein Gels (Thermo Fisher) and ran at 150V for 70 minutes in 1X NuPAGE Tris-Acetate running buffer (Thermo Fisher).

Immobilion polyvinylidene fluoride (PVDF) (Millipore) was soaked in methanol for 2 minutes to activate. Gels were then transferred to the PVDF in transfer buffer with 10% methanol by running at 30V for 120 minutes. PVDF blots were then blocked in 5% milk for 60 minutes and incubated with primary antibody overnight (see dilutions in table 2.4). Following overnight incubation, blots were washed in TBST three times for 5 minutes before being incubated with the appropriate secondary antibody with conjugated HRP for 60 minutes at room temperature. Blots were then washed again in TBST three times for 5 minutes before being visualised with Pierce enhanced chemiluminescence (ECL) Plus (Thermo Fisher) on a ChemiDoc (Biorad).

Antibody	Manufacturer	Catalog No.	Dilution
ARID1A	Cell Signalling	#12354	1/1000
ARID1B	Cell Signalling	#92964	1/1000

AR	Santa Cruz	N-20 , sc-816	1/1000
BRG1	Abcam	ab4081	1/1000
HSC70	Abcam	ab154415	1/1000
PAKT	Cell Signalling	#9272	1/1000
PTEN	Cell Signalling	#9559	1/1000
Anti-mouse HRP	Cell Signalling	#7076	1/400
Anti-rabbit HRP	Cell Signalling	#7074	1/400

Table 2.4 - Immunoblotting antibodies and dilutions

2.5 RNA Analysis

2.5.1 RNA extraction

RNA was extracted from cell lines grown in 6-well plate or from cell pellet using RNAeasy Mini Kit (Qiagen) as per manufacturer's instructions. The optional step to remove genomic DNA using RNase-free DNase (Qiagen) was also included in all samples. RNA was eluted in final step of extraction into RNase-free molecular grade water and quantified by Nanodrop (Thermo Fisher).

2.5.2 Reverse transcription of RNA to produce cDNA

2µg of extracted RNA was used in the High Capacity cDNA Transcription Kit (Applied Biosystems) according to manufacturer's instructions. Once the mastermix of RNA, RT-buffer, dNTP, random primers, and Multiscribe Reverse transcriptase was made up, the reaction volume was made up to 20µl using RNase free water. This mix was then placed in a thermocycler (MJ Research - PTC200) with the program: 25 °C for 10 minutes, 37 °C for 120 minutes, 85 °C for 5 minutes, before remaining at 4 °C until the samples are removed. DNA samples could then be stored at 4 °C, -20 °C, or -80 °C depending on immediacy of use.

2.5.3 Quantitative real-time polymerase chain reaction (QRT-PCR)

Roche universal probe library was used for all primer design (Table 2.5). QRT-PCR was performed using a 96-well plate with a 20µl reaction volume made up of 10µl 2X Taqman universal PCR Master Mix (Applied Biosystems), 0.2µl Universal library probe as per primer requirement (Roche), 0.2µl of forward and reverse primers respectively, 2µl cDNA (routinely diluted (1:10)), and 7.4µl nuclease free water. The 96-well plate was then sealed with microamp optical adhesive film (Thermo Fisher) and centrifuged prior to thermal cycling. The plate was then inserted into ThermoFisher Scientific Quant Studio 3 Real- Time PCR system and ran on following conditions: 50°C for 2 minutes, 95°C for 10 minutes, followed by 40 cycles between 95°C for 15 minutes and 60C for 1 minute. *ROX* was used as a passive reference and *CASC3* gene was used as the endogenous reference as a housekeeping gene for all genes examined.

Primers	Forward	Reverse	Probe
Human			
<i>ARID1A</i>	CCAACAAAGGAGGCACCA	CCATTGGACTGGATGACTGA	32
<i>CASC3</i>	CACCTCCTCATCTGTATCCTAACA	CTGGGCGGGTTATAGTAAGT	8
Mouse			
<i>Arid1a</i>	CCAACAAAGGAGGCACCA	CCATTGGACTGGATGACTGA	32
<i>Axin2</i>	GAGAGTGAGCGGCAGAGC	CGGCTGACTCGTTCTCCT	16
<i>Casc3</i>	AAGAAACAGGTGGCTGCCTT	GGGTTGTCCAGTCTCTGCTC	27
<i>Ccnd1</i>	TTTCTTTCCAGAGTCATCAAGTGT	TGACTCCAGAAGGGCTTCAA	72
<i>Cd44</i>	TCCTTCTTTATCCGGAGCAC	CCTGGAGTCCTTGGATGAGT	49
<i>Dkk1</i>	CCGGGAAGTACTGCAAAAAT	CCAAGGTTTTCAATGATGCTT	76
<i>Myc</i>	CCTAGTGCTGCATGAGGAGA	TCCACAGACACCACATCAATTT	77
<i>Tiam1</i>	GGAATATTTGATGACTGTTCCA	GGTGGACTGGGTAAGACC	7

Table 2.5 - QRT-PCR primer sequences. Probe column indicates which Roche Universal Probe was used with which primer set for detection of cDNA by QPCR using Taqman.

2.6 Omics techniques

2.6.1 RNA-Sequencing

For mouse prostate samples the prostate tumour was used as tissue of interest for RNA-Seq which results in inevitable comparison of anterior vs dorsolateral prostate tumours in instances where GEM model causes these different tumours to emerge. In the murine RNA-Seq anterior lobes were the primary lobe for *Pb-Cre+ve Pten^{fl/fl} Arid1a^{+/+}* and *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/+}*, while dorsolateral lobes were the primary lobe of *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* mice. RNA was extracted from these mouse prostate samples or DU145 clones using Qiagen kit as described in 2.4.1 including optional DNase stepRNA sequencing (RNA-Seq) with poly-A-tailed mRNA selection-based RNA-seq library preparation was performed by our in house Molecular Technology Services.

For RNA quality control 5µl of 50ng/µl per sample was loaded into tapestation. 1ug for each sample was used in library preparation. Quality and quantity of the libraries were assessed on an Agilent 2100 bioanalysed and Qubit (Thermo Fisher Scientific). The libraries were subsequently run on an Illumina Next Seq 500 using the High Output 75 cycle kit (2 x 36 cycles, paired-end reads, single index) at 50 million reads. Three biological replicates were used for mouse prostate samples and five biological replicates for DU145 clones. Quality checks on the raw RNA-seq data files were performed by Mr William Clark (Core Sequencing Services, Beatson Institute) using fastqc (<http://bioinformatics.barbraham.ac.uk/projects/fastqc/>). Analysis of the RNAseq datasets was carried out by Mr Robin Shaw (Informatician, Beatson Institute).

Metacore was used to identify pathway or network alterations in experimental groups with a fold change as indicated and an adjusted P-value <0.05.

2.6.2 ChIP-Seq

2.6.2.1 ChIP

ChIP was performed using the SimpleChIP Enzymatic Chromatin IP Kit with Magnetic Beads (Cell Signalling). The protocol was followed according to manufacturer's instructions, with an input of 25mg of murine prostate tumour from *Pb-Cre+ve Pten^{fl/fl}* mice leading to prostate tissue with an anterior origin. sample homogenised on ice for 5-10 minutes with a dounce homogeniser (Corning). Step III Nuclei preparation and Chromatin digestion included an optimisation step recommended by manufacturer's instruction to ensure appropriate digestion of chromatin by micrococcal nuclease to give a range of DNA fragments. These steps were included, testing 0µl, 0.25µl, 0.5 µl, 0.75 µl, and 1 µl of the micrococcal nuclease. This optimisation showed that 0.5 µl performed best, giving a range of DNA fragments without over digestion. Samples were incubated with micrococcal nuclease as indicated, then samples were run on a High Sensitivity DNA Chip (Agilent) on a Bioanalyser (Agilent) to measure sizes of DNA fragments.

All other steps were followed as recommended. The ARID1A D2A8U (Cell Signalling) antibody was used for the pulldown of ARID1A. The Histone H4 2592 (Cell Signalling) antibody was used for the pulldown of Histone H4 which functioned as a positive control. No antibody was included in negative control samples. QRT-PCR was performed to quantify relative amount of DNA from each pulldown using SimpleChIP Universal qPCR Master Mix (Cell Signalling).

2.6.2.2 ChIP library preparation

Libraries were prepared using NEBNext Multiplex Oligos for Illumina Index Primer Set 1 (New England Biolabs) as the index primers, using the NEBNext Ultra II DNA Library Prep Kit for Illumina (New England Biolabs) library preparation kit.

Samples were initially quantified using a Qubit (ThermoFisher Scientific) with the Qubit dsDNA high-sensitivity assay kit (ThermoFisher Scientific). From here, 500pg of DNA from each ChIP was used in library preparation. All steps were included as necessary, and the kits were used following manufacturer's instructions. As we had smaller yields for ARID1A compared to Histone H4 pulldowns, we did not use the optional clean up steps to avoid library loss as was recommended. Indexes used for each sample are indicated below in table 2.6.

Mouse ID	Antibody	Index	Index Sequence
HPAA 32.2A	Histone H3	1	5'- CAAGCAGAAGACGGCATAACGAGATCGTGATGTGACTG GAGTTCAGACGTGTGCTCTTCCGATC-s-T-3'
HPAA 32.2A	ARID1A	2	5'- CAAGCAGAAGACGGCATAACGAGATACATCGGTGACTG GAGTTCAGACGTGTGCTCTTCCGATC-s-T-3'
HPAA 32.2D	Histone H3	3	5'- CAAGCAGAAGACGGCATAACGAGATGCCTAAGTGACTG GAGTTCAGACGTGTGCTCTTCCGATC-s-T-3'

HPAA 32.2D	ARID1A	4	5'- CAAGCAGAAGACGGCATACGAGATTGGTCAGTGACTG GAGTTCAGACGTGTGCTCTTCCGATC-s-T-3'
HPAA 33.3E	Histone H3	5	5'- CAAGCAGAAGACGGCATACGAGATCACTGTGTGACTG GAGTTCAGACGTGTGCTCTTCCGATC-s-T-3'
HPAA 33.3E	ARID1A	6	5'- CAAGCAGAAGACGGCATACGAGATATTGGCGTGACTG GAGTTCAGACGTGTGCTCTTCCGATC-s-T-3'

Table 2.6 - Indexes for ChIP samples. Index sequences are ligated to DNA libraries and indicate a 'barcode' used to identify samples uniquely as part of a particular library. This allows multiple libraries to be 'pooled' and sequenced together and then separated during downstream analysis.

2.6.2.3 ChIP Sequencing

Following library preparation samples were prepared and sequenced on NextSeq 2000 (Illumina) with 30million 2x100bp paired end reads by Glasgow Polyomics' next generation sequencing sequencing and transcriptomics service.

2.6.3 ATAC-Seq

The protocol for Omni-ATAC Sequencing was adapted from ATAC seq protocol by (Buenrostro, Giresi *et al.* 2013). The protocol for Omni-ATAC sequencing comprises of 9 steps summarised in Figure 2.2. Reagents are summarised in Table 4.2.

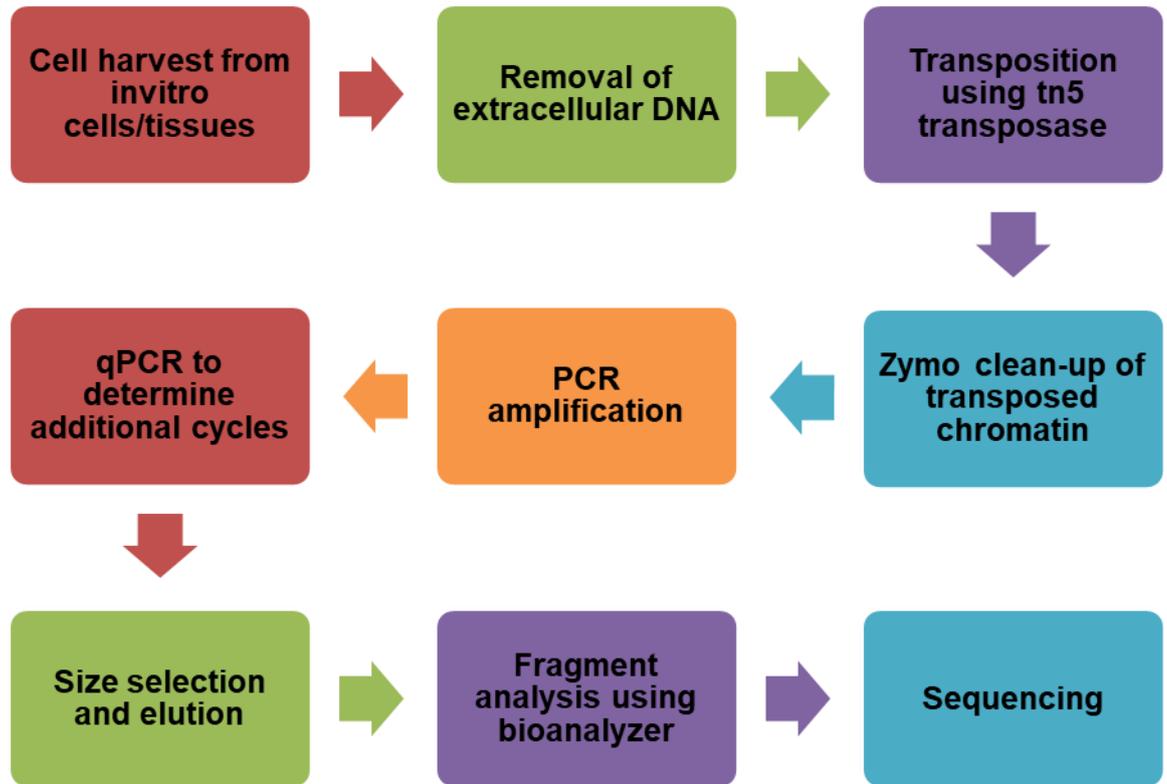


Figure 2.2 - ATAC-Seq flowthrough

Description	Catalogue	Supplier	Concentration stock	Concentration working
DNase	A37780010	Esmer	3755U/mg 5mg/ml 15000U/ml	200U/ml 1:75 in water

Digitonin	G9441	Promega	2% in DMSO	100x (1%) in water
Tween 20	P7949-500ml (Bioextra)	Sigma	10% (100x)	10% (100x)
NP40/IGEPAL	18896_100ml	Sigma	10% (100x)	10% (100x)
Zymo DNA clean and concentrator:5 kit	D4014	Zymo	-	-
NEBNext® High-Fidelity 2X PCR Master	M0541L	NEB	2X	2X
Sybr Gold(in DMSO)			-	25X
1M Tris HCL	T5941-1007	Sigma	1M	1M
5M NaCl	S3014	Sigma	5M	5M
1M MgCl ₂	M8266-100G	Sigma	1M	1M

Table 2.7 - ATAC-Seq Reagents

2.6.3.1 Making ATAC-Seq Reagents

1. 1M Tris HCL(Sigma cat*T5941-1007)

To make 1M Tris-HCL Ph 7.4

- a. Add 15.76 g of Tris.HCL in 70 ml of autoclaved milliq.
- b. Add NaOH to adjust Ph to 7.4 and add more milliq to a total of 100 ml
- c. Filter sterilize and store at RT

2. 5M NaCl (Sigma cat*S3014)

To make 5M NaCl

- a. Add 29.2 g of NaCL in 80 ml of autoclaved milliq.
- b. Add more milliq to a total of 100 ml

c. Filter sterilize and store at RT

3. 1M MgCl₂ (sigma cat*M8266-100g)

To make 5M MgCl₂

- a. Add 80 ml of milliq drop by drop to 9.52 g of MgCl₂
- b. Add more milliq to a total of 100 ml
- c. Filter sterilize and store at RT

4. ATAC-Resuspension Buffer (RSB):

Add 500 µl of 1M Tris-HCl pH 7.4, 100 µl 5M NaCl, 150 µl 1M MgCl₂ to 49.25 ml sterile water to make 50 ml ATAC-RSB.

5. RSB1 (50 µl /rxn):

Add 0.1% NP40, 0.1% Tween-20, and 0.01% Digitonin to 50 µl cold ATAC-Resuspension Buffer (RSB)

6. RSB2 (1ml/rxn):

Add 0.1% Tween-20 to 1 ml of cold ATAC-RSB containing 0.1% Tween-20

2.6.3.2 Clear out extracellular DNA

25mg of prostate tissue was homogenised on ice and cells were treated with DNase(Esmer) at 200U/ml in serum for 30 minutes at room temperature. DNase was washed off and cells were washed thoroughly but gently with PBS.

2.6.3.3 Transposition

Cells were pelleted at 500 RCF at 4°C for 5 min in a fixed angle centrifuge in a 0.5ml low binding tube (Fisher Scientific). Supernatant was removed and 50µl cold ATAC-Resuspension Buffer (RSB1) containing 0.1% NP40, 0.1% Tween-20, and 0.01% Digitonin was added. Cell suspension was pipetted up and down three times and incubated on ice for 3 minutes. Lysis buffer was washed out with 500µl cold ATAC-RSB 2 containing 0.1% Tween-20 and tube was inverted 3 times gently to mix the buffers. Nuclei was pelleted at 500 RCF for 10 min at 4°C. Supernatant was aspirated carefully avoiding the visible pellet. Extracted nuclei was checked using trypan blue assay (Thermo Fisher) to check the efficiency of lysis. Cell pellet was suspended in 50µl of transposition mixture by pipetting up and down 6 times in 1.5 ml low binding Eppendorf. Transposition reaction was incubated at 37°C for 30 minutes in a thermomixer with 1000 RPM mixing in 1.5 ml low binding Eppendorf.

2.6.3.4 Zymo clean up

DNA was cleaned up using Zymo DNA Clean and Concentrator-5 Kit (Zymo). 100µl of DNA Binding Buffer was added to 50µl of DNA sample) in a 1.5 ml low binding tube. Sample solution was vortexed briefly and transferred to Zymo-Spin™ Column2 with a collection tube. Column along with the sample was centrifuged at 13,000 x g for 30 seconds and the flow through was discarded. 200 µl DNA wash buffer was added to the column, centrifuged at 13,000 x g for 30 seconds and the flow through was discarded. This step was repeated twice followed by the addition of 21 µl DNA Elution Buffer 3 onto the membrane and incubated for 1 minute at room temperature. DNA column was transferred to a fresh 1.5 ml Eppendorf tube and centrifuged for 30 seconds at 13,000 x g to elute the DNA.

2.6.3.5 Library construction

For the construction of library, 20 μ l of cleaned sample was taken and amplified for 5 cycles using NEBNext 2x MasterMix. Each reaction contained 2.5 μ l of 25 μ M i5 primer, 2.5 μ l of 25 μ M i7 primer, 25 μ l 2x NEBNext master mix, and 20 μ l transposed/cleaned up sample. Reaction mix was taken in a 200 μ l 8 strip PCR tubes with flat lids (Thermo Fisher) with a total volume of 50 μ l reaction. Sample was initially amplified in a thermocycler (MJ Research - PTC200) using conditions in Table 2.8. After initial amplification was complete, tubes were removed from thermocycler and stored on ice.

Step	Number of cycles	Temperature (°C)	Time
1	1	72	5 minutes
2	1	98	30 seconds
3	5	98	10 seconds
		63	30 seconds
		72	1 minute
4	1	4	hold

Table 2.8 - ATAC-Seq initial library construction

2.6.3.6 RT-PCR to determine additional cycles

qPCR was performed to determine additional cycles required for library synthesis 5 μ l (10%) of the pre-amplified library was taken to be run on qPCR 480 thermocycler (Roche) in a 15 μ l reaction. To 5 μ l sample, 3.76 μ l sterile water, 0.5 μ l 25 μ M i5 primer, 0.5 μ l 25 μ M i7 primer, 0.24 μ l 25x SYBR Gold (in DMSO), 5 μ l 2x NEBNext master mix were added. The reaction mix was added in a 384 well plate and run on a thermocycler (MJ Research - PTC200). Libraries were made using adapters to barcode samples before sequencing (Table 2.9).

Sample	Barcode adapter	Sequence	Cohort
Gla 01-HPAA5.3J	Ad 2.10 _CGAGGCTG	CAAGCAGAAGACGGCATAACGAGATCAGCCTCGGTCTCGTGGGCTCGGAGATGT	Pten ^{fl/fl} Arid1a ^{+/+}
Gla 02 – HPAA31.1D	Ad 2.11 _AAGAGGC A	CAAGCAGAAGACGGCATAACGAGATTGCCTCTTGTCTCGTGGGCTCGGAGATGT	Pten ^{fl/fl} Arid1a ^{+/+}
Gla 03 – HPAA31.1E	Ad 2.12 _GTAGAGG A	CAAGCAGAAGACGGCATAACGAGATTCTCTACGTCTCGTGGGCTCGGAGATGT	Pten ^{fl/fl} Arid1a ^{+/+}
Gla 04 – HPAA5.3D	Ad 2.13 _GTCGTGAT	CAAGCAGAAGACGGCATAACGAGATATCACGACGTCTCGTGGGCTCGGAGATGT	Pten ^{fl/fl} Arid1a ^{fl/+}
Gla 05 – HPAA10.1B	Ad 2.14 _ACCACTGT	CAAGCAGAAGACGGCATAACGAGATACAGTGGTGTCTCGTGGGCTCGGAGATGT	Pten ^{fl/fl} Arid1a ^{fl/+}
Gla 06 – HPAA10.1F	Ad 2.16 _CCGTTTGT	CAAGCAGAAGACGGCATAACGAGATACAAACGGGTCTCGTGGGCTCGGAGATGT	Pten ^{fl/fl} Arid1a ^{fl/+}
X1 – HPAA15.1F	Ad 2.21 _TGGGTTTC	CAAGCAGAAGACGGCATAACGAGATGAAACCCAGTCTCGTGGGCTCGGAGATGT	Pten ^{fl/fl} Arid1a ^{fl/f} I
X2 – HPAA20.2C	Ad 2.22 _TGGTCACA	CAAGCAGAAGACGGCATAACGAGATTGTGACCAGTCTCGTGGGCTCGGAGATGT	Pten ^{fl/fl} Arid1a ^{fl/f} I
X2 – HPAA15.2H	Ad 2.23 _TTGACCCT	CAAGCAGAAGACGGCATAACGAGATAGGGTCAAGTCTCGTGGGCTCGGAGATGT	Pten ^{fl/fl} Arid1a ^{fl/f} I

Table 2.9 - ATAC-Seq barcodes in each library

The amplification profiles were assessed manually and the required number of additional cycles to amplify, after qPCR amplification was determined. Rn value, was used to calculate the additional number of cycles required by plotting linear Rn versus cycle number. The number of cycles (n) that corresponds to ¼ of maximum fluorescent intensity was taken and the preamplified library was further amplified for number of cycles with the following thermocycler conditions in the same tube. Table of amplification of library shown in Table 2.10

Step	Number of cycles	Temperature (°C)	Time
1	1	98	30 seconds

2	n	98	10 seconds
		63	30 seconds
		72	1 minute
3	1	4	hold

Table 2.10 - ATAC-Seq PCR steps for library amplification

2.6.3.7 Size selection and elution

Libraries were double-sided bead purified using XP Beads (AMPure) to remove primer dimers and fragments larger than 1000 bp. 30µl of the sample was transferred to an Eppendorf tube and 15µl (0.5 X) AMPure XP beads were added to each tube and pipetted up and down 8X to mix thoroughly. Samples were incubated at room temperature for 10 minutes and pulse spun. Eppendorf tubes were kept on a magnetic rack up to 2-3 minutes until supernatant was clear of beads. Supernatant was transferred to a new tube and 58.5 µl (1.3X original volume) AMPure XP beads was added to each tube. Samples were pipetted up and down 8x times and incubated at room temperature for 10 minutes before being placed on a magnetic rack for 2-3 minutes until supernatant was clear of beads. Supernatant was removed and beads were washed with fresh 80% ethanol (made with nuclease free water). Ethanol was removed and beads were allowed to dry for 1-2 mins (made sure not to over dry the beads). DNA was eluted in 20 µl nuclease-free H₂O and stored at -20°C.

2.6.3.8 Fragment analysis and DNA quantification

After library generation, DNA library is quantified using a fluorescent-based Qubit assay. Libraries are diluted based on the concentration obtained from Qubit and then ran on Bioanalyser (Agilent) to analyse the structure of nucleosome and the fragment size generated using transposition.

2.6.3.9 RT-PCR for Omni-ATAC

Following fragment analysis, libraries are run on a RT-PCR to assess the enrichment of closed region vs open regions and contamination of mitochondrial DNA. Libraries are diluted in the ratio of 1:10 using nuclease free water. 10 μ l PCR mix reaction is used by adding 1 μ l of forward and reverse primer each to 5 μ l of Sybr green (Sigma Aldrich) in the 384 well plate. To this mix, 3 μ l of diluted library is added and the plate is sealed with plate sealing tape (Thermo Fisher) and run in a thermocycler (MJ Reseach - PTC200). Plate was ran at 95 $^{\circ}$ C for 5 minutes to denature DNA. DNA amplification steps were performed for 40 cycles, beginning with 95 $^{\circ}$ C or 30 seconds, 60 $^{\circ}$ C for 30 seconds, and finally 32 $^{\circ}$ C for 30 seconds before repeating. A melt curve step was included by increasing temperature gradually up to 40 $^{\circ}$ C for 1 minutes. Samples were then held at 4 $^{\circ}$ C.

2.6.3.10 ATAC-Sequencing

Following library preparation samples were prepared and sequenced on NextSeq 2000 (Illumina) with 1,500 million 2x100bp paired end reads by Glasgow Polyomics' next generation sequencing and transcriptomics service. Analysis performed by Mr Robin Shaw (Informatician, Beatson Institute) using analysis pipeline and expertise of Holly Brunton (Beatson Institute).

2.7 Statistics

All data values shown are presented as mean +/- standard error of the mean (SEM) from three experimental replicates unless otherwise stated. To determine statistical significance the appropriate statistical test was performed. In comparing two mean values of two groups, Wilcoxon rank or Mann-Whitney tests were used. When comparing the mean values of more than two independent

groups, a one-way analysis of variance (ANOVA) test was used. When comparing Kaplan-Meier survival curves the Log Rank (Mantel Cox) test was used. A p-value of >0.05 (*) was considered the threshold for significance, whereas $P<0.01$ (**), $P<0.001$ (***), $P<0.0001$ (****) represent higher degrees of significance. Graphical representations produced using either GraphPad Prism version 9.3.1 or R-Studio version +492.

3 Loss of *Arid1a* Cooperates With Loss of *Pten* to Drive Prostate Tumour Growth in a Murine Model

3.1 Introduction

In humans, PTEN is mutated in around 20% of cases of primary PC, however it is mutated in around 40-60% of metastatic PC (Robbins, Tembe *et al.* 2011, Jamaspishvili, Berman *et al.* 2018). This indicates that PTEN loss is more likely to be associated with advanced disease and metastasis. This was modelled in mice using a Cre-recombinase system on a prostate-specific *Pb* promoter. *Pten* can be engineered out from mouse genome to produce heterozygous or homozygous knockout in prostate cells (Figure 3.1). This recombination only occurs in prostate luminal epithelial cells, as basal epithelial cells do not express *Probasin* (Wu, Wu *et al.* 2001). Recombination occurs in all lobes of the prostate, though the recombination rate is highest in lateral and ventral lobes. Mice with a *Pb Cre-recombinase* (*Pb-Cre*) positive, *Pten flox/+* (*fl/+*) genotype rarely develop prostate tumour (Wang, Gao *et al.* 2003). Those that do have often lost the second copy of *Pten* or have gained mutations in other genes. *Pb-Cre +ve, Pten flox/Pten flox* (*fl/fl*) mice typically develop prostate tumours that reach a clinical endpoint within 9-12 months. These tumours are often large,

cystic, and rarely metastatic.

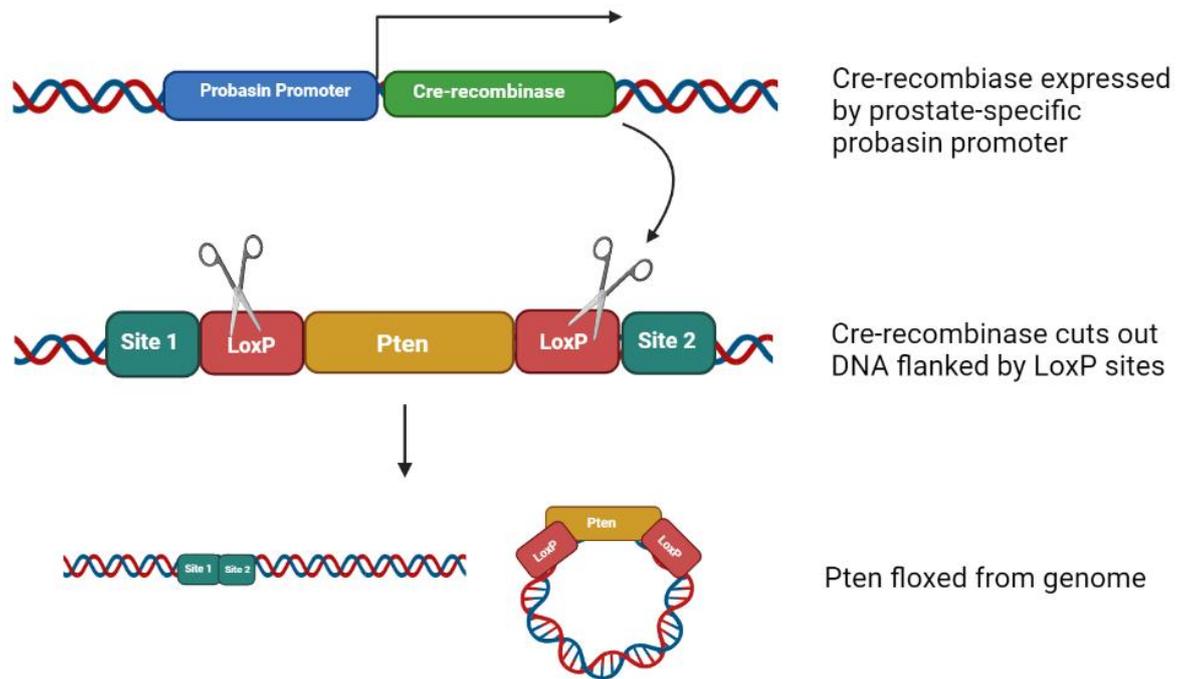


Figure 3.1 - Probasin Cre-recombinase system in mouse prostate

Cre-recombinase enzyme is under control of prostate-specific *probasin* promoter. Following *Cre-recombinase* expression, it cuts out DNA elements which are flanked by LoxP sites can fuse DNA back together. Circular DNA contain *Pten* element which is then later degraded.

This discrepancy, that PTEN drives advanced disease and metastasis in humans, but causes slow, none-metastatic tumour development in mice indicates that PTEN loss alone may not be sufficient to drive advanced PC.

A transposon-based mutagenesis screen was carried out to investigate which other genes cooperate with the loss of PTEN to drive PC (Ahmad, Mui *et al.* 2016). In this study, the SB system uses a transposon (*T2Onc3*) and a transposase (*Rosa*) to randomly disrupt gene expression. The transposase inserts the transposon randomly into a gene. If the insertion orientation is inserted into the antisense DNA strand it can downregulate gene expression or upregulate expression if inserted into the sense strand. *T2Onc3* and *Rosa* were bred onto the *Pb-Cre+ve Pten^{fl/fl}* mouse line with the breeding scheme shown in Figure 3.2.

Mice were aged until clinical endpoint then had the tumours sequenced to see where the transposon inserted. Phenotypic changes in tumour development such as accelerated tumour growth or elevated metastasis could be correlated with which genes had transposons inserted. From here, genes that cooperated with loss of *Pten* to drive tumorigenesis could be identified. The resultant genes could then be altered in a genetically engineered mouse (GEM) model to understand their function.

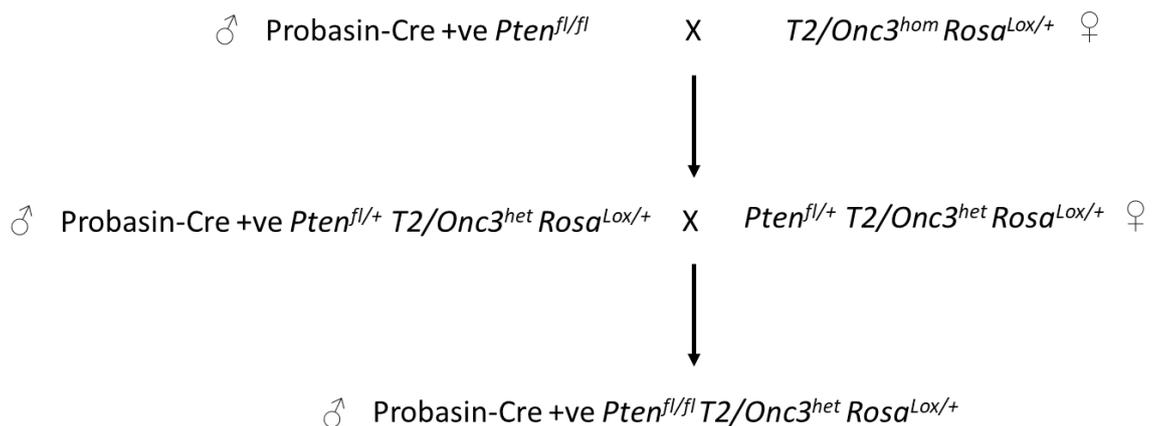


Figure 3.2 – *Sleeping Beauty* screen breeding scheme

Probasin-Cre+ve *Pten*^{fl/fl} mouse is bred with mouse with *Sleeping Beauty* transposon (*T2/Onc3*^{hom}) and transposase (*Rosa*^{Lox+}). This produces *Pten*^{fl/+} mice with the *Sleeping Beauty* system. Breeding these mice together produces *Pten*^{fl/fl} mouse with *Sleeping Beauty* system (Ahmad, Mui et al. 2016).

3.2 Hypothesis and Aims

As *Pten* alone does not drive metastatic disease there are likely many other genes which can cooperate with it to drive tumorigenesis. This chapter addresses the hypothesis that genes would cooperate with *Pten* by converging on common pathways known to be affected by *Pten* such as PI3K-AKT signalling or fatty acid synthesis.

The aims of this chapter are as follows:

- Investigate the effect and location of transposon insertions in a mouse model of prostate cancer to identify new prostate cancer genetic events.
- Generate a model in which Probasin-Cre+ve *Pten^{fl/fl}* mouse model can be combined with a driver mutation identified in SB screen
- Validate the role of the SB screen driver in a mouse model of PC and characterise its ability to collaborate with *Pten*.

3.3 Results

3.3.1 Sleeping Beauty – A Transposon-based Forward Mutagenesis Screen

From the SB screen, *Arid1a* was identified as a potential driver of PC in combination with *Pten* loss with four mice out of twenty-one having the SB transposon inserted into the *Arid1a* gene (Figure 3.3). Each of these transposon insertions were antisense which likely reduces *Arid1a* expression (DeNicola, Karreth *et al.* 2015). Other targets were identified from Sleeping Beauty screen, including peroxisome proliferator activated receptor gamma (*Pparg*), membrane bound transcription factor site-2-protease (*Mbtps2*), and cyclin E1 (*CCNE1*) were identified as potential drivers.

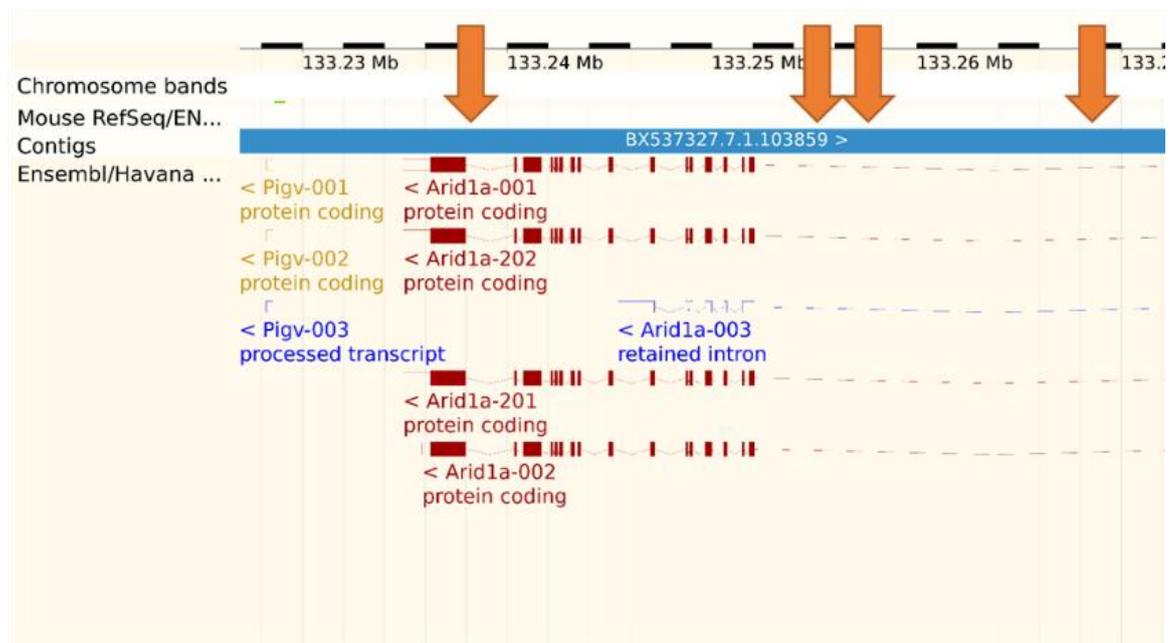
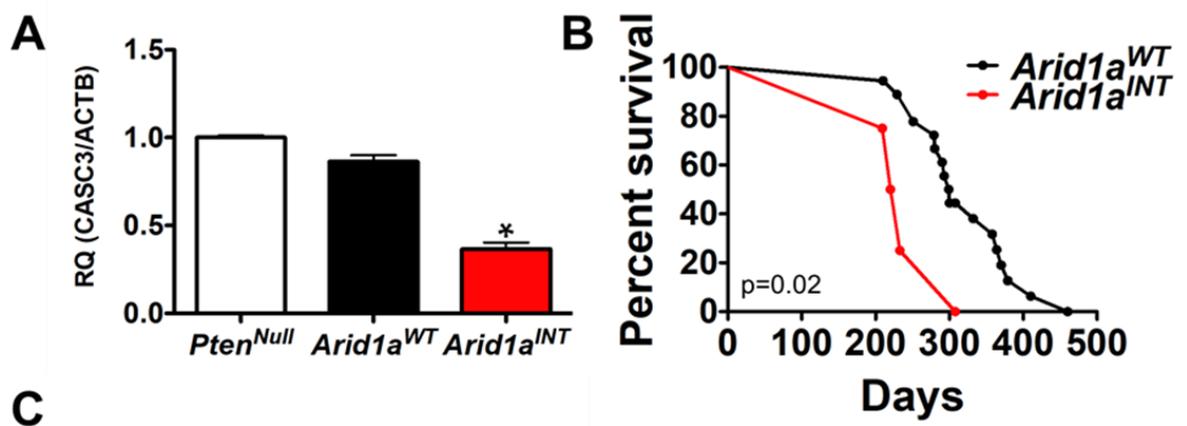


Figure 3.3 - *Sleeping Beauty* transposon insertion sites in *Arid1a* gene

Insertions sites indicated by orange arrow shows location of transposon insertion against Ensembl schematic of *Arid1a* gene. All insertions were antisense which likely reduces *Arid1a* expression.

To investigate the impact of insertions in the *Arid1a* gene on gene expression, QRT-PCR was performed to compare *Arid1a* expression levels in *Pb-Cre+ve*

Pten^{fl/fl} mice compared to *Pb-Cre+ve Pten*^{fl/fl} *T2/Onc3*^{het} *Rosa*^{fl/+} mice with and without *Arid1a* insertions. Results confirmed that mice with the transposon inserted into the *Arid1a* gene had a significant decrease in expression of *Arid1a* (Figure 3.4A). As expected, mice without *Arid1a* insertions had comparable *Arid1a* levels to *Pb-Cre+ve Pten*^{fl/fl} mice. Importantly, mice with *Arid1a* insertions had a poorer survival when compared to mice without *Arid1a* insertions (Figure 3.4B). Within the cohort of mice with reduced *Arid1a*, 75% (3/4) developed distal metastasis to the lung, and 100% (4/4) developed metastasis to lymph nodes (LN) (Figure 3.4C). Other SB screen mice without *Arid1a* insertions saw LN metastasis in 65% (11/17) of cases, and lung metastasis in 17% (3/17) of cases. These mice had higher instance of metastasis compared to *Pb-Cre+ve Pten*^{fl/fl} as they had transposon insertions in other genes. This indicates that reduced levels of *Arid1a* may lead to decreased survival and promoted metastasis.



	Lymph node metastasis	Lung metastasis
Arid1a insertion	100%	75%
No Arid1a insertion	65%	17%

Figure 3.4 - Reduced expression of *Arid1a* reduces survival and promotes metastasis

A. Mice with insertions (INT) in the *Arid1a* gene have reduced mRNA detected by qRT-PCR $P=0.05(*)$ tested by one-way ANOVA. **B.** Mice with INT in *Arid1a* gene have significantly reduced survival ($p=0.02$) tested by Log-Rank (Mantel-Cox). **C.** 20% of total mice had insertions in the ARID1A. Of those 75% had lung metastasis and 100% had lymph node metastasis (Ahmad, Mui et al. 2016)

3.3.2 Recombination of *Arid1a* and *Pten* Leads to Loss of Protein in Murine Tumours

To investigate how *Arid1a* and *Pten* genetically interact, a mouse line was generated which would contain a knockout of *Arid1a* and *Pten* alleles in various combinations by using the Cre-recombinase LoxP system. The *Pten* gene was knocked out of the mouse genome by recombination of exon 5 which was flanked by LoxP sites to cause degradation of the *Pten* transcript. Similarly, exon 8 of *Arid1a* was flanked by LoxP sites to degrade the transcript (2.1.1). *Pb-Cre* expression begins as the mice enter puberty at around 2 months where *Probasin* is most strongly expressed in the lateral and ventral lobes of the prostate epithelium. Following *Pb-Cre* expression, *Arid1a* and *Pten* begin being deleted. IHC of *Arid1a* shows progressively decreased levels of nuclear staining in *Arid1a*^{fl/+} and *Arid1a*^{fl/fl} genotypes confirming that the recombination of the *Arid1a* gene leads to a loss of protein (Figure 3.5). Staining intensity was quantified using Halo Image Analysis software.

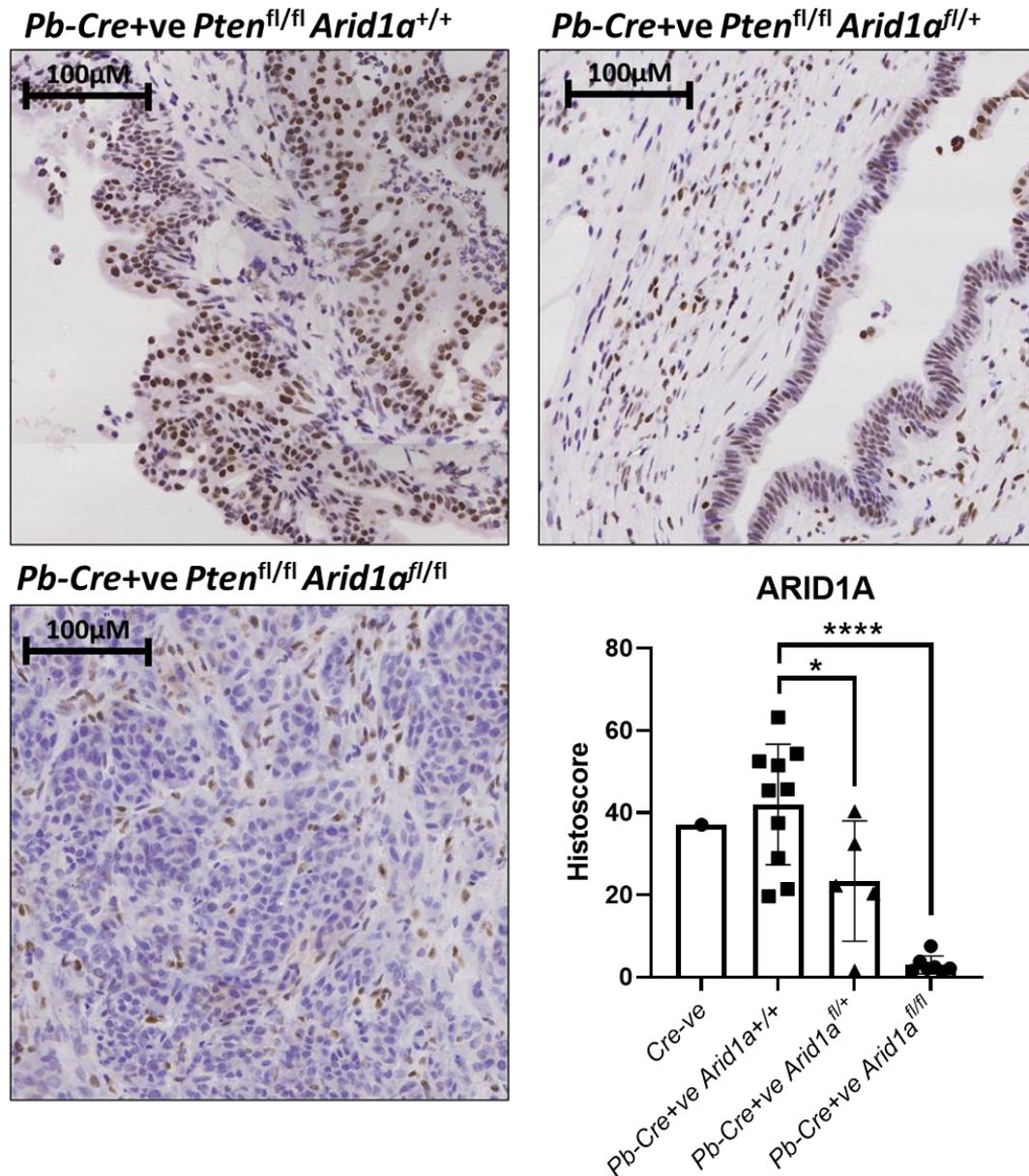


Figure 3.5 – ARID1A nuclear staining is reduced following Pb-Cre expression to delete the gene in a mouse model.

Mouse prostate tumour tissue of indicated genotype were stained for ARID1A images show representative images of each genotype (shown in above image). Each dot represents one mouse, with *Pb-Cre+ve Pten^{fl/fl} Arid1a^{+/+}* showing 10 mice, *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/+}* showing 5 mice, and *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* showing 7 mice. Stain intensity was quantified for ARID1A into a histoscore using Halo Image analysis software P-value <0.05 (*), P-value <0.0001(****) tested by one-way ANOVA. Histoscore determined by formula: (% cells low intensity)+2(%cells medium intensity)+3(%cells high intensity)=Histoscore.

To confirm PTEN loss, IHC was performed on the tumours PTEN. However, staining revealed inconsistency in PTEN levels between tumours as well as some PTEN stain in *Pb-cre+ve Pten^{fl/fl}* mice. Quantifying this stain with Halo Image

Analysis software shows PTEN levels remain detectable following recombination. As this *Pten* model is well characterised, we believed this staining to be an artefact of the antibody as it does not show a reduction in PTEN levels as expected (Figure 3.6)(Ahmad, Patel *et al.* 2011, Patel, Brzezinska *et al.* 2020).

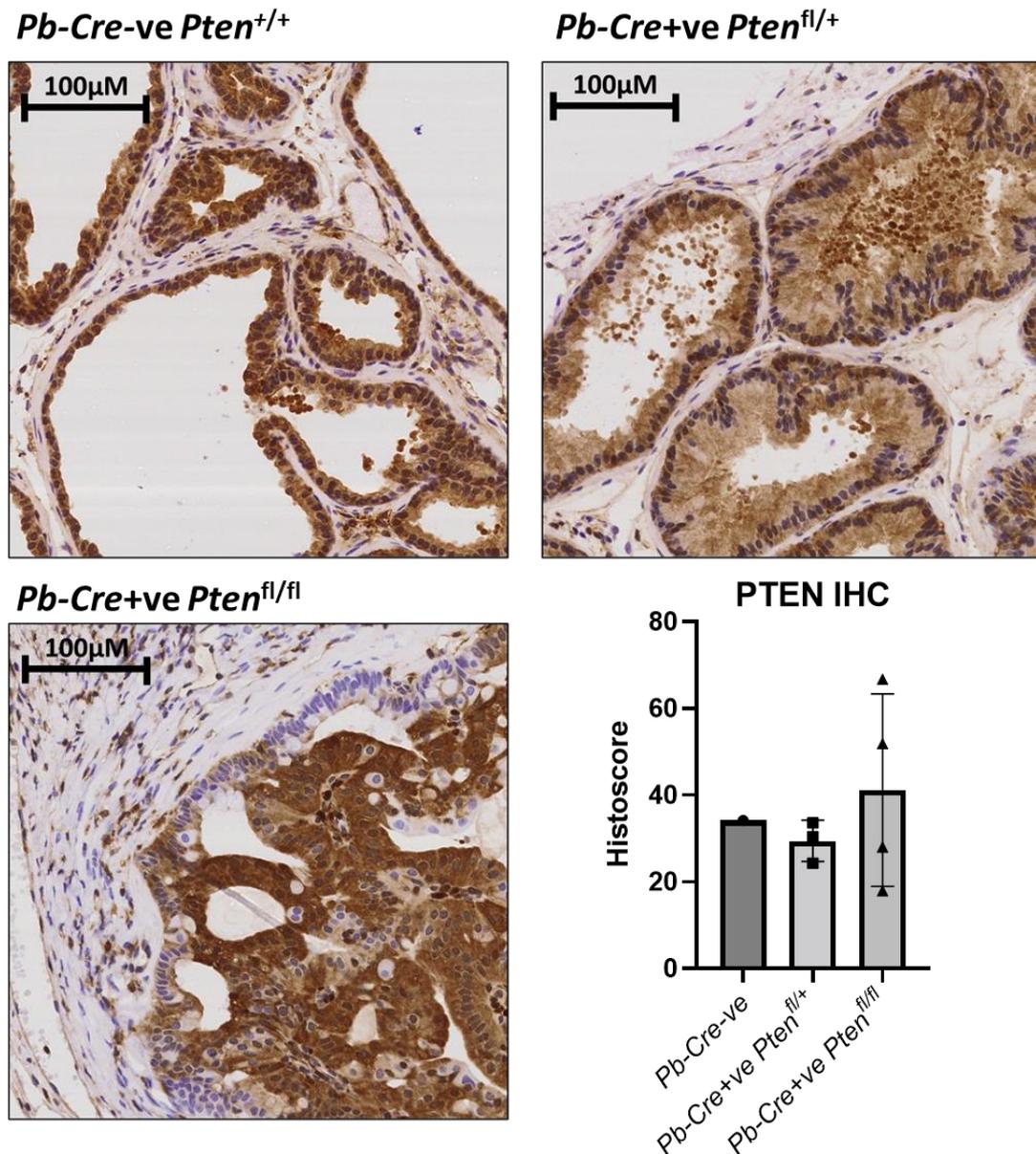


Figure 3.6 - Recombination of *Pten* did not lower detectable levels of PTEN protein by IHC

Mouse prostate tissue of indicated genotype were stained for PTEN, images show representative images of each genotype (shown above the image). Each dot represents one mouse, with *Pb-Cre-ve* showing 1 mice, *Pb-Cre+ve Pten*^{fl/+} showing 3 mice, and *Pb-Cre+ve Pten*^{fl/fl} showing 4 mice. Stain intensity was quantified for PTEN into a histoscore using Halo Image analysis software. Histoscore determined by formula: (% cells low intensity)+2(%cells medium intensity)+3(%cells high intensity)=Histoscore.

However, to confirm PTEN was lost, IHC was performed for a downstream target of PI3K which becomes active following PTEN loss. For this we used phosphorylation of AKT on serine473 which should become phosphorylated following PTEN loss. As expected, complete loss of PTEN was required before there is an elevation in pAKT staining by IHC. Following *Pten* recombination, 70-80% of epithelial cells in the tumours show strong pAKT staining which is not altered by *Arid1a* status (Figure 3.7).

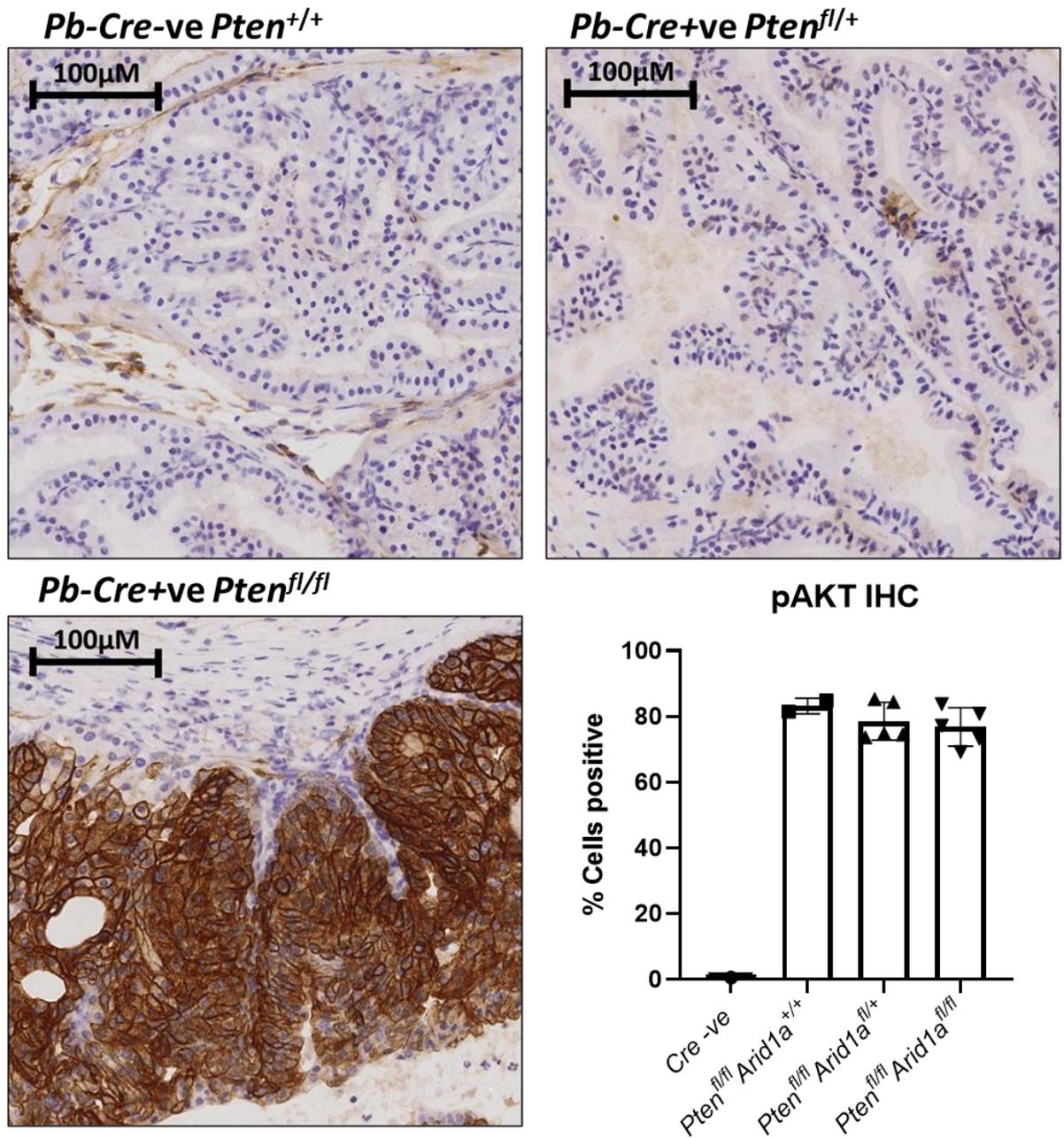


Figure 3.7 - Loss of PTEN leads to pAKT phosphorylation which is unaltered by *Arid1a* status

Mouse prostate tissue of indicated genotype were stained for pAKT. Images show representative images of each genotype (shown above the image). Each dot represents one mouse, with *Pb-Cre-ve* showing 1 mouse, *Pb-Cre+ve Pten^{fl/fl} Arid1a^{+/+}* showing 2 mice, *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/+}* showing 5 mice, and *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* showing 5 mice. Stain intensity was quantified for pAKT into a histoscore using Halo Image analysis software

3.3.3 Loss of *Arid1a* Accelerates Tumorigenesis in *Pten* Null Mice

Pb-Cre+ve Pten^{fl/fl} mice are well characterised for developing large cystic tumours of the anterior prostate that typically reach clinical endpoint between 9 - 12 months.

Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/+} (Figure 3.8B) mice developed prostate tumours very similar to control *Pb-Cre+ve Pten^{fl/fl}* mice (Figure 3.8A) with large cystic tumours of the anterior prostate. Overall, there was no significant difference in the survival of these two cohorts indicating that losing a single copy of *Arid1a* does not impact survival (Figure 3.9). However, six *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/+}* mice succumbed to tumourigenesis before 7 months which is uncommon for *Pb-Cre+ve Pten^{fl/fl}*. In contrast, four *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/+}* mice came down after 11.5 months with the latest at 13.5 months. This indicates that loss of one copy of *Arid1a* introduces variability in the cohort.

Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl} mice however developed very differently to the other genotypes. Loss of both copies of *Arid1a* with *Pten* loss caused development of tumours of both the dorsolateral and anterior lobes of the prostate (Figure 3.8C). In these mice, clinical endpoint was reached due to bladder distension around 3.5 months (Figure 3.9). The tumours themselves were also very dense without any cystic fluid which is typically observed in a *Pb-Cre+ve Pten^{fl/fl}* mouse, particularly in the dorsolateral lobes, but also in anterior lobes (Figure 3.10). The tumour of the *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* dorsolateral prostate also had an invasive phenotype, with a breakdown of the normal basement membrane of the prostate gland as seen in normal *Pb-Cre+ve Pten^{fl/fl} Arid1a^{+/+}* (Figure 3.11).

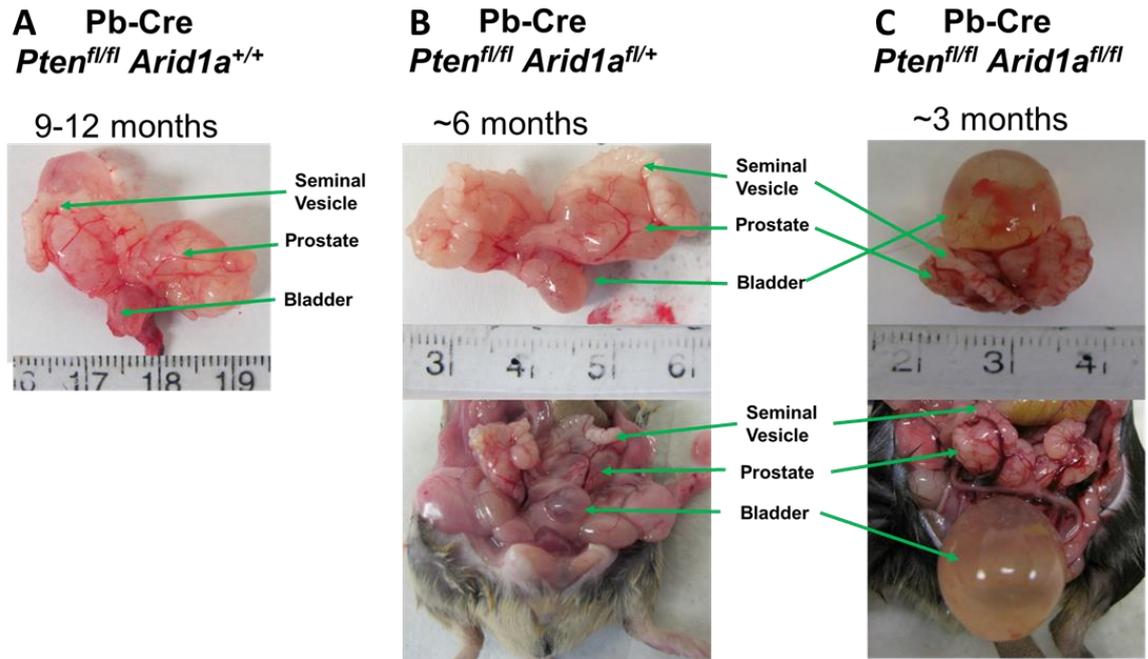


Figure 3.8 – Tumour morphology in Pb-Cre+*Pten*^{fl/fl} mice with different *Arid1a* status

Representative tumour morphology of Pb-Cre+ve; *Pten*^{fl/fl} mice with different *Arid1a* genotypes; A) *Arid1a*^{+/+} (n=10); B) *Arid1a*^{fl/+} (n=19); C) *Arid1a*^{fl/fl} (N=8) either dissected (top images) or in situ (lower images). Key structures are indicated by arrows. Age of representative animals shown here is given (months) above the images.

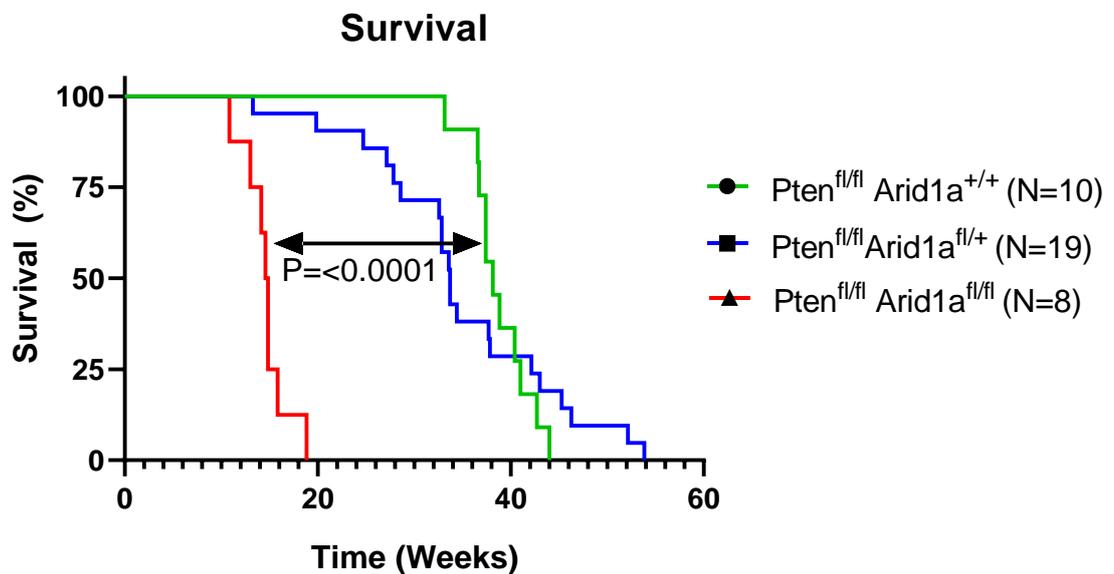


Figure 3.9 – Loss of *Arid1a* reduces survival in *Pten* null mice

Kaplan-Meier survival curve comparing *Pb-Cre+ve Pten*^{fl/fl} *Arid1a*^{+/+} (N=10)(green line), *Pb-Cre+ve Pten*^{fl/fl} *Arid1a*^{fl/+} (N=19)(blue line), and *Pb-Cre+ve Pten*^{fl/fl} *Arid1a*^{fl/fl} (N=8)(red line) mice. P-value shows significance between *Pb-Cre+ve Pten*^{fl/fl} *Arid1a*^{+/+} and *Pb-Cre+ve Pten*^{fl/fl} *Arid1a*^{fl/fl} as shown. P-value=<math>< 0.0001</math> tested by log-rank (Mantel-lox) test.

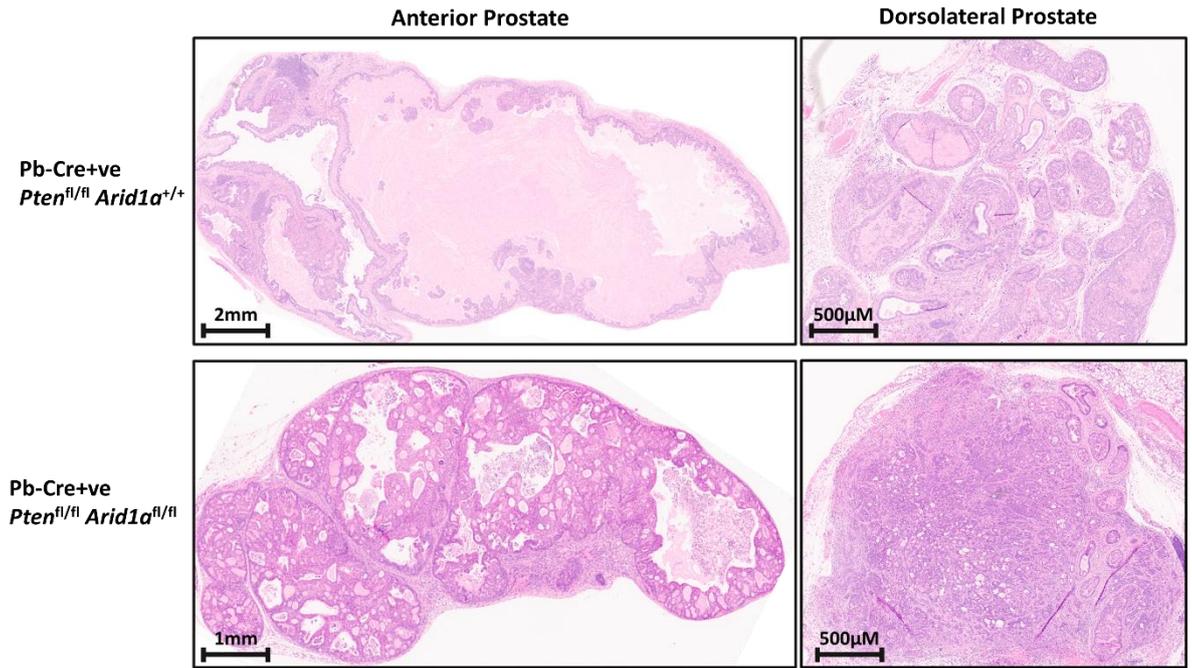


Figure 3.10 – Loss of *Arid1a* alters tumour morphology in anterior and dorsolateral prostate lobes

Prostate tumours of indicated lobe and genotype were stained for haematoxylin and eosin (H&E) to show tumour density. *Pten^{fl/fl} Arid1a^{+/+}* mouse shown was 10 months of age at endpoint shown, whereas *Pten^{fl/fl} Arid1a^{fl/fl}* was 3.5 months.

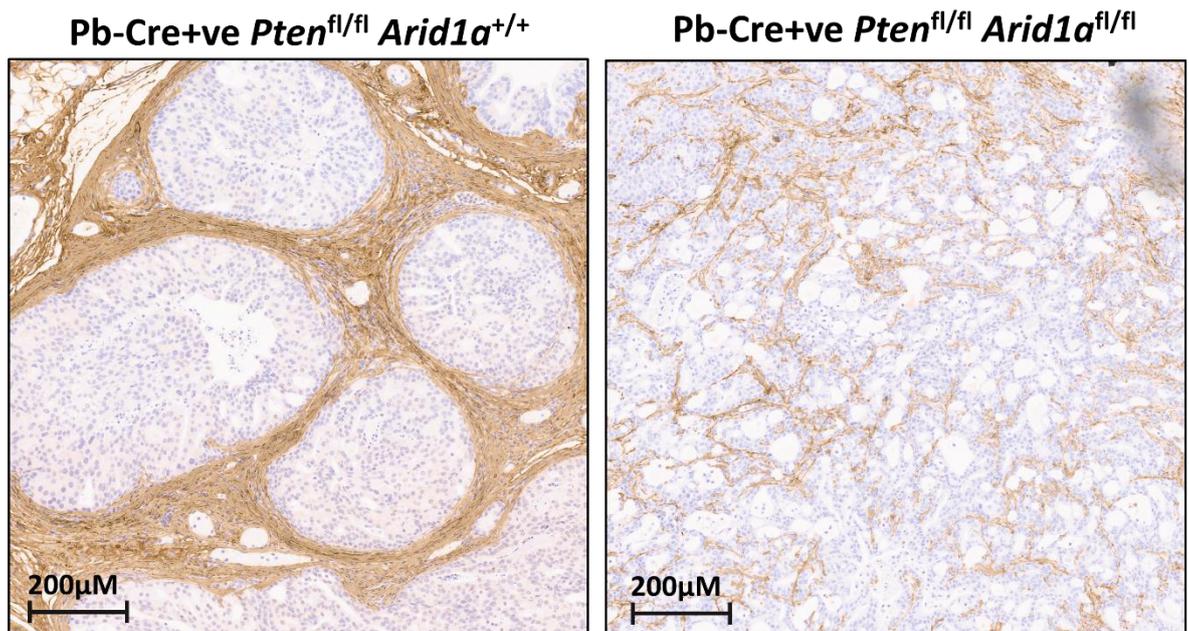


Figure 3.11 - Loss of *Arid1a* causes invasive phenotype of dorsolateral prostate

Dorsolateral prostate of indicated genotype stained for pan-collagen to show integrity of prostate gland. *Pten^{fl/fl} Arid1a^{+/+}* mouse shown was 10 months of age at endpoint shown, whereas *Pten^{fl/fl} Arid1a^{fl/fl}* was 3.5 months.

Tumour weights from the different cohorts were recorded as wet weight, which includes cystic fluid and infiltrate; and dry weight, which is recorded after draining fluid and infiltrate so represents the true mass of the tumour itself (Figure 3.12). This will show if the tumour weights were composed of cystic fluid or necrotic tissue giving insight into the tumour morphology.

There was no significant difference in the wet weight between *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/+}* and *Pb-Cre+ve Pten^{fl/fl}* mice as these cohorts both develop cystic tumours. There is no wet weight for *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* mice as these mice develop solid tumours with no cystic fluid. There was also no significant difference in the dry tumour weights from these cohorts indicating a similar tumour mass contributed to the clinical endpoint of these mice.

Altogether, these data indicate that losing one copy of *Arid1a* in a *Pten* null background does not accelerate tumorigenesis, and the morphology of these tumours is comparable to *Pten* null alone tumours. In comparison, losing both copies of *Arid1a* in a *Pten* null background accelerates tumorigenesis and produces solid tumours. *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* also reach a similar tumour mass at endpoint as compared to *Pb-Cre+ve Pten^{fl/fl} Arid1a^{+/+}* despite a difference in 6 months in survival.

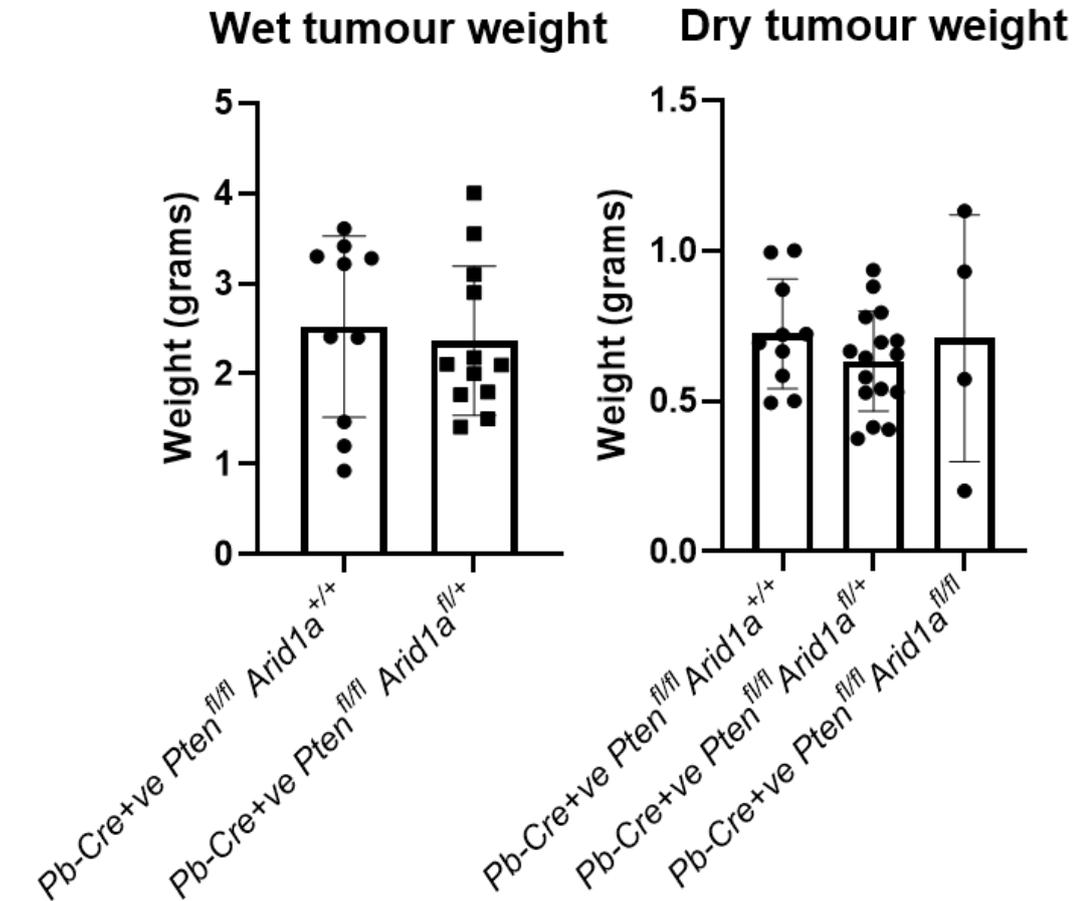


Figure 3.12 – Tumour weight at clinical endpoint for each genotype

Graph shows wet and dry tumour weights of indicated genotypes. Wet prostate weight includes cystic fluid and infiltrate whereas dry weight has had fluid drained. Each dot represents one mouse, with *Pb-Cre+ve Pten^{fl/fl} Arid1a^{+/+}* showing 15 mice, *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/+}* showing 4 mice, and *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* showing 7 mice. No significance found between any cohorts. No wet weight available for *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* due to solid tumour.

3.3.4 *Pten* Loss is a Prerequisite to *Arid1a* Loss Being Tumorigenic

To investigate the specific role of *Arid1a* in prostate development and tumour development, mice were bred with either *Pb-Cre+ve Pten^{fl/+}* or *Pb-Cre+ve Pten^{+/+}* genotypes to see the effect of *Arid1a* loss alone.

Pb-Cre+ve Pten^{fl/+} mice are phenotypically very similar to wild type mice (*Pb-Cre-ve Pten^{+/+}*) as they do not often develop prostate tumours. Those *Pb-Cre+ve Pten^{fl/+}* mice that do develop prostate tumours have often lost the second copy of *Pten* causing an equivalent of *Pb-Cre+ve Pten^{fl/fl}* phenotype with a typically smaller cystic tumour (Ratnacaram, Teletin *et al.* 2008). This loss of the second copy of *Pten* is often spontaneous so can occur at any age.

Pb-Cre+ve Pten^{fl/+}Arid1a^{fl/+} developed very similarly to *Pb-Cre+ve Pten^{fl/+}* mice with only one of the 19 mice developing a small prostate tumour by the 15-month aging endpoint (Figure 3.13). Like the previous *Pb-Cre+ve Pten^{fl/fl}Arid1a^{fl/+}* cohort described in 3.3.3, loss of one copy of *Arid1a* did not seem to alter the *Pten^{fl/+}* phenotype.

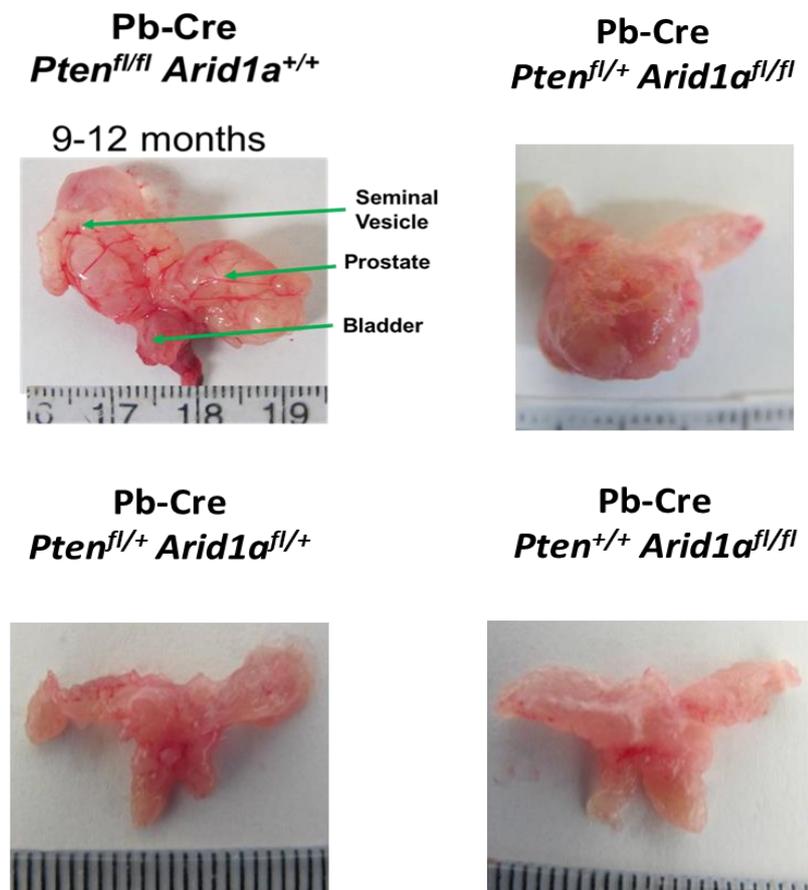


Figure 3.13- Complete loss of *Pten* is required for *Arid1a* loss to be tumorigenic

Representative dissected prostates of mice with following genotypes: *Pb-Cre+ve Pten^{fl/fl} Arid1a^{+/+}* (N=10), *Pb-Cre+ve Pten^{fl/+} Arid1a^{fl/fl}* (N=10), *Pb-Cre+ve Pten^{fl/+} Arid1a^{fl/+}* (N=19), and *Pb-Cre+ve Pten^{+/+} Arid1a^{fl/fl}* (N=18). Key structures are indicated by arrows. *Pten^{fl/fl} Arid1a^{+/+}* mouse shown was 11 months of age at endpoint shown, whereas *Pten^{fl/+} Arid1a^{fl/fl}* was 6 months, *Pten^{fl/+} Arid1a^{fl/+}* was 15 months, *Pten^{+/+} Arid1a^{fl/fl}* was also 15 months.

Pb-Cre+ve Pten^{fl/+}Arid1a^{fl/fl} however were phenotypically distinct from *Pb-Cre+ve Pten^{fl/+}* mice. Three of the eleven mice developed endpoint sized prostate tumours at 4, 9, and 15 months respectively (Figure 3.14). Two mice were also found-dead with prostate tumour at endpoint. The other mice in this cohort reached the 15-month aging endpoint without developing a prostate tumour. The tumours that had developed were similar to the *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* mice with small solid tumours of the dorsolateral prostate which caused bladder distension (Figure 3.15A and Figure 3.15B). IHC was performed on these tumours which showed that the second copy of *Pten* had been lost confirming that these tumours were similar to the *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* as indicated by the phenotype. PTEN stain was retained in intact glands adjacent to the tumour, indicating this was a spontaneous event (Figure 3.15C). Interestingly, PTEN staining was retained following recombination (Figure 3.6), but not present following spontaneous loss. This may be due to recombination retaining some detectable PTEN protein. However, this detectable protein is lost

through the spontaneous event.

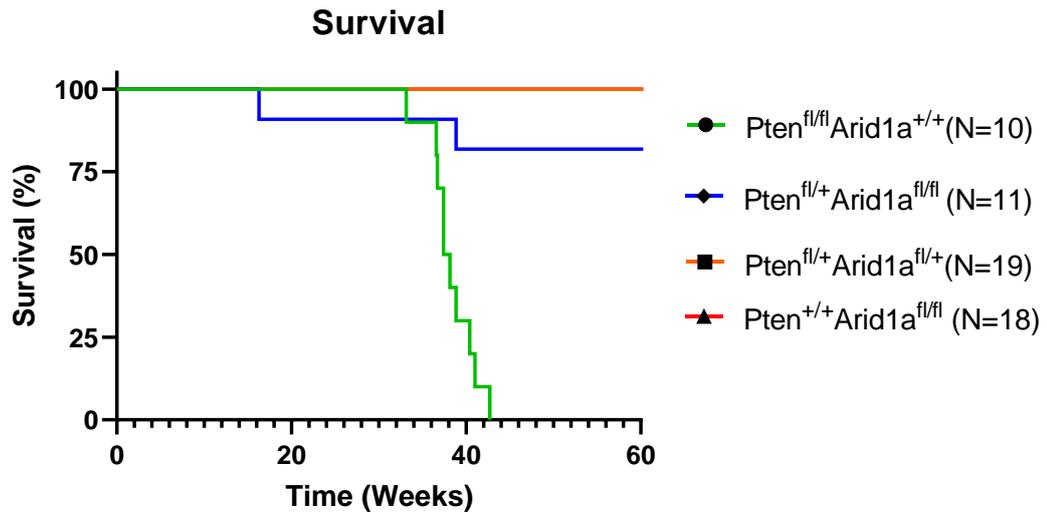


Figure 3.14 – *Pten* loss is a prerequisite to *Arid1a* loss being tumorigenic

Kaplan-Meier survival curve comparing *Pb-Cre+ve Pten^{fl/fl} Arid1a^{+/+}* (green line; N=10), *Pb-Cre+ve Pten^{fl/+} Arid1a^{fl/fl}* (blue line; N=10), and *Pb-Cre+ve Pten^{fl/+} Arid1a^{fl/+}* (orange line; N=19), and *Pb-Cre+ve Pten^{+/+} Arid1a^{fl/fl}* (red line; N=18) genotypes. *Pb-Cre+ve Pten^{fl/+} Arid1a^{fl/fl}* (N=10) vs *Pb-Cre+ve Pten^{fl/+} Arid1a^{fl/+}* P=0.0001 as determined by log-rank (Mantel-lox) test.

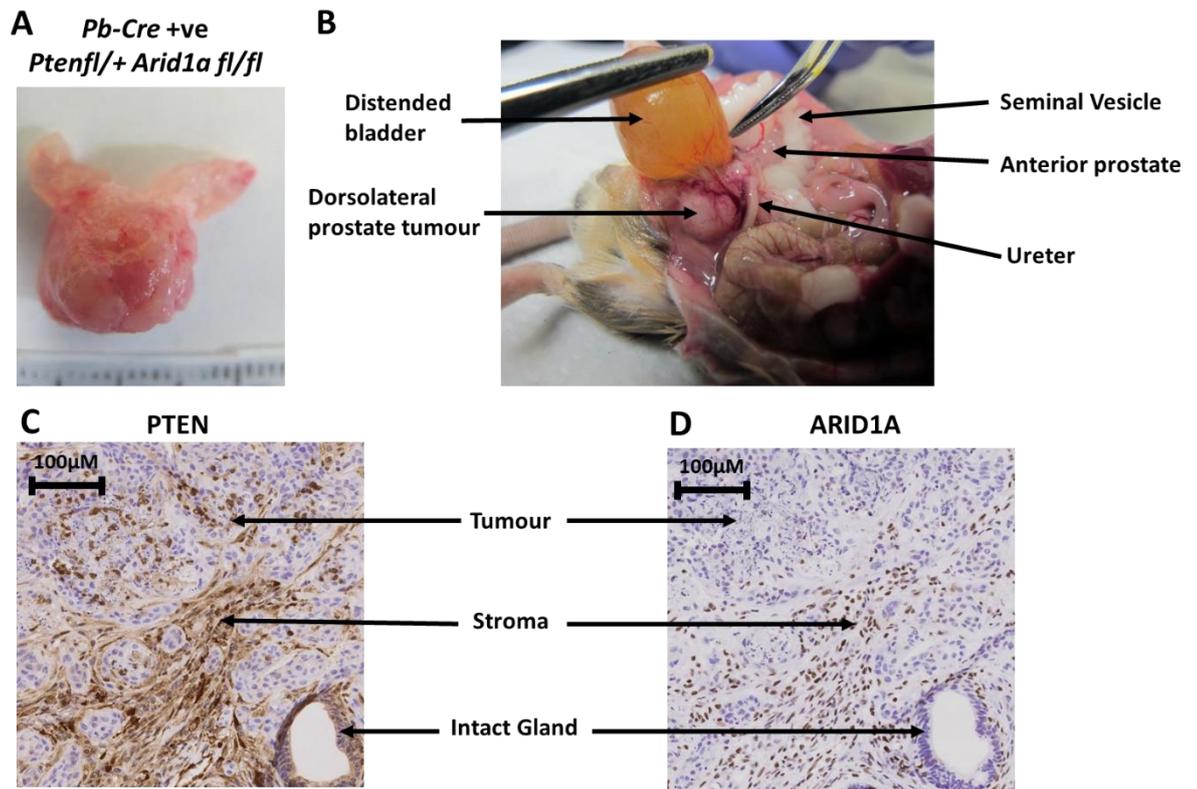


Figure 3.15 – Spontaneous loss of PTEN in *Pb-Cre+ve Pten^{fl/+} Arid1a^{fl/fl}* dorsolateral tumour

A. Representative tumour of *Pb-Cre+ve Pten^{fl/+} Arid1a^{fl/fl}* with spontaneous tumour development also shown in 3.13. B. In situ picture shows tumour and distended bladder at endpoint with key structures indicated. C. Immunohistochemistry for PTEN with structures labelled. D. Immunohistochemistry for ARID1A with structures labelled. *Pten^{fl/+} Arid1a^{fl/fl}* mouse shown in these images was 6 months old at clinical endpoint.

Due to the sporadic and spontaneous nature of the loss of other *Pten* allele in the *Pb-Cre+ve Pten^{fl/+} Arid1a^{fl/fl}* cohort combined with the aggressive tumours of the *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* genotype, breeding of this cohort was halted as we had 2 mice breaching clinical endpoint (being found dead with a prostate tumour).

None of the 18 mice in the *Pb-Cre+ve Pten^{+/+} Arid1a^{fl/fl}* cohort developed prostate tumour potentially indicating that the loss of *Arid1a* alone is insufficient to drive tumorigenesis and a complete loss of *Pten* is a prerequisite to *Arid1a* loss being tumorigenic. Alternatively, it may be that *Arid1a* requires another driver, not necessarily loss of *Pten*.

Overall, prostate weights remained comparable between each of these cohorts. *Pb-Cre+ve Pten^{fl/+}Arid1a^{fl/+}* has one mouse which developed a small prostate tumour, and *Pb-Cre+ve Pten^{fl/+}Arid1a^{fl/fl}* had five mice total with prostate tumours with only three of these sampled as the others were found-dead mice (indicated in red on the graph of Figure 3.16). Only dry weight was recorded as neither the normal prostate nor *Pb-Cre+ve Pten^{fl/+}Arid1a^{fl/fl}* contain cystic fluid.

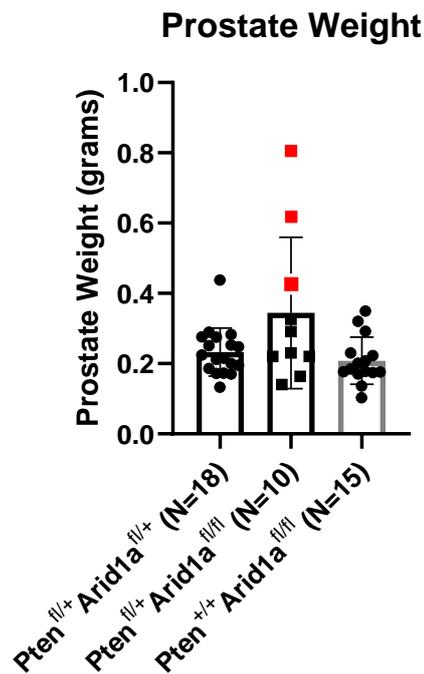


Figure 3.16 - Tumour weights in *Pb-Cre+ve Pten^{fl/+}* cohorts showing spontaneous loss of PTEN is required for tumour growth in *Arid1a^{fl/fl}* prostates

Graph shows tumour weights of indicated genotypes with each data point representing one mouse, *Pb-Cre+ve Pten^{fl/+} Arid1a^{fl/+}* (N=18). Mice of *Pb-Cre+ve Pten^{fl/+} Arid1a^{fl/fl}* (N=10) genotype in red indicate mice which developed a tumour and bladder distension endpoint. Those mice that were found-dead do not have a tumour weight. *Pb-Cre+ve Pten^{+/-} Arid1a^{fl/fl}* (N=15).

3.3.5 Characterising Systemic Effects of *Pb-Cre +ve Pten^{fl/fl} Arid1a^{fl/fl}* mice in blood biochemistry

The *Pb-Cre +ve Pten^{fl/fl} Arid1a^{fl/fl}* genotype produces aggressive and rapidly growing tumours of the dorsolateral prostate. Unfortunately, 5 mice as part of this cohort also died between the ages of 2 and 3.75 months old without an endpoint sized prostate tumour.

To characterise if these mice were experiencing any systemic effects because of this genotype, 10 mice aged between 1 and 3.75 months were taken to investigate if any changes occurred as these mice aged. The prostate morphology as these mice aged is shown in Figure 3.17 relative to a 12-week *Pb-Cre-ve* prostate control with an N=1 for each timepoint. These images allow a reference to prostate size relative to age. The 15-week HPAA18.1A represents the clinical endpoint for this genotype with bladder distension shown.

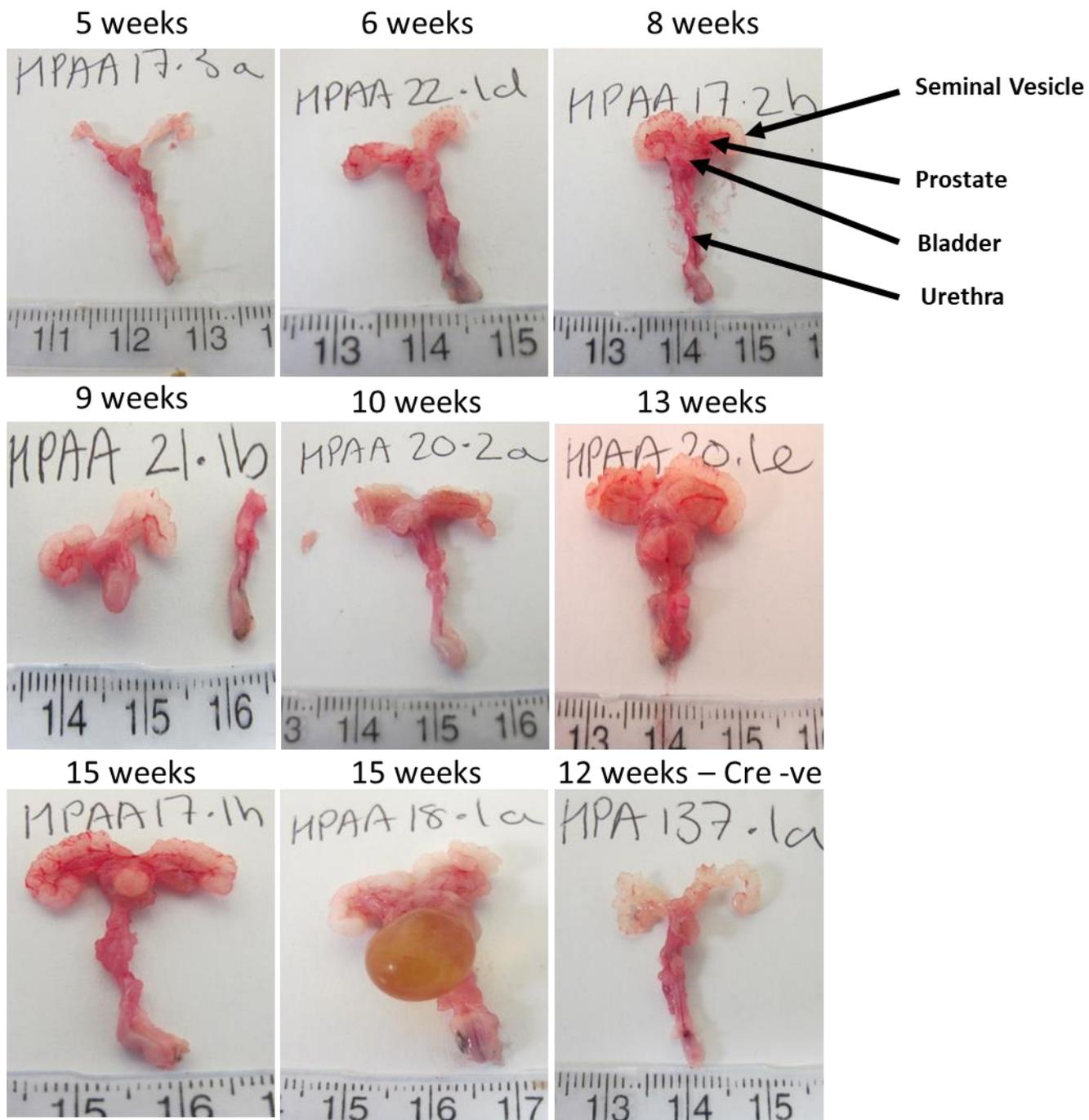


Figure 3.17 - *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* prostate morphology over time

Dissected *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* prostates of shown age in weeks with N=1 for each timepoint with centimetre scale at bottom of each image.. Example of key structures indicated in 8 week timepoint and representative throughout. 15 week sample (HPAA18.1A) shows clinical endpoint sample with bladder distension at endpoint. 12 week *Pb-Cre* -ve sample (HPA137.1a) shows wildtype mouse prostate.

Blood chemistry from these animals were analysed by the Veterinary Diagnostic Service, a veterinary pathology lab which can give a biochemical profile of the sample as well as red and white blood cell counts. With this, organ function and blood cell levels could be correlated with the age of the mice as the *Pb-Cre* system becomes active.

Haematology was assessed by 6 parameters by measuring red cell count, platelet count and haematocrit, as well as corpuscular attributes (Figure 3.18). Red cell count, haematocrit, and platelet count were all within the normal range, except for the endpoint mouse where the values were all lower. These lower values likely are as a result of the impact in kidney function damaging red blood cells. This is likely an acute event related to bladder distension as it is not observed in any of the other mice. Corpuscular haemoglobin, volume and concentration values all remain in normal range even in the endpoint mouse indicating that red cell production is not affected.

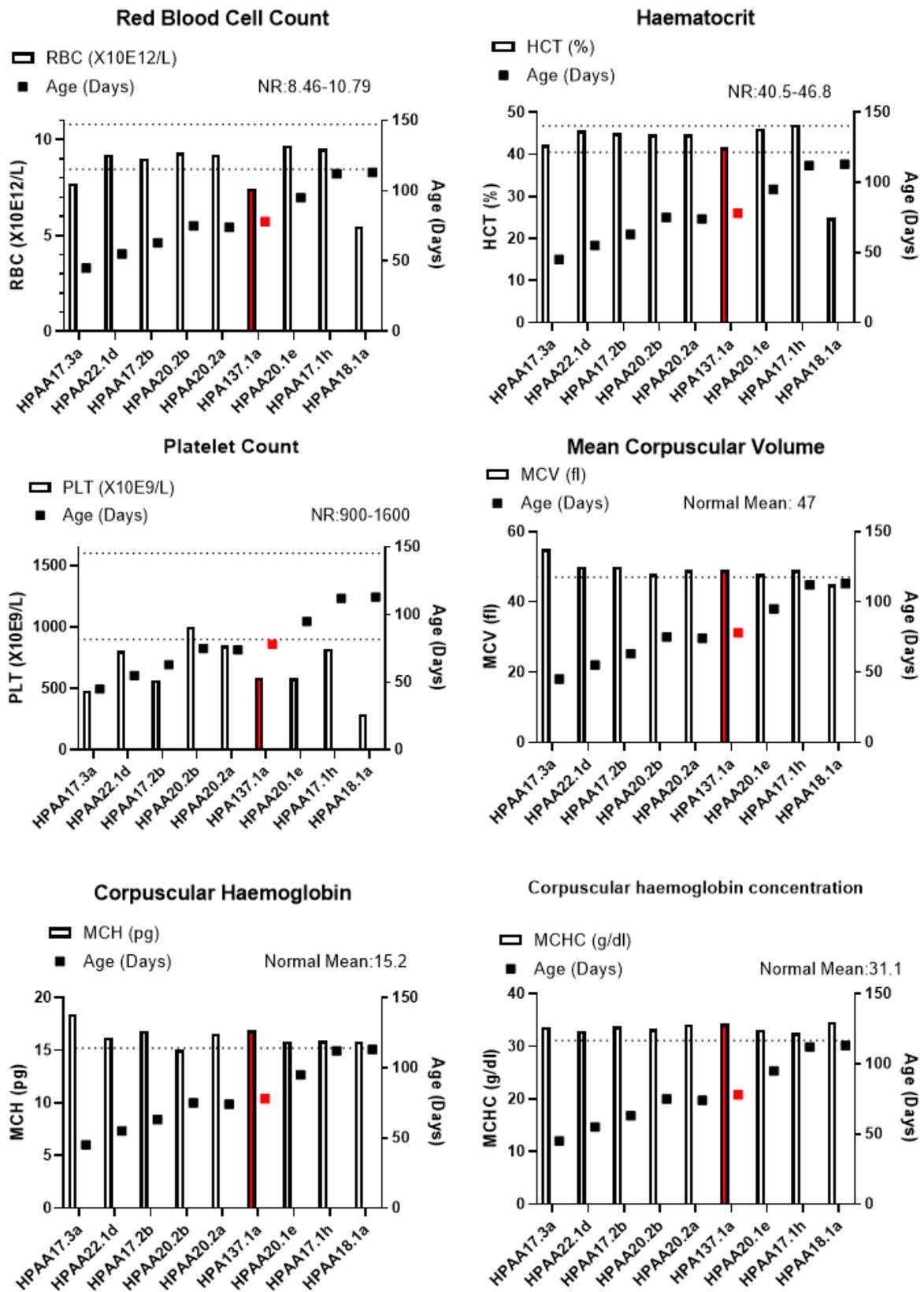


Figure 3.18 - Blood biochemistry describing haematology of *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* mice over time

Graphs show the indicated blood parameter of each of the *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* or *Pb-Cre-ve* mice. Mice are ordered left to right by age indicated by a square point with N=1 for each point. The indicated red mouse being *Pb-Cre-ve*, and HPAA18.1A has clinical endpoint tumour. Bars represent the indicated blood parameter compared to a normal range (NR) or normal mean represented by the dotted line obtained from (Otto, Rathkolb *et al.* 2016).

Metabolite testing was also performed on the blood samples. Liver function was assessed by measuring levels of 5 different parameters (Figure 3.19). Total bilirubin levels above the normal range indicate potential liver issue as the liver is unable to clear bilirubin. The older mice had elevated levels of total bilirubin however the *Pb-Cre-ve* mice had the highest levels indicating this is likely not related to genotype. The other 4 parameters directly examine levels of alkaline phosphatase, alanine aminotransferase, gamma-glutamyl transpeptidase, and aspartate aminotransferase. Elevated levels of these proteins indicate liver injury or problems as these proteins are leaked into the blood. Alkaline phosphatase and alanine aminotransferase were low however, which indicates none of the *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* had liver injury. Altogether this data indicates the *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* mice have intact liver function even at clinical endpoint.

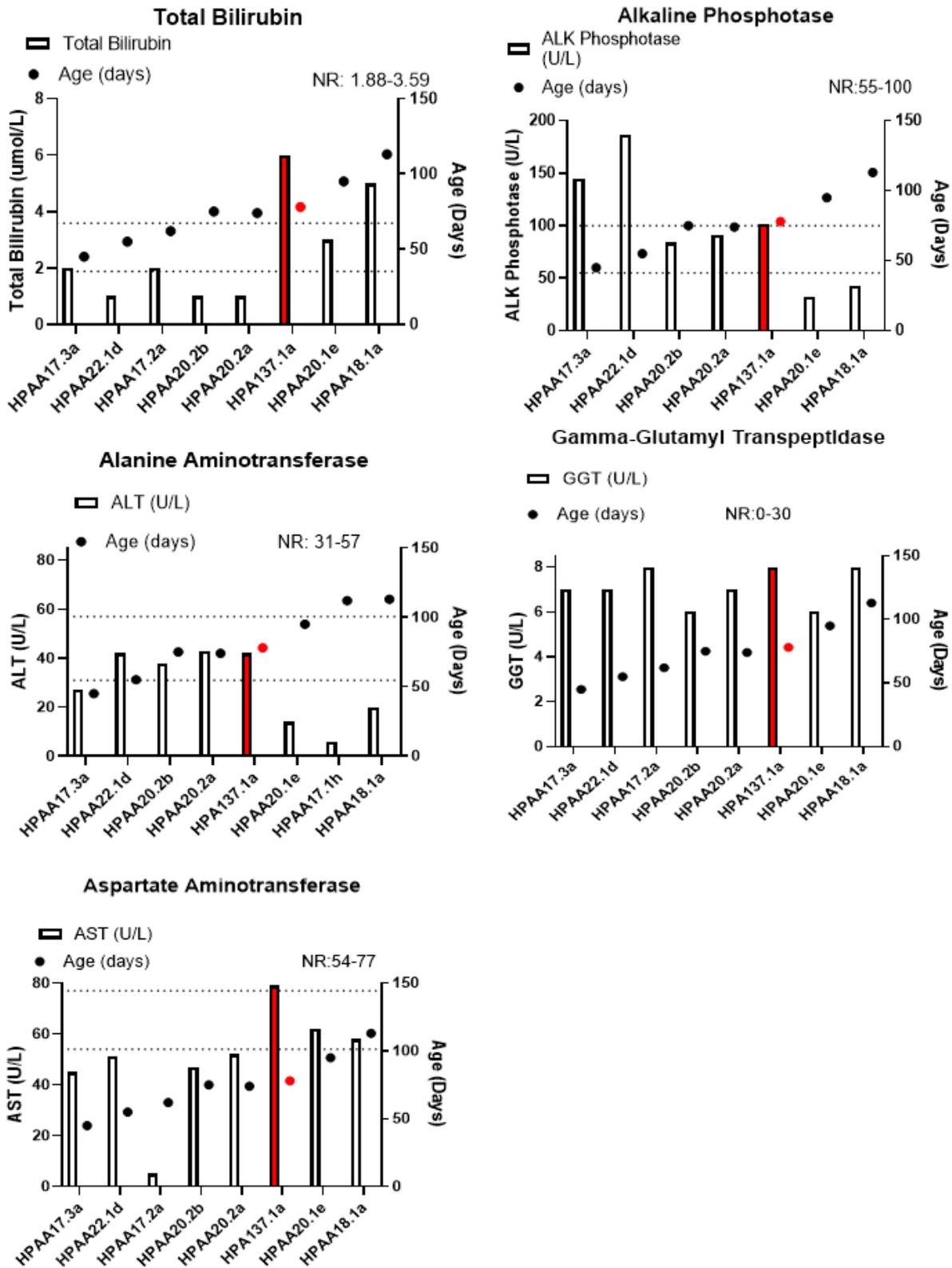


Figure 3.19 - Blood biochemistry indicating liver function of *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* mice over time

Graphs show the indicated liver function parameters of each of the *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* or *Pb-Cre-ve* mice. Mice are ordered left to right by age indicated by a square point with N=1 for each point. The indicated red mouse being *Pb-Cre-ve*, and HPAA18.1A. Bars represent the indicated liver function parameter compared to a normal range (NR) or normal mean represented by the dotted line obtained from (Otto, Rathkolb *et al.* 2016).

To assess kidney function, salt and urea levels in the blood were also measured (Figure 3.20). As bladder distension was the clinical endpoint for the *Pb-Cre* +ve *Pten^{fl/fl} Arid1a^{fl/fl}* genotype, it may have been expected to see either a gradual change in these values as the mice age, or an acute event specifically in the endpoint mice. Blood salt content were mostly in or around the normal range even in the endpoint mouse. Urea content however was higher in the endpoint mouse indicating there was stress on the kidneys which is concurrent with the bladder distension. This data indicates that while kidney function is affected at clinical endpoint it is an acute event and consistent with the tumour size. Meanwhile the other non-endpoint mice do not progressively lose kidney function. This likely indicates that only an endpoint sized tumour impacts kidney function. Therefore as the mice were found-dead without a prostate tumour, it is not likely related to their kidney function.

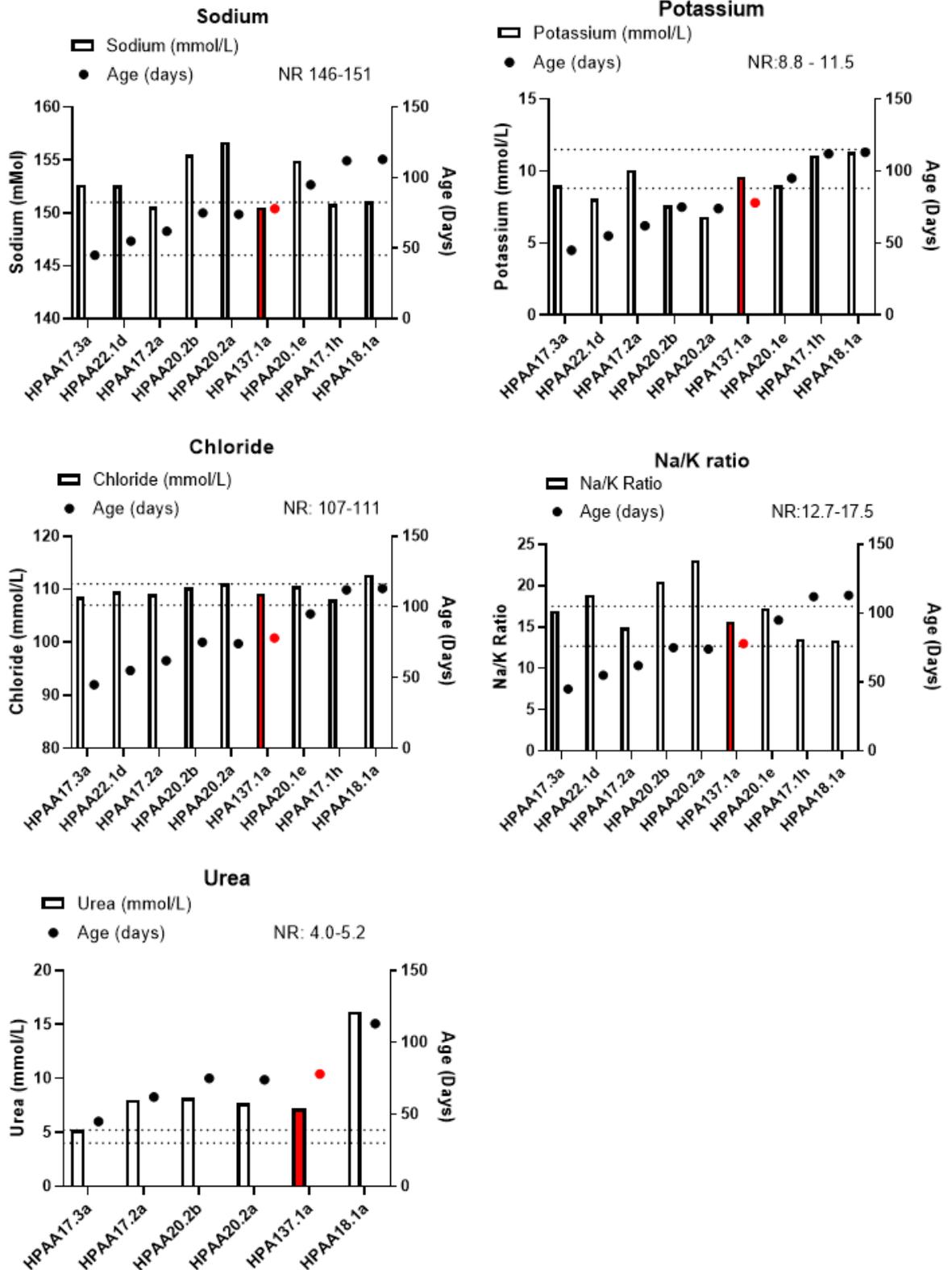


Figure 3.20 - Blood biochemistry indicating kidney function

Graphs show the indicated kidney function parameter of each of the *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* or *Pb-Cre-ve* mice. Mice are ordered left to right by age indicated by a square point with N=1 for each point. The indicated red mouse being *Pb-Cre-ve*, and HPAA18.1A having a clinical endpoint tumour. Bars represent the indicated kidney function parameter compared to a normal range (NR) or normal mean represented by the dotted line obtained from (Otto, Rathkolb *et al.* 2016).

To understand if there were any alterations in immune cell populations, immune cells were assessed by counts of total white cells, neutrophil, eosinophil, lymphocytes, and monocytes (Figure 3.21). Some mice lack these values due to insufficient blood volume during analysis. All mice with the exception of the endpoint mouse had similar white cell counts. The endpoint mouse had reduced overall white cell count and lymphocytes, while having elevated neutrophils. As these counts are only altered in the endpoint mouse this is likely a consequence of the tumour and distended bladder activating the immune system due to stress and damaging cells similar to the haematology findings.

Overall, the blood biochemistry of these mice indicates that there is no alteration in liver, kidney, red cells, or immune cells as these mice age. The endpoint mouse however has a decrease in kidney function likely because of the bladder distension. In turn, this likely has impacted red and white cell counts. Overall, this indicate that the 5 mice that were found dead without endpoint sized tumours likely did not have blood biochemistry because of their genotype.

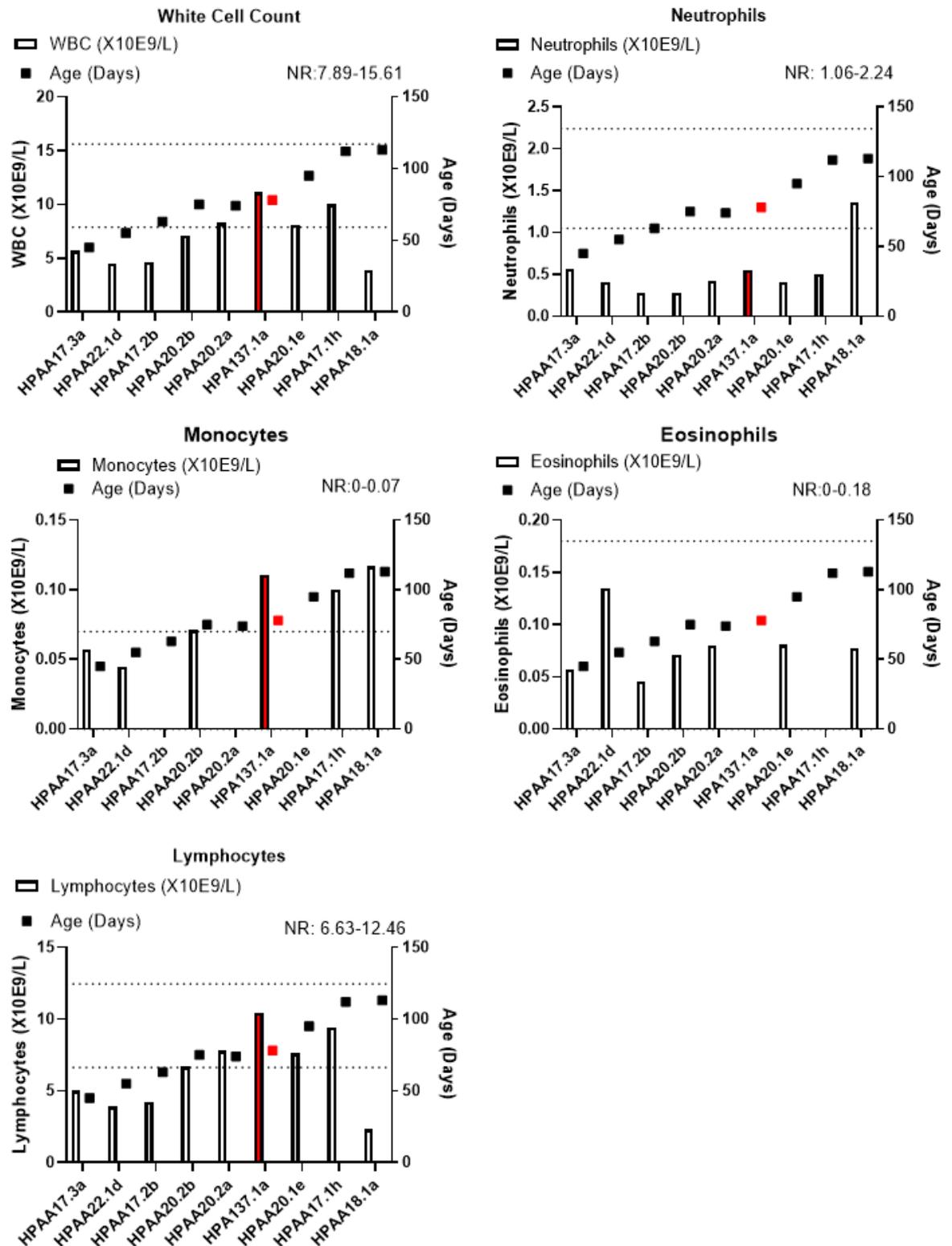


Figure 3.21 - Blood biochemistry indicating white cells function

Graphs show the indicated immune cell parameters of each of the *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* or *Pb-Cre-ve* mice. Mice are ordered left to right by age indicated by a square point with N=1 for each point. The indicated red mouse being *Pb-Cre-ve*, and HPAA18.1A having clinical endpoint tumour. Bars represent the indicated immune cell count parameter compared to a normal range (NR) or normal mean represented by the dotted line obtained from (Otto, Rathkolb *et al.* 2016).

3.3.6 Characterising Systemic Effects of *Pb-Cre +ve Pten^{fl/fl} Arid1a^{fl/fl}* mice with tissue staining for Arid1a

One other explanation for the *Pb-Cre +ve Pten^{fl/fl} Arid1a^{fl/fl}* genotype to have caused the 5 pre-endpoint dead mice could be due to an unexpected recombination in an off-target tissue which did not alter blood chemistry. This could have resulted in *Arid1a* being lost in unexpected tissues. Multiple organs including, testis, brain, liver, lymph node (L/N), heart, kidney, spleen, gut, and lung were all taken from these mice and stained for ARID1A to confirm there was no unintended recombination events in these other tissues (Figure 3.22, 3.23, 3.24).

These tissue stains show that ARID1A is present in each of the compartments in the *Pb-Cre +ve Pten^{fl/fl} Arid1a^{fl/fl}* genotype as in the *Pb-Cre -ve* control mouse. This data therefore indicates that *Arid1a* recombination did not occur in any of these harvested tissues and so did not contribute to the death of these mice.

In all, this data characterising the systemic effects of the *Pb-Cre +ve Pten^{fl/fl} Arid1a^{fl/fl}* mouse has not found evidence of unintended effects of the genotype either in systemic blood biochemistry or alternate tissue recombination. As we cannot determine the cause of death for the 5 pre-endpoint dead mice, we stopped breeding the *Pb-Cre +ve Pten^{fl/fl} Arid1a^{fl/fl}* genotype.

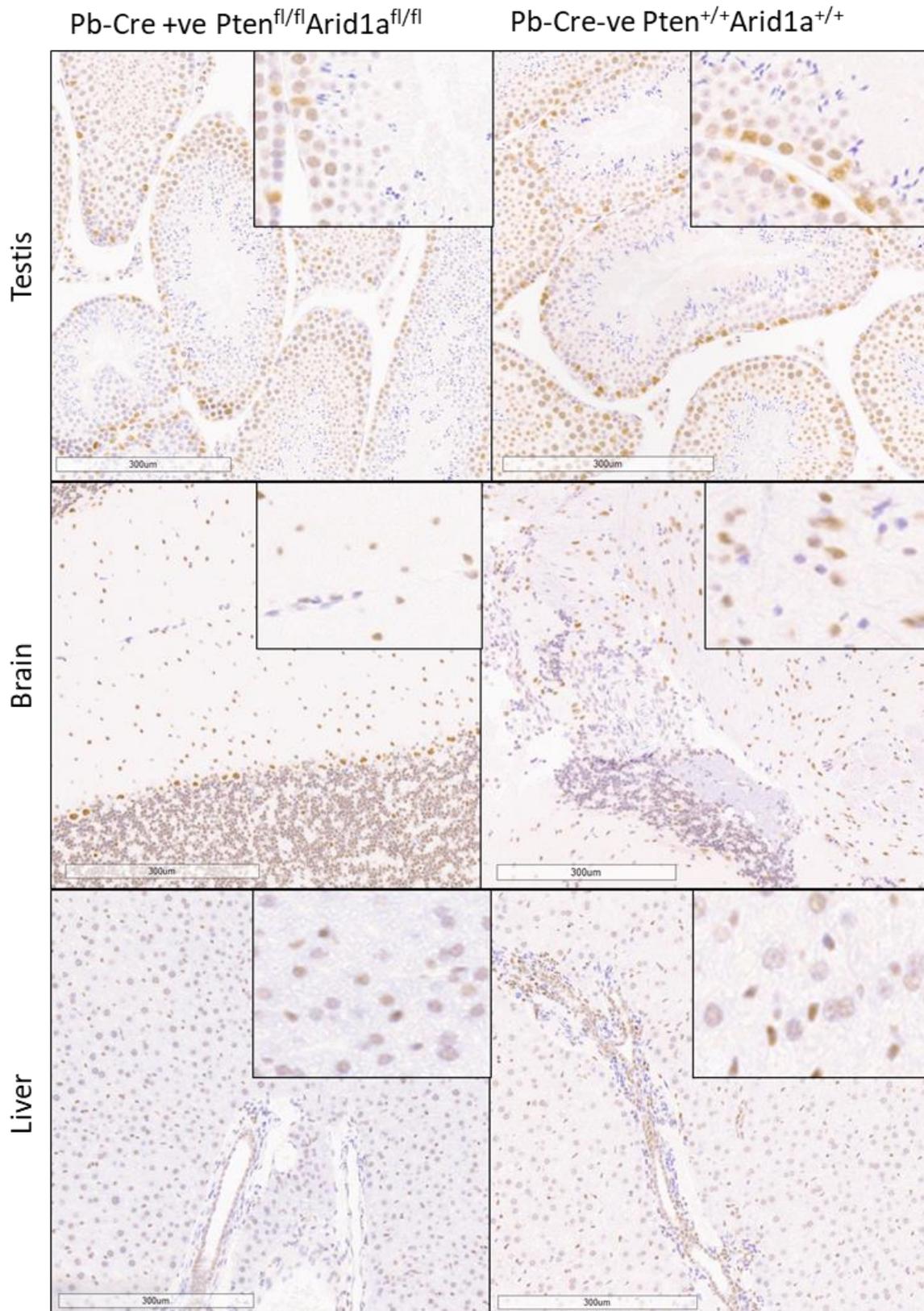


Figure 3.22 – ARID1A protein levels in testis, brain, and liver

Indicated tissues were stained in indicated genotypes for ARID1A representative of N=5. Presence of ARID1A nuclear signal in each instance will indicate ARID1A has not been lost from tissue. This shows there have been no off-target effects for the probasin cre-recombinase system.

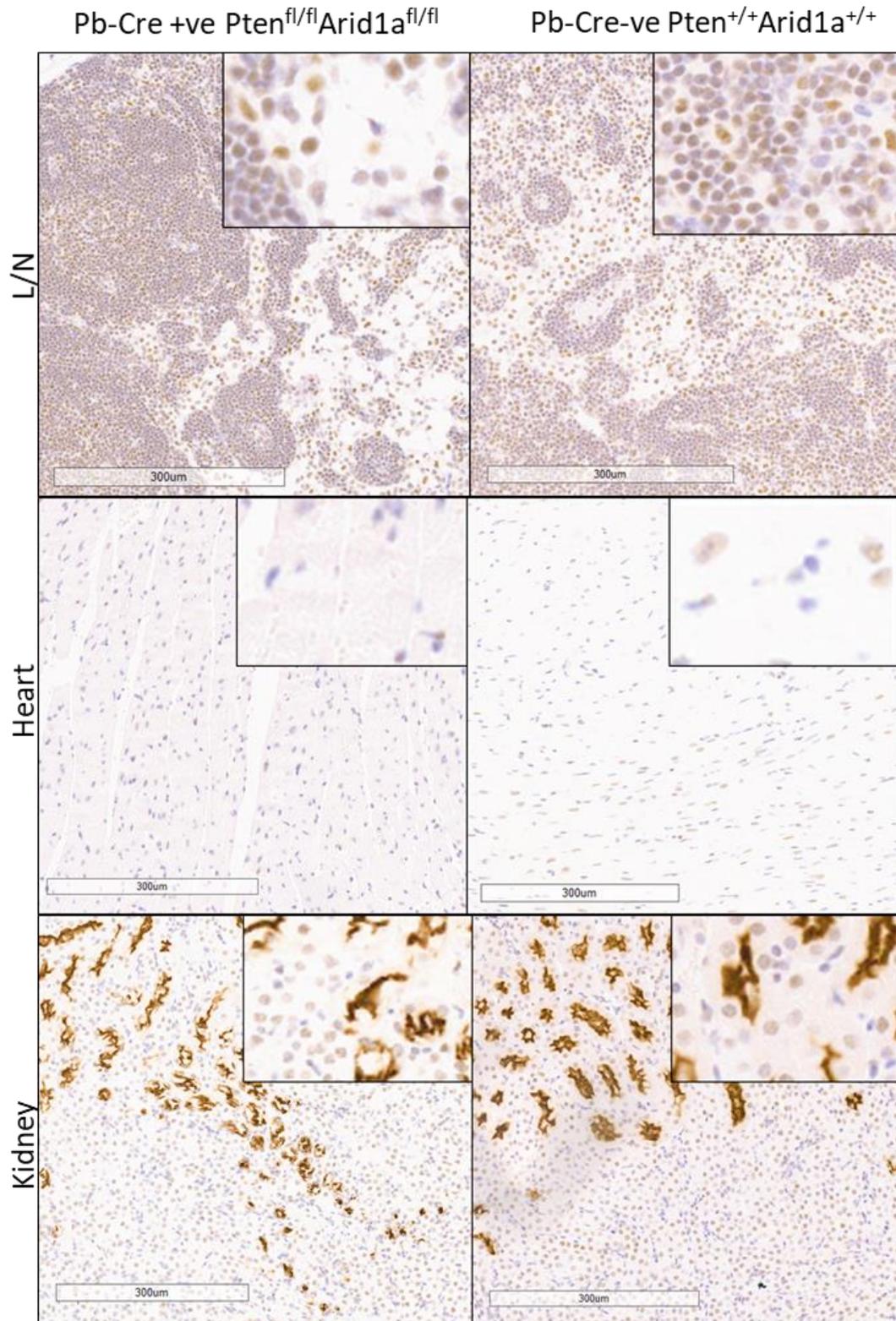


Figure 3.23 – ARID1A protein levels in lymph node (L/N), heart, and kidney

Indicated tissues were stained in indicated genotypes for ARID1A representative of N=5. Presence of ARID1A nuclear signal in each instance will indicate ARID1A has not been lost from tissue. This shows there have been no off-target effects for the *Probasin Cre-recombinase* system.

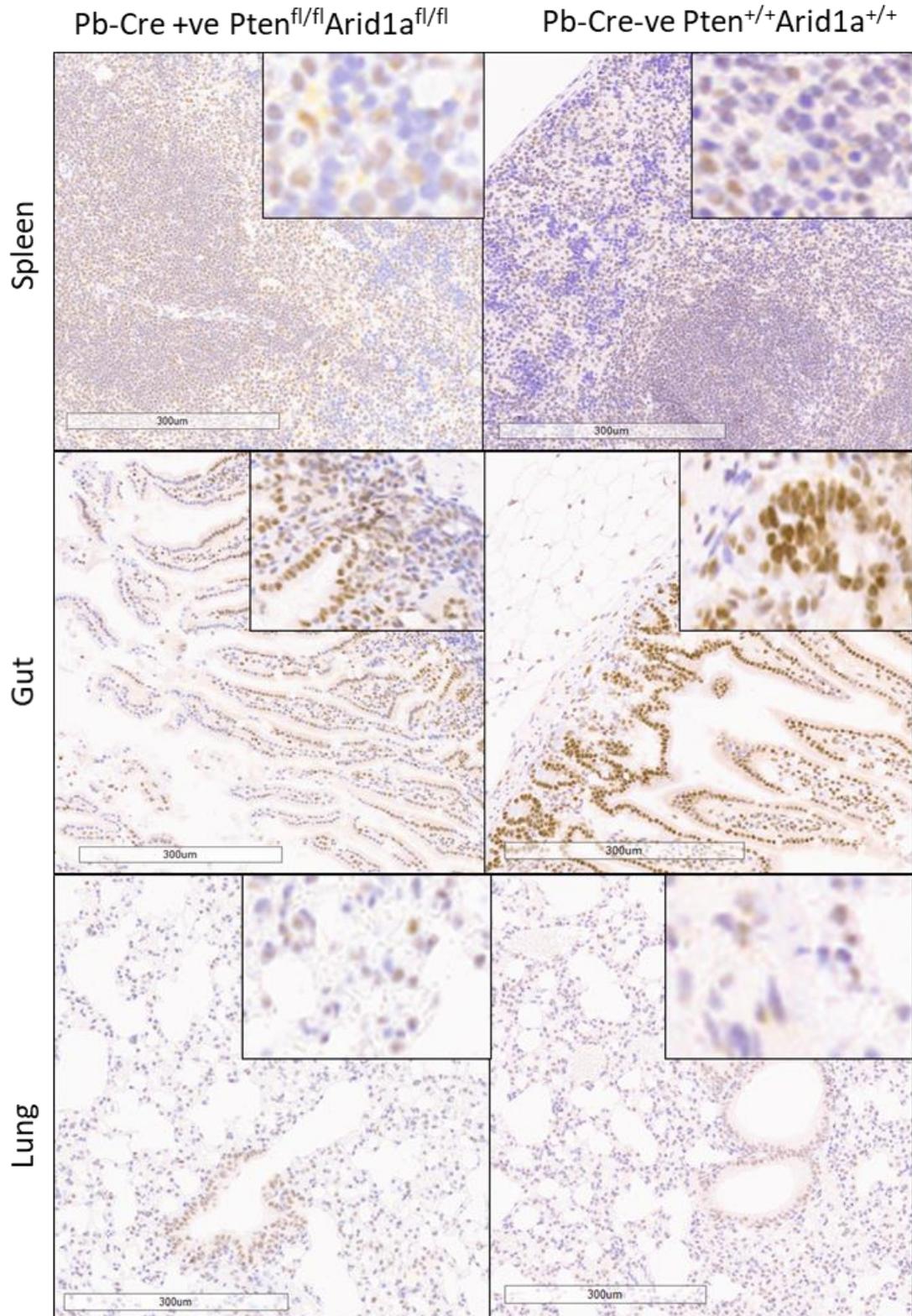


Figure 3.24 – ARID1A protein levels in spleen, gut, and lung

Indicated tissues were stained in indicated genotypes for ARID1A representative of N=5. Presence of ARID1A nuclear signal in each instance will indicate ARID1A has not been lost from tissue. This shows there have been no off-target effects for the probasin cre-recombinase system.

3.4 Discussion

3.4.1 ARID1A and PTEN in PI3K-AKT Signalling

Most of the research regarding ARID1A originates from work on gynaecological cancers, particularly ovarian clear cell carcinoma (OCCC) where the mutation rate of ARID1A is highest, in around 57% of cases (Jones, Wang *et al.* 2010). Of these cases, ARID1A mutations were frequently found to co-occur with activating mutations in the PI3K/AKT pathway, such as those causing loss of PTEN, or gain of function in PIK3CA (Bosse, ter Haar *et al.* 2013, Cancer Genome Atlas Research, Kandoth *et al.* 2013). In ovarian cancers, it has been shown that loss of ARID1A correlates with an increase in phosphorylation of AKT (Wiegand, Hennessy *et al.* 2014). However, loss of ARID1A alone with intact PIK3CA and PTEN does not impact pAKT levels.

ARID1A has been shown to act on the PI3K/AKT pathway in two different ways. In ovarian cell lines, ARID1A and enhancer of zeste homology 2 (EZH2) were shown to positively regulate PIK3IP1 (Bitler, Aird *et al.* 2015). PIK3IP1 can bind to p110, the catalytic subunit of PI3K, thereby reducing PI3K activity (He, Zhu *et al.* 2008). This occurs by PIK3IP1 having similar homology to the p85 regulatory subunit of PI3K. Secondly, in breast cancer, ARID1A has been shown to negatively regulate ANXA1 which increased phosphorylation of AKT mediated drug resistance (Berns, Sonnenblick *et al.* 2016). These two systems demonstrate two different mechanisms of ARID1A regulating PI3K/AKT.

However, it is worth noting that ARID1A loss does not predict synthetic lethality with PI3K/AKT inhibition. ARID1A status alone is likely a poor predictor of PI3K inhibition efficacy. While ARID1A loss is often observed with PI3K pathway

aberrations, and it certainly appears to synergise with these mutations, ARID1A loss alone has consistently been shown to have little impact on downstream P-AKT (Huang, Lin *et al.* 2014, Wiegand, Hennessy *et al.* 2014).

This shows that our mouse data finding that loss *Arid1a* cooperates with loss of *Pten* to drive prostate tumour growth is consistent with other published data. These findings show that loss of *Arid1a* synergises with *Pten* loss to cause rapid tumour development. Such an effect, however, was not observed without complete *Pten* loss indicating that *Pten* loss is required before loss of *Arid1a* loss is tumourigenic. However, this type of synergism as observed in OCCC is likely due to a *Pten* loss and *Arid1a* loss converging on PI3K/AKT signalling. As demonstrated by phosphor-Ser473 on AKT, there is no hyperactivation of this pathway following the loss of *Arid1a* (Figure 3.7). This may indicate that *Arid1a* and *Pten* loss cooperate in a different way than observed in these other cancers.

3.4.2 ARID1A in murine cancer models

3.4.2.1 *Arid1a* and *Pten* in gynaecological cancers

ARID1A is mutated in gynaecological cancers more than any other cancer type, particularly in combination with PI3K/AKT pathway mutations such as PTEN (Jones, Wang *et al.* 2010). To date, all mouse models investigating the interaction of *Arid1a* and *Pten* have been in these endometrial cancers. This interaction was first investigated in 2014 by Guan *et al.* who showed that loss of either *Pten* or *Arid1a* was insufficient to cause ovarian cancer, a double-knockout of both *Pten* and *Arid1a* produced tumour development or hyperplasia in all mice (Guan, Rahmanto *et al.* 2014). This phenotype was later explored by Zhai *et al.* (2016) who showed that *Arid1a* loss did not cooperate with *Apc* loss

(Zhai, Kuick *et al.* 2016). This likely is due to the requirement of *Arid1a* for mediating β -catenin signalling (Mathur, Alver *et al.* 2017). This was most recently investigated by Rahmanto *et al.* in 2020 who showed that loss of *Arid1a* led to a loss of transforming growth factor-beta (TGF- β) tumour suppressor activity due to genome-wide transcriptional reprogramming of a *Pten* null cell (Suryo Rahmanto, Shen *et al.* 2020).

These studies show similarities to our PC murine model, with a similar inability for *Arid1a* alone to drive tumorigenesis without loss of *Pten* or other PI3K/AKT driving mutation.

3.4.2.2 *Arid1a* in other murine cancer models

In a murine model of colorectal cancer, loss of *Arid1a* alone is sufficient to drive an invasive colon adenocarcinoma (Mathur, Alver *et al.* 2017). As previously mentioned, this loss of *Arid1a* did not cooperate with *Apc* loss which typically also drives colorectal carcinomas through β -catenin signalling. This was also investigated by Hiramatsu *et al.* in 2019 who showed a similar phenotype with loss of *Arid1a* driving an invasive carcinoma (Hiramatsu, Fukuda *et al.* 2019). They also showed that loss of *Arid1a* and subsequent β -catenin signalling impaired stem cell maintenance together with a loss of Wnt signalling. These findings also demonstrate a double-edged ability of *Arid1a* to mediate potentially oncogenic β -catenin signalling while its loss can also mediate tumorigenesis.

This is also clearly demonstrated in hepatocellular carcinoma (HCC), where *Arid1a* loss in early disease hinders tumour initiation by impairing reactive oxygen species (ROS) formation (Sun, Tawfik *et al.* 2007). However, in an

established tumour, either heterozygous or homozygous *Arid1a* loss could drive metastasis and advanced disease through altered cell cycle and migration signalling.

This is not observed in our PC murine model, as we show loss of *Arid1a* with *Pten* is immediately oncogenic without any impact on tumour initiation. This indicates that our *Arid1a* function may be more like endometrial cancers than colorectal or HCC models, potentially showing more similarities between hormone-driven cancers.

3.4.3 Methodological changes and future directions

There are still many unknowns that must be investigated about how *Arid1a* drives PC. The first is a full understanding of *Arid1a* in driving advanced stages of prostate cancer, namely metastasis, as this current model replicates an invasive early disease. Unfortunately, with the *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* tumours arising in the dorsolateral prostate and that the tumours can be fatal even at a very small size, we would be unable to lengthen investigations with this genotype. One option to slow PC growth clinically is ADT which removes androgen to slow prostate cell growth (Gunner, Gulamhusein *et al.* 2016). However, as *Pb-Cre* is activated by AR signalling in response to androgen, the mice would need to progress through puberty prior to castration to allow gene floxing to occur. This typically occurs within 2-2.5 months and could be considered an intervention step to slow growth of a *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* tumour.

Investigating how *Arid1a* interacts with other genes may also be a consideration. In keeping with the idea of *Arid1a* and PI3K signalling, we could combine *Arid1a*

with a PI3K activating mutation such as *Pik3ca*^{H1047R} (Hare, Phesse *et al.* 2014). Alternatively, *Arid1a* could be combined with oncogenes such as *Kras* or *Myc* to investigate combining *Arid1a* loss with other drivers (Clegg, Couto *et al.* 2011, Aytes, Mitrofanova *et al.* 2013). These models would allow us to build a more complete picture of *Arid1a* in PC. However, in line with the 3Rs (Replace, Reduce, Refine), these methods could also be performed on cell lines initially to investigate the phenotype and only the most interesting validated in a mouse model.

3.4.4 Conclusions

In this chapter, I have demonstrated that *Arid1a* can function as a tumour suppressor gene in prostate cancer in combination with loss of *Pten*. This homozygous loss of both genes produces aggressive and locally invasive cancer to develop from the dorsolateral glands of the murine prostate. Additional work will focus on understanding the mechanism by which *Arid1a* loss drives tumorigenesis in the murine model and investigating the role of ARID1A in established human PC cell lines.

4 Loss of *Arid1a* upregulates cell cycle and migration signalling in *Pten* null mice

4.1 Introduction

4.1.1 ARID1A as a tumour suppressor in cancer

ARID1A is often a tumour suppressor in cancer, exemplified by it being the most mutated epigenetic regulator in cancer (Wu and Roberts 2013). However, the specific way that it is a tumour suppressor is context-dependent and therefore different between cancers.

Loss of ARID1A dysregulates chromatin organisation around specific genes leading to their up or down regulation depending on if chromatin was opened or closed. These genes therefore often confer a growth or survival advantage to the cancer if ARID1A loss is tumorigenic. However, the genes or at least the main effector genes, are context-dependent. As mentioned, ARID1A has no DNA-binding specificity therefore its 'target genes' are dependent on the transcription factors in the system. Because of this, the mechanism of ARID1A's tumour suppressor activity is often reported as being very different in each of the different cancers.

As mentioned, in OCCC ARID1A loss cooperates with PI3K-AKT signalling to produce hyperactivation of that growth signalling pathway (Wiegand, Hennessy *et al.* 2014). In pancreatic intraepithelial neoplasm (PanIN), loss of ARID1A attenuates Kirsten Rat Sarcoma Virus (KRAS)-induced senescence allowing cancer progression (Liu, Cao *et al.* 2021). In colorectal cancer, ARID1A loss often co-occurs with defects in DNA-damage machinery leading to high microsatellite instability (MSI) tumours along with other classic oncogenes like v-Raf murine sarcoma viral oncogene homolog B (BRAF) and Catenin beta-1 (CTNNB1) (Tokunaga, Xiu *et al.* 2020). In gastric cancer, ARID1A loss downregulates E-

cadherin causing an increase in migration and invasion (Yan, Wang *et al.* 2013). These examples show the variety of ways in which ARID1A loss can support tumorigenesis, always by altering unique gene sets in each of the different cancers. To understand how ARID1A loss may drive a cancer, it is required to understand which genes are under its control and how those genes are altered following its loss. ARID1A loss often accentuates oncogenic pathways which are normally found in a tumour. In the example of OCCC, it causes hyperactivation of an already mutated PI3K-AKT pathway, and in PanIN it cooperates with already mutated KRAS. Each of these pathways being already highly mutated in their respective cancer. As such, to understand how *Arid1a* loss may drive *Pten* null PC, we need to understand the signalling pathways endogenous to *Pten* null tumours.

4.1.2 Senescence in *Pb-Cre+ve Pten^{fl/fl}* mice

Pten loss in the prostate alone *in vivo* is characterised as creating large, cystic tumours which often remain non-invasive. One explanation for the inability for these cancers to fully transform into an invasive cancer is due to cellular senescence which opposes this transformation (Chen, Trotman *et al.* 2005). Senescence is a cellular process which leads to a change in many aspects of a cell including reduced proliferation potential, inhibition of apoptosis, and metabolic rewiring (Wyld, Bellantuono *et al.* 2020).

In *Pb-cre+ve Pten^{fl/fl}* mice, this senescence is a p53-dependent process that limits tumour growth following *Pten* loss. This type of senescence was later shown to be caused by the enhanced proliferative signalling of p-AKT leading to elevated replicative stress which triggers a sustained DNA-damage response

which stabilises P53 (Parisotto, Grelet *et al.* 2018). As such, *Pb-cre+ve Pten^{fl/fl} Trp53^{fl/fl}* mice avoid this induced senescence and develop invasive carcinoma within 1 month of age with an endpoint typically around 7 months (Chen, Trotman *et al.* 2005).

However, overcoming P53-dependent senescence does not require loss of P53. Upregulating other oncogenic signalling pathways can bypass the induced senescence of *Pten* loss. One such oncogenic pathway includes overexpression of human epidermal growth factor receptor 2 (HER2), an RTK which is mutated in 20% of PC (Saal, Johansson *et al.* 2007). Upregulation of HER2 leads to a signalling cascade which elevates MAPK signalling which leads to abrogation of the senescence signalling (Ahmad, Patel *et al.* 2011).

4.2 Hypothesis and Aims

As discussed, loss of ARID1A can drive tumorigenesis in a variety of ways. To understand these changes, we need to understand where ARID1A binds and how those genes are dysregulated following ARID1A's loss. As *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* mice develop aggressive rapidly growing tumours, we may expect changes in migration and growth signalling. In combination with deleted *Pten*, *Arid1a* loss may alter other known aspects of *Pten* deficient tumours such as their senescent phenotype.

The aims of this chapter are as follow:

- Perform and analyse RNA-Seq of *Pb-Cre+ve Pten^{fl/fl} Arid1a^{+/+}* vs *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* to understand signalling pathways and networks dysregulated following loss of *Arid1a* in PC
- Perform and analyse chromatin immunoprecipitation (ChIP)-Seq to understand *Arid1a* 'target genes' in a *Pb-Cre+ve Pten^{fl/fl}* mouse model
- Overlay ChIP-Seq and RNA-Seq datasets to understand which signalling pathways were altered as a result in changes in *Arid1a* regulation.
- Understand the changes in key oncogenic signalling pathways in *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* tumours

4.3 Results

4.3.1 RNA-Sequencing shows upregulated cell cycle signalling following *Arid1a* loss

To investigate how loss of *Arid1a* drove tumorigenesis, and which pathways and genes may be altered in different genotypes, RNA-Sequencing (RNA-Seq) was performed across all *Arid1a* genotypes in a *Pb-Cre+ve Pten^{fl/fl}* background. This should identify how loss of one or both copies of *Arid1a* would impact gene expression.

RNA was extracted from three mouse prostate tumours of *Pb-Cre+ve Pten^{fl/fl} Arid1a^{+/+}*, *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/+}*, and *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* genotypes and processed into a library. The quality was assessed on tapestation (Figure 4.1A) to validate the library generation. This gel shows a single large band in each of the samples indicating a uniform library was produced. These samples were then sequenced and variance analysed by principle component analysis (PCA) (Figure 4.1B). Analysis was performed by Mr Robin Shaw (Informatician, Beatson Institute for Cancer Research). *Pb-Cre+ve Pten^{fl/fl} Arid1a^{+/+}* and *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/+}* tumours clustered together indicating a large amount of similarity between these two cohorts. *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* however clustered separately from the other two cohorts indicating a larger difference compared to control.

The number genes significantly altered in each of the experimental cohorts compared to *Pb-Cre+ve Pten^{fl/fl} Arid1a^{+/+}* is shown in Figure 4.1C. Significant genes were defined as having P-value <0.05 and 1.5-fold change compared to control. This data shows that the *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/+}* cohort only has 183 genes significantly altered compared to the *Arid1a* wild type control, while

Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl} has 1540 genes altered. Of the 183 altered genes in the *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/+}* cohort, 132 are shared with the *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* cohort. This indicates that this loss of the second copy of *Arid1a* is having a much larger impact on gene expression as expected given the phenotype of these cohorts. This indicates there is not a haploinsufficient effect and *Arid1a* functions like other tumour suppressors and requires both copies to be lost to have a pro-tumorigenic effect.

The fold change in the significant genes was plotted as a volcano plot to understand if there was a bias to gene upregulation or downregulation following loss of *Arid1a* (Figure 4.1D and 4.1E). In both experimental cohorts, gene downregulation occurred more often than upregulation, however this is more prominent in the *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* cohort. The ratio of upregulation to downregulation is 1:1.4 and 1:2.9 for *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/+}* and *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* cohorts respectively. This indicates that loss of *Arid1a* causes more genes to be downregulated than upregulated, potentially indicating that *Arid1a* opens chromatin more than it closes chromatin. This is consistent with what has been previously observed with ARID-containing proteins, with ARID1B and ARID2 preferentially closing chromatin and ARID1A preferentially opening chromatin (Raab, Resnick *et al.* 2015).

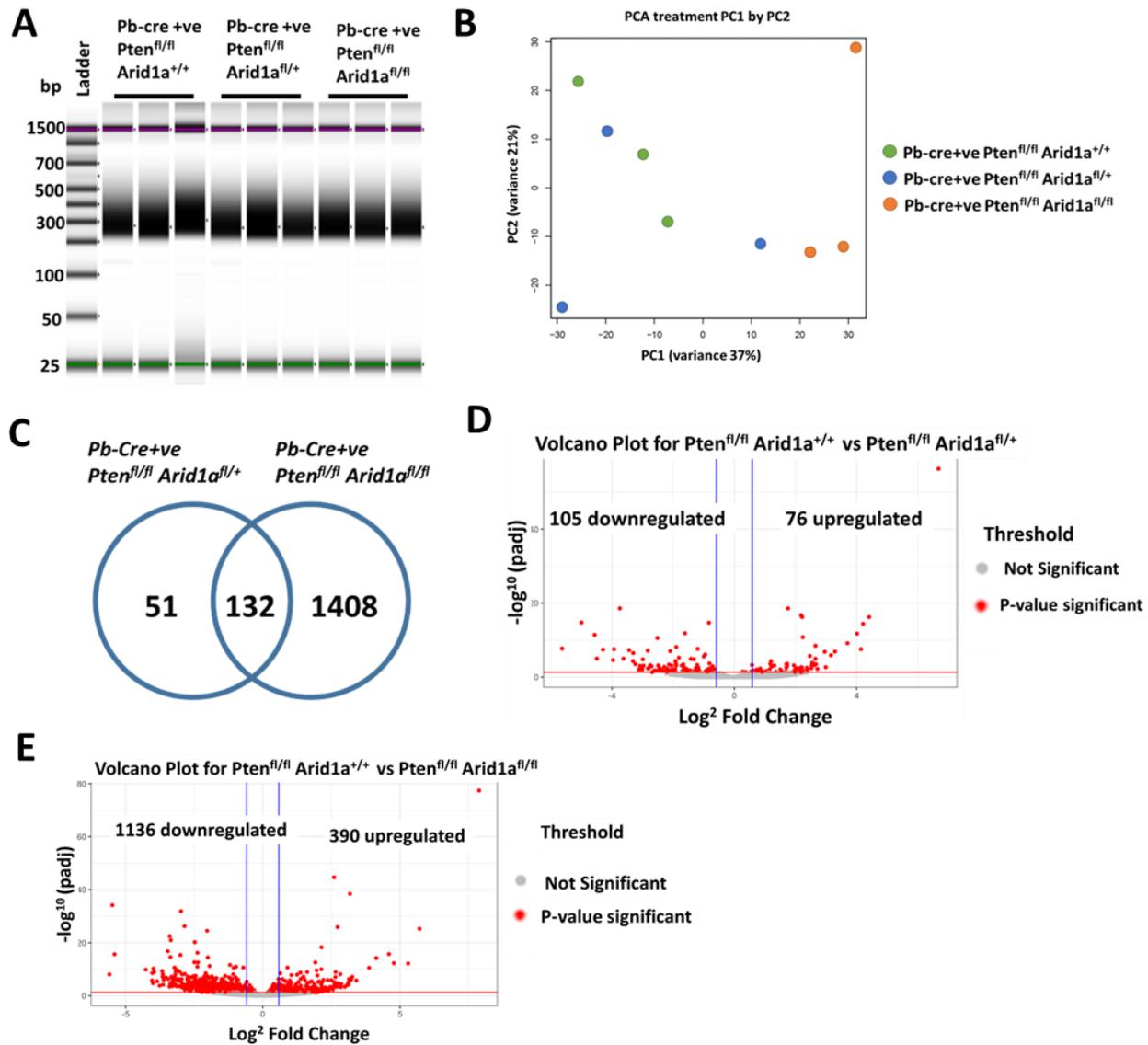


Figure 4.1 Quality assessment for RNA-Seq

A. RNA quality assessed on tapestation gel. First lane shows DNA ladder with sizes indicated in basepairs (bp). **B.** Principal component analysis (PCA) plot shows variance between each sample. **C.** Venn-diagram shows significantly altered genes each genotype compared to *Pb-Cre+ve Pten^{fl/fl} Arid1a^{+/+}*. **D.** Volcano plot of significantly altered genes when comparing *Pb-Cre+ve Pten^{fl/fl} Arid1a^{+/+}* to *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/+}*. Threshold of 1.5 with p-value <0.05. **E.** Volcano plot of significantly altered genes when comparing *Pb-Cre+ve Pten^{fl/fl} Arid1a^{+/+}* to *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}*. Threshold of 1.5 with p-value <0.05.

4.3.2 Loss of one copy of *Arid1a* does not highly dysregulated *Pten^{fl/fl}* prostate tumour signalling

The significantly enriched genes from the RNA-Sequencing of *Pb-Cre +ve Pten^{fl/fl} Arid1a^{fl/+}* tumours were imported into Metacore, a data analysis software, to

investigate whether the genes were part of similar pathways or signalling networks.

Of the most upregulated networks and pathways there did not appear to be any consensus. Some unusual networks such as 'Platelet aggregation' and 'Muscle Contraction' appear to be highly upregulated for instance. As there are more genes downregulated than upregulated in this dataset, there is a greater degree of significance of these networks and pathways. Likewise, as the *Pb-Cre +ve Pten^{fl/fl} Arid1a^{fl/+}* mice do not have a strong phenotype we would not necessarily expect highly upregulated pathways.

However, interestingly the most upregulated network was 'Signal transduction androgen receptor nuclear signalling'. This network may have some biological relevance, however there were no upregulated pathways showing a meaningful impact on androgen receptor signalling. This indicates that while genes that promote AR signalling were upregulated, AR signalling itself was not upregulated.

Similarly, there was little consensus in the most downregulated networks of this genotype although as mentioned the degree of significance is higher due to more genes being downregulated. However, of the most downregulated pathways there was some consensus with many pathways involving the themes of lipid metabolism. These pathways focus primarily on lipoproteins which bind lipids. The functional relevance of this is unclear as this effect does not appear to strongly impact the phenotype of these tumours. However, this may have an impact in a castration setting where lipid and cholesterol metabolism become more important(Škara, Huđek Turković et al. 2021).

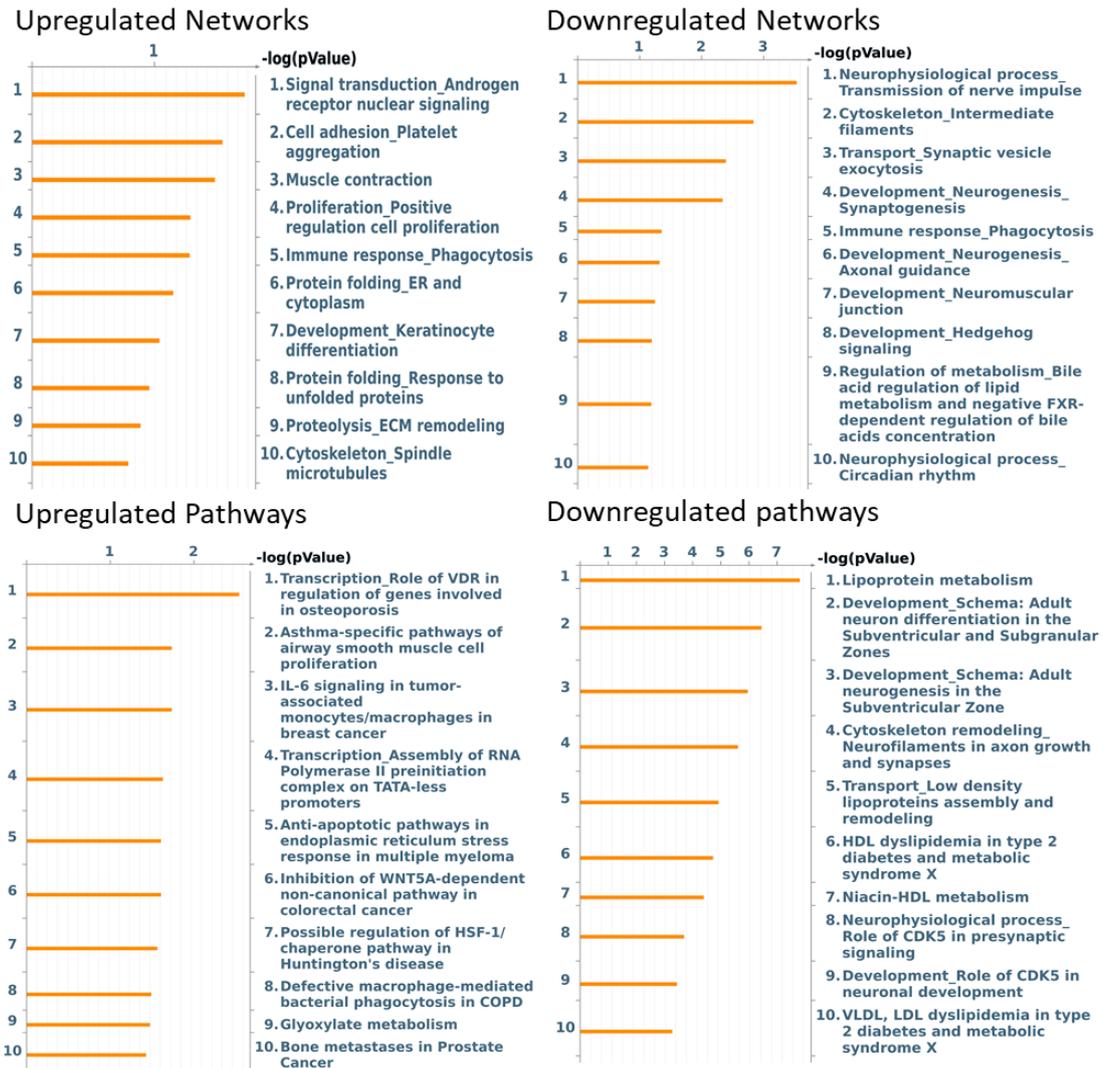


Figure 4.2 – Network and Pathway analysis of *Pb-Cre+ve Pten^{fl/fl} Arid1a^{+/-}* vs *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* prostate tumours

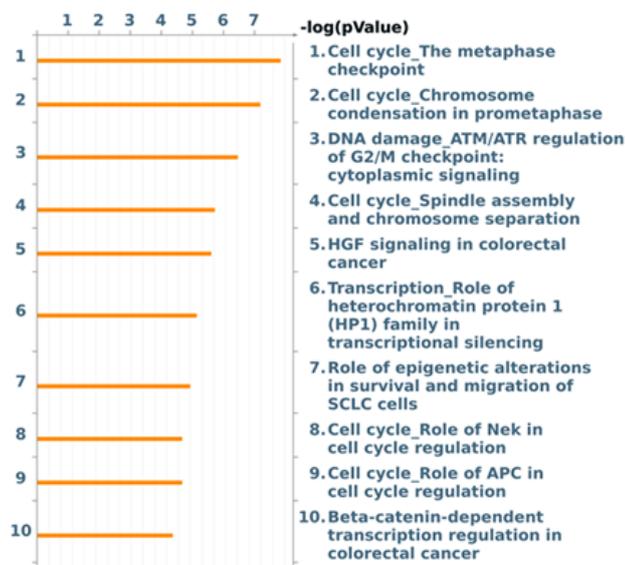
Significantly altered genes between *Pb-Cre+ve Pten^{fl/fl} Arid1a^{+/-}* and *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* tumours were uploaded to Metacore and pathway and network analysis was performed. Genes were filtered by threshold of 1.5 and P-value <0.05.

4.3.3 Loss of *Arid1a* upregulates cell cycle signalling in *Pten^{fl/fl}* prostate tumours

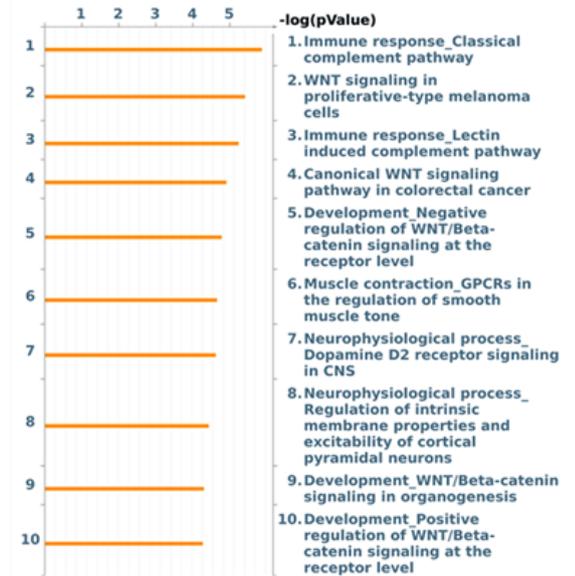
The significantly enriched genes from the RNA-Sequencing in the *Arid1a*-deleted tumours compared to the *Arid1a*-proficient tumours were also imported into Metacore to investigate whether the genes were part of similar pathways or signalling networks.

As the *Pb-Cre +ve Pten^{fl/fl} Arid1a^{fl/fl}* mice had a strong tumorigenic phenotype they also had many more genes significantly altered compared to the *Pb-Cre +ve Pten^{fl/fl} Arid1a^{fl/+}* mice as expected. Many of the upregulated pathways and networks showed an increase in cell cycle signalling, particularly signalling around the G2/M phase checkpoint rather than G1/S phase. 7/10 upregulated pathways and 6/10 upregulated networks relate to changes in cell cycle (Figure 4-3). Key G2/M phase regulators such as Aurora Kinase A (AURKA), Polo-like kinase 1 (PLK1), and cyclin dependent kinase 1 (CDK1) were upregulated indicating the loss of *Arid1a* leads to an elevation in cell cycle signalling. The most downregulated pathways showed an unexpected consensus around neural signalling with 7/10 having some relation to neurophysiology.

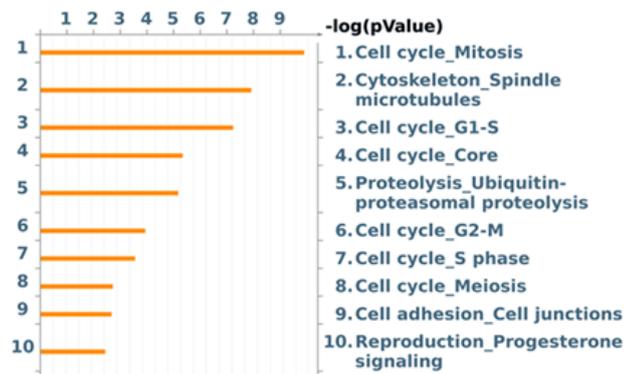
Upregulated Networks



Downregulated Networks



Upregulated pathways



Downregulated pathways

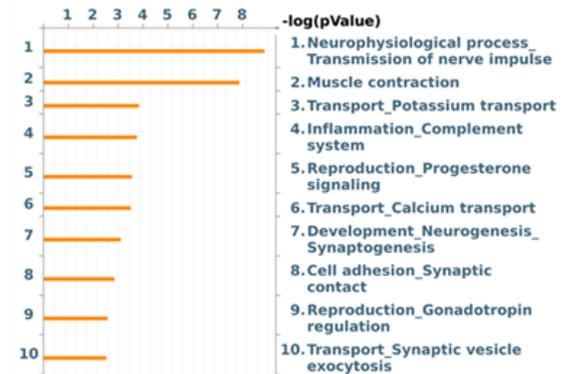


Figure 4.3 – Network and Pathway analysis of *Pb-Cre+ve Pten^{fl/fl} Arid1a^{+/+}* vs *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* prostate tumours

Significantly altered genes between *Pb-Cre+ve Pten^{fl/fl} Arid1a^{+/+}* and *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* tumours were uploaded to Metacore and pathway and network analysis was performed. Genes were filtered by threshold of 1.5 and P-value <0.05.

To further investigate the alterations in cell cycle, gene set enrichment analysis (GSEA) was performed (Figure 4.4). These genesets confirm that G2/M checkpoint is upregulated, particularly genes involved in mitotic spindle formation indicating both pro-growth signals and cell cycle machinery is upregulated. Furthermore, key regulators of cell cycle such as E2F and MYC have their targets upregulated in *Pb-Cre +ve Pten^{fl/fl} Arid1a^{fl/fl}* mice compared to *Pb-Cre +ve Pten^{fl/fl} Arid1a^{+/+}*. This may indicate then that Arid1a represses target

genes of each of these transcription factors and following its loss these genes are upregulated.

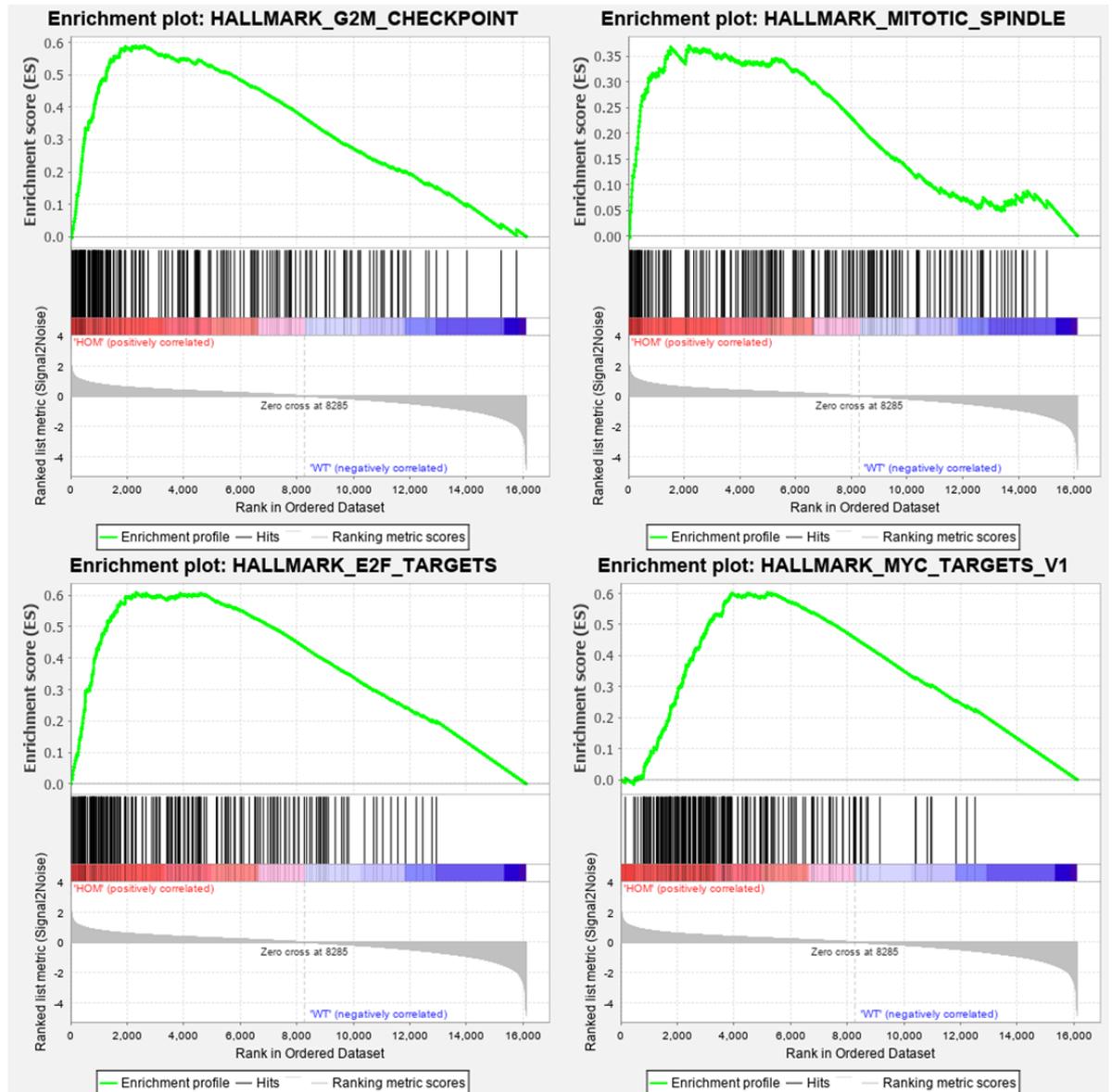


Figure 4.4 – Loss of *Arid1a* and *Pten* upregulates pro-growth signalling

Significantly altered genes between *Pb-Cre+ve Pten^{fl/fl} Arid1a^{+/+}* and *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* tumours were investigated to show enriched hallmark datasets with geneset enrichment analysis (GSEA). Hallmark_G2M NES: 1.58, P=0.000, Mitotic_Spindle NES:1.35, P=0.208, E2F_Targets NES: 1.44, P=0.097, MYC_Targets NES: 1.41 P=0.084.

4.3.4 ChIP-Seq performing and optimisation

4.3.4.1 Optimisation of chromatin digestion

To understand how loss of *Arid1a* causes tumorigenesis in the *Pb-Cre+ve Pten^{fl/fl}Arid1a^{fl/fl}* mice we needed to identify which genes are normally bound by ARID1A in a *Pb-cre+ve Pten^{fl/fl} Arid1a^{+/+}* mouse. Some of these genes will have been dysregulated following loss of *Arid1a*.

ChIP-Seq was performed as indicated in section 2.3.2, including the optimisation of chromatin digestion by micrococcal nuclease. This step is important to produce a range of DNA fragment size without over digestion which can impact many of the later steps including the DNA-pulldown, QRT-PCR, and sequencing, as small fragments are preferentially sequenced but not pulled down. The ideal range is to have fragment sizes between 150 - 900bp. Stock micrococcal nuclease was diluted 1:10 and incubated with sample for 20 minutes. Samples were then diluted and ran on a Bioanalyser to visualise DNA fragments (Figure 4.5).

This DNA gel shows that as micrococcal nuclease quantity increases, there is an increase in the 150bp band. Other bands also appear around 400-600bp. As one nucleosome is approximately 150bp, these represent a range of digestion leaving between 1 - 4 nucleosomes in length. As we want a range in fragments sizes, 1µl of the nuclease would not be preferred as it over digests the sample giving 150bp as the strongest band. 0.5µl however gives a good range in all other band sizes without under digestion of the sample. For this reason, all other experimental samples were incubated with 0.5µl of micrococcal nuclease.

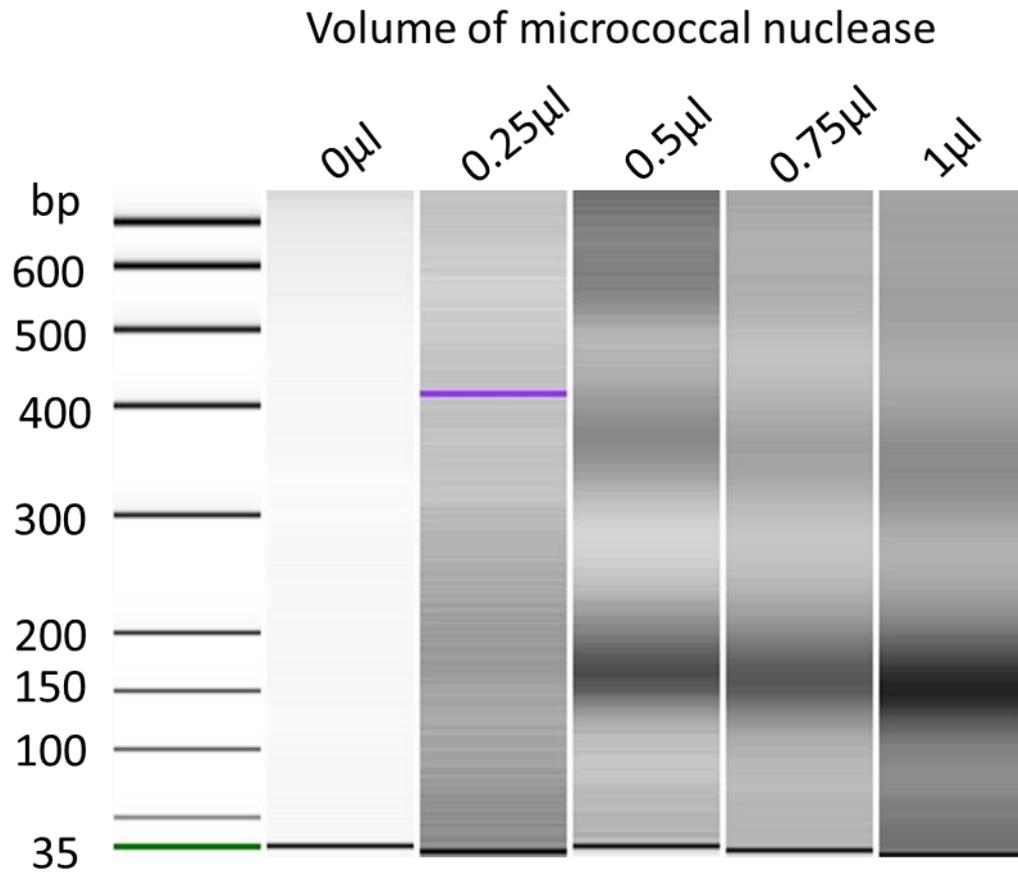


Figure 4.5 – Mnase digest optimisation for ChIP

Tumour cells derived from three *Pb-Cre+ve Pten^{fl/fl} Arid1a^{+/+}* prostate tumours were incubated with various volumes of micrococcal nuclease to optimise chromatin digestion. DNA ladder in first lane with band sizes labelled in basepairs (bp). 0µl micrococcal nuclease show no degradation with more degradation as nuclease volume increases.

4.3.5 Quantifying ChIP DNA pulldown by QRT-PCR

To ensure the chromatin pulldown has been successful, qRT-PCR was performed to quantify the amount of DNA in each sample. Data was represented relative to histone H4 pulldown which is very efficient and often recovers the most specific protein-bound DNA (Bao, Vinciotti *et al.* 2013)(Figure 4.6). This showed that the ARID1A pulldown recovered ~10x less DNA than histone H4 but still recovered ~10x more than the negative control. This gave final concentrations for Histone H4 samples around 5 - 10 ng/µl and ARID1A samples around 0.35 - 0.45ng/µl. This gave enough DNA to begin library preparation and sequencing. Prepared

libraries were run on Bioanalyser to ensure clean preparation and removal of all adapters (Figure 4.7). This showed a single peak in each sample indicating a clean library had been produced. Samples were then sent to Glasgow Polyomics for sequencing.

DNA pull-down relative to Histone H4

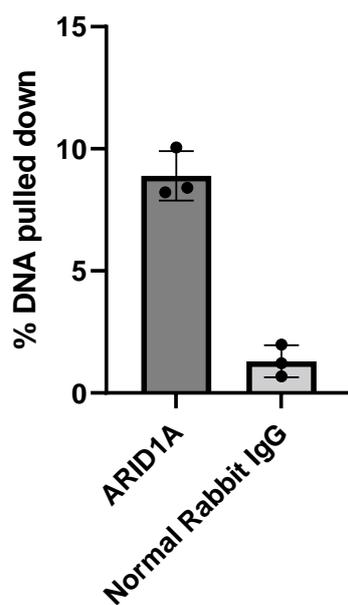


Figure 4.6 - QRT-PCR for ChIP

Quantitative real time – PCR (QRT-PCR) quantified amount of DNA from Histone H4 as positive control, ARID1A, and normal rabbit IgG as negative control. DNA pulled down represented as amount of DNA relative to positive control. Each point represents one independent experimental replicate each made up of three technical replicates.

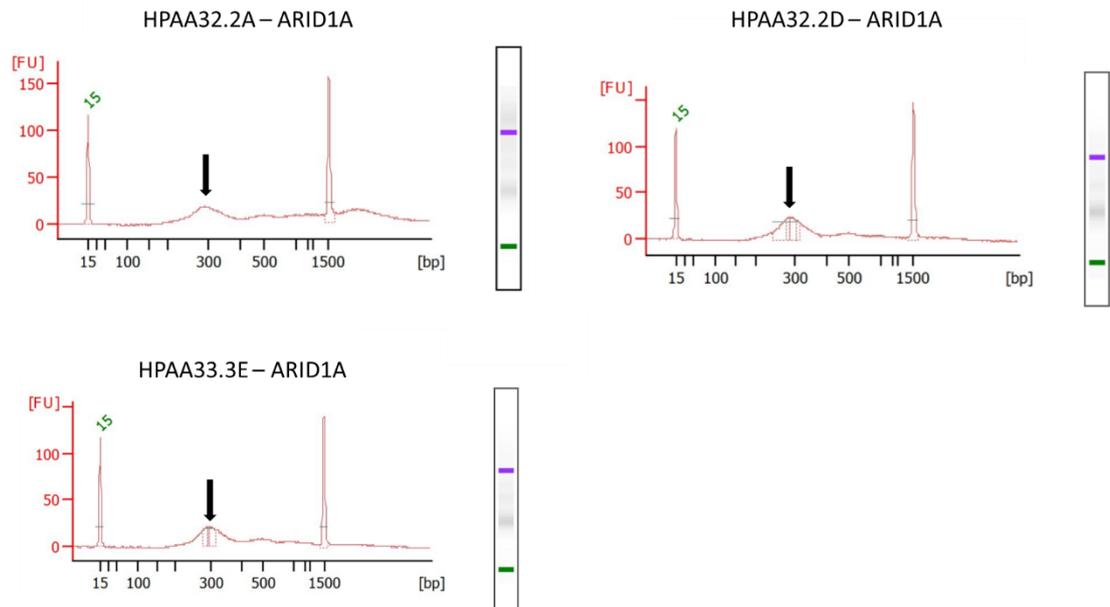


Figure 4.7 - Library preparation for ChIP

Libraries were prepared for each of the ARID1A ChIP pulldowns. To evaluate library preparation, libraries were ran on high sensitivity DNA chip Bioanalyser. Gels shows DNA quantity by arbitrary fluorescence unit (FU) and size of sample by basepairs (bp).

4.3.6 ChIP-Seq results

After sequencing, reads were trimmed, mapped, and aligned with informatics support from Robin Shaw, Beatson Institute. From here, between 18 - 24M reads per biological replicate of the ARID1A ChIP pulldown was received. Using deepTools, we were able to visualise where ARID1A bound around the genes TSS and transcription end site (TES) compared to Histone H4 as a positive control (Figure 4.8). This shows that ARID1A typically bound genes most frequently around -500kb from TSS, likely in a promoter region. This peak gradually reduces towards the end of the gene suggesting ARID1A is also involved in remodelling chromatin during transcription.

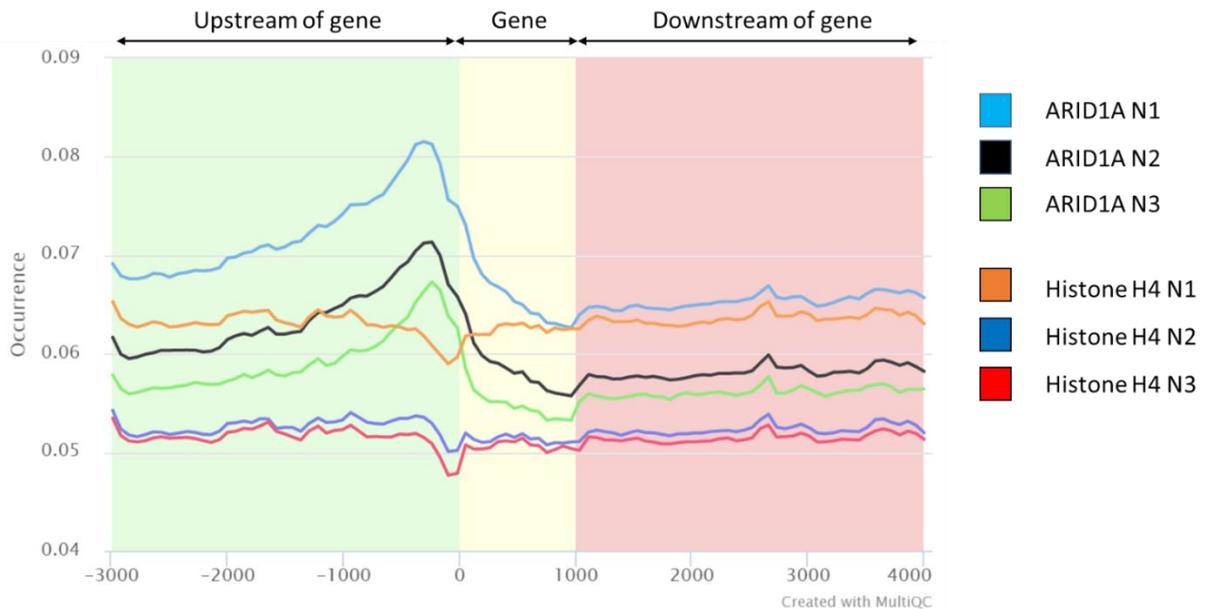


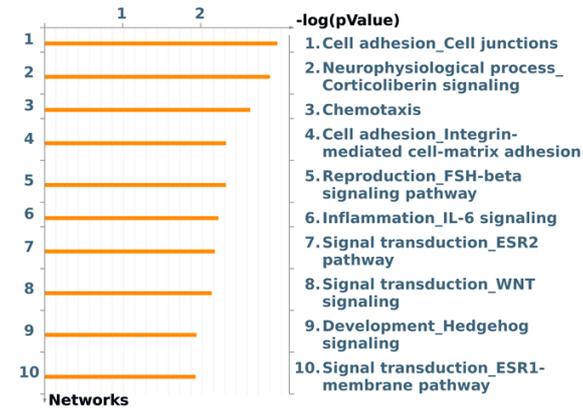
Figure 4.8 - Proximity of ARID1A binding to genes

Relative occurrence of ARID1A or Histone H4 binding compared to gene sequence. Binding to DNA is determined by ChIP-Seq of *Pb-Cre+ve Pten^{fl/fl} prostate* tumours. ARID1A N1-3 indicated by light-blue, black, and green line respectively. Histone H4 N1-3 indicated by orange, dark-blue, and red lines. Graph made using MultiQC.

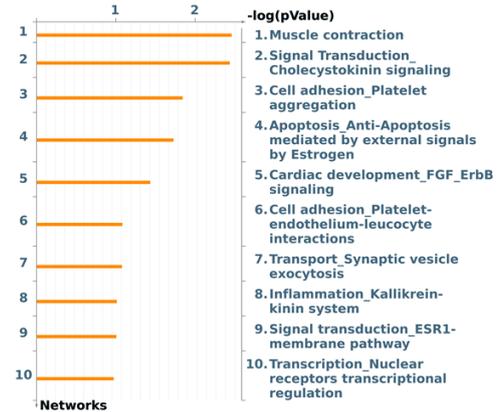
Following this, it was important to understand which of the genes that are bound by ARID1A are dysregulated following its loss. This would allow the understanding of how the genes that *Arid1a* directly controls contribute to the phenotype. To do this, the ChIP-Seq datasets of ARID1A were overlaid with significantly altered in the *Pb-Cre+ve Pten^{fl/fl} Arid1a^{+/+} vs Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* experiment from the RNA-Seq. Out of the 1590 genes that are dysregulated following *Arid1a* loss, 229 of them are directly bound by *Arid1a*. 113 of these genes were downregulated following *Arid1a* loss, and 116 were upregulated following *Arid1a* loss. This distribution is not seen in the overall RNA-Seq data which showed a bias towards downregulating genes after *Arid1a* loss.

Performing network analysis on these 229 genes shows that the upregulated genes are associated with migration and chemotaxis with key regulator of invasion and metastasis such as Plasminogen activator urokinase receptor (PLAUR), Laminin subunit gamma 2 (LAMC2), Claudin 4 (CLDN4), Partitioning defective 6 homology alpha (PARD6), Chemokine ligand 2 (CXCL2), and Chemokine ligand 3 (CXCL3) all being bound by *Arid1a* and upregulated following *Arid1a* loss (Figure 4.9 and Figure 4.10). This indicates that *Arid1a* suppresses pro-invasion genes indicating why its loss drives locally invasive dorsolateral tumours. Downregulated networks however did not show any strong consensus, similar to the RNA-Seq data, as this identified networks such as muscle contraction and cholecystokinin signalling.

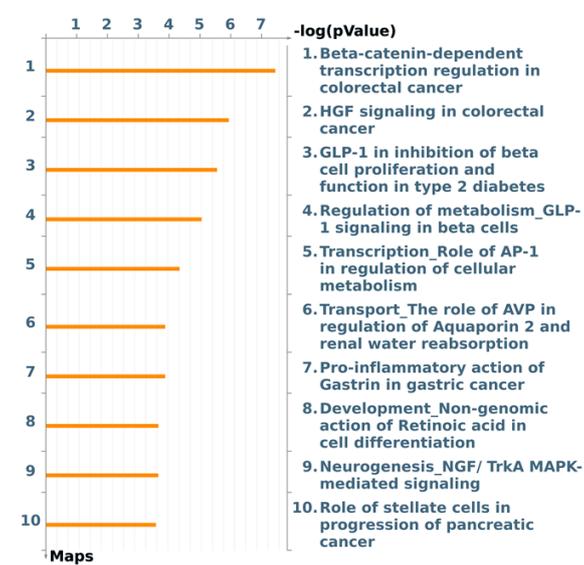
Upregulated Networks



Downregulated Networks



Upregulated Pathways



Downregulated Pathways

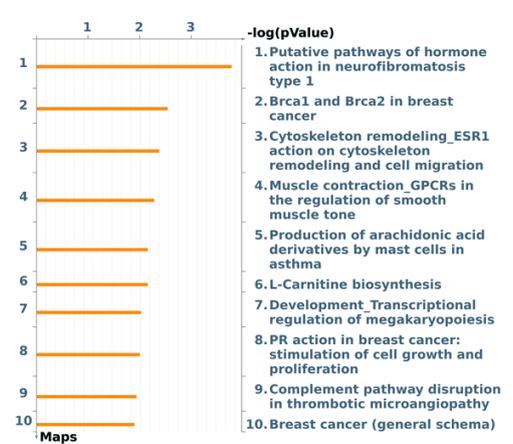


Figure 4.9 - Networks and pathway analysis of ARID1A-bound genes

Genes which were bound by ARID1A in clinical endpoint of three *Pb-Cre+ve Pten^{fl/fl}* mice were identified by ChIP-Seq. Of these genes, those that were significantly altered following ARID1A loss in clinical endpoint *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* were analysed by Metacore. Pathway and network analysis was performed, with genes were filtered by threshold of 1.5 and P-value <0.05.

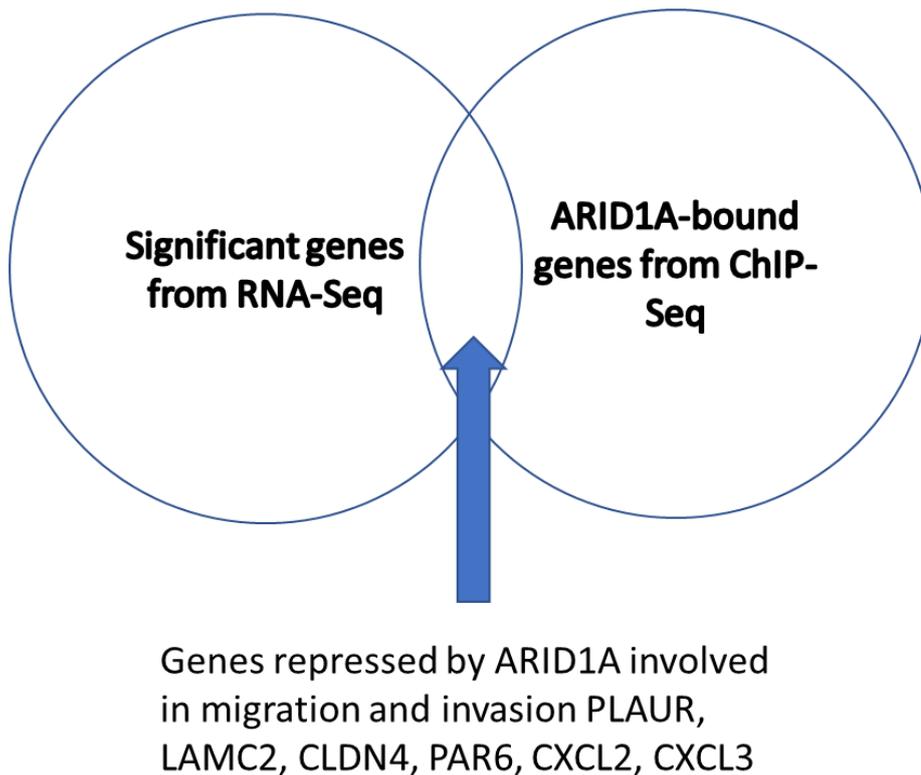
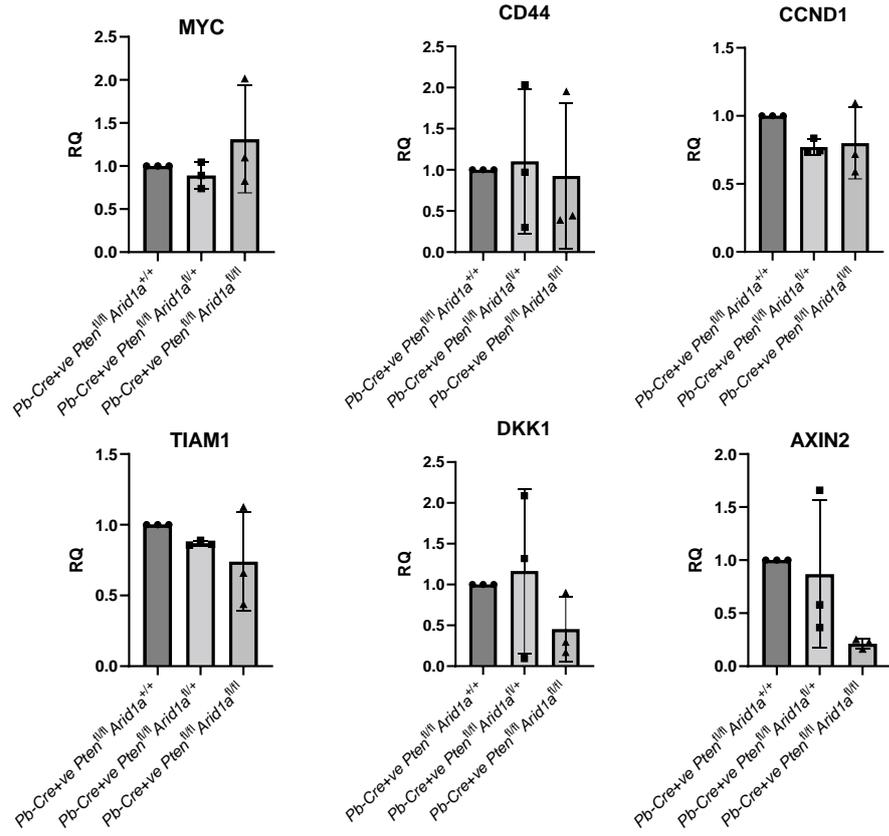


Figure 4.10 - ARID1A binds and represses drivers of migration and invasion

Overlaying significantly altered genes between *Pb-Cre+ve Pten^{fl/fl} Arid1a^{+/+}* vs *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* and ARID1A ChIP-Seq of *Pb-Cre+ve Pten^{fl/fl} Arid1a^{+/+}* prostate tumours. Overlap of these groups shows ARID1A regulated genes.

Analysis of upregulated pathways suggested that *Plaur* and *Lamc2* were both positively regulated by β -catenin, which was also bound by *Arid1a* and upregulated following *Arid1a* loss. This suggested that loss of *Arid1a* upregulated β -catenin signalling leading to elevated migration signalling. This seemed unusual as *Arid1a* has been demonstrated to be a positive regulator of β -catenin, with its loss abrogating β -catenin dependent pathways (Hiramatsu, Fukuda et al. 2019). Similarly, β -catenin signalling was shown to be downregulated in *Pb-Cre+ve Pten^{fl/fl} Arid1a^{+/+}* vs *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* pathway analysis from RNA-Seq. To validate this, QRT-PCR was performed on each of the *Pten* null mouse tumours and investigated the expression of Wnt target genes, *Myc*, *Cd44*,

CyclinD1 (*Ccnd1*), T-cell lymphoma invasion and metastasis-inducing protein 1 (*Tiam1*), Dickkopf WNT Signalling Pathway Inhibitor 1 (*Dkk1*), and axis inhibition protein 2 (*Axin2*). This showed that only *Myc* was upregulated in *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* tumours although this was not significant. All other genes remained comparable to *Pb-Cre+ve Pten^{fl/fl} Arid1a^{+/+}* tumours or reduced (Figure 4.11). This indicates that β -catenin signalling is not elevated in these tumours, and the *Plaur* and *Lamc2* are upregulated in another way.



β-Catenin Target Genes

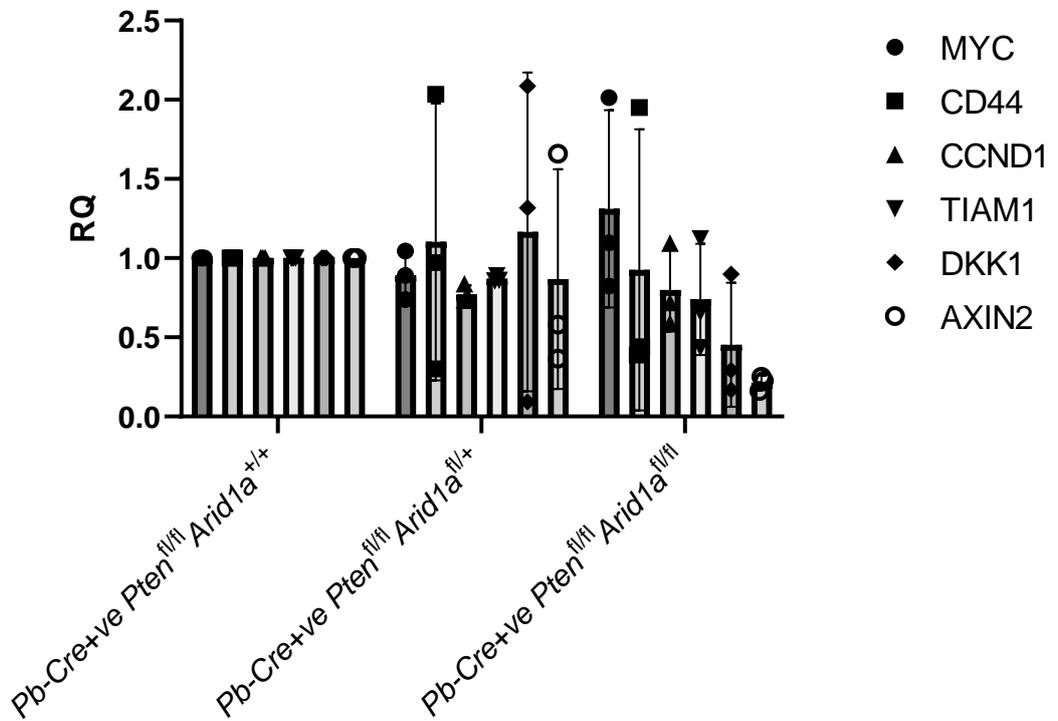


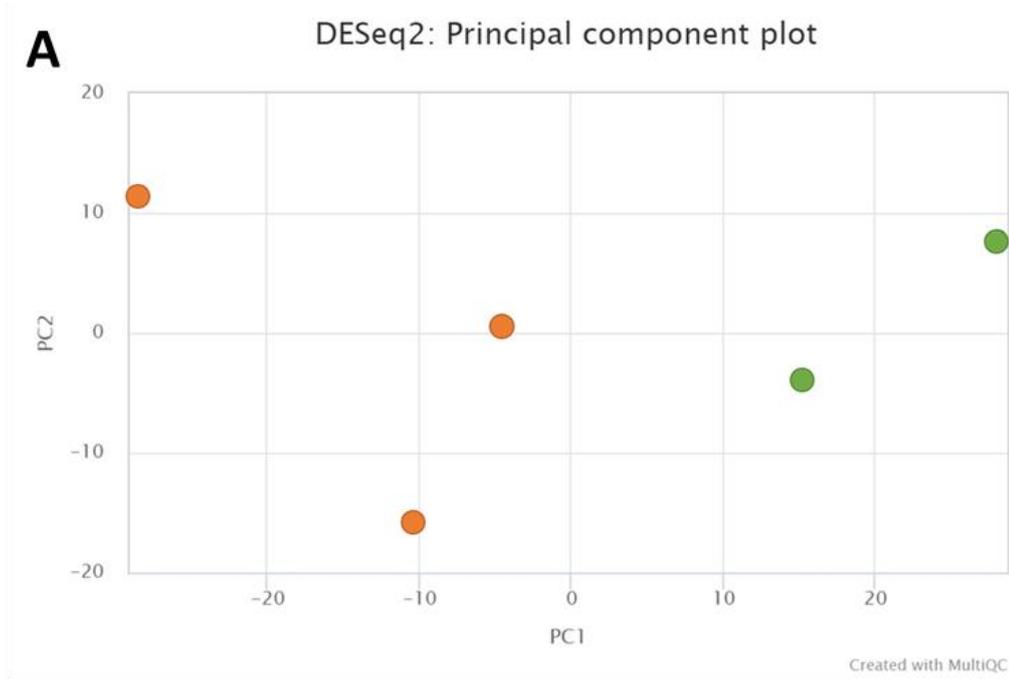
Figure 4.11 - β-catenin signalling is downregulated following loss of Arid1a

Relative quantification (RQ) of Wnt target genes compared between *Pb-Cre+ve Pten^{fl/fl} Arid1a^{+/+}*, *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/+}* and *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* mouse prostate tumours. Each point represents one independent experimental replicates each made up of three technical replicates. Error bars show standard error of the mean.

4.3.7 Using ATAC-Sequencing to investigate changes in chromatin accessibility after *Arid1a* loss

To investigate how the loss of *Arid1a* altered chromatin organisation, ATAC-Seq was performed by Syed Umbreen from the University of Belfast to compare *Pb-Cre+ve Pten^{fl/fl} Arid1a^{+/+}* and *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* mouse prostate tumours. Following sample processing in Belfast, samples were returned to Glasgow where they were then sequenced by Glasgow Polyomics. Analysis was then performed by Mr Robin Shaw (Informatician, Beatson Institute), using an analysis pipeline created by Holly Hall, Beatson Institute.

Following sequencing, one of the *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* samples could not be used as very few reads were sequenced in that sample. The variance between each sample of the remaining samples were then assessed by PCA (Figure 4.12A) and by DESeq2 (Figure 4.12B). The PCA shows that *Pb-Cre+ve Pten^{fl/fl} Arid1a^{+/+}* cluster separately to *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* samples, indicating a difference between each of these samples. DESeq2 clusters the samples respectively by genotype showing similarities within each genotype and differences between the two genotypes.



● *Pb-Cre+ve Pten^{fl/fl} Arid1a^{+/+}*

● *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}*

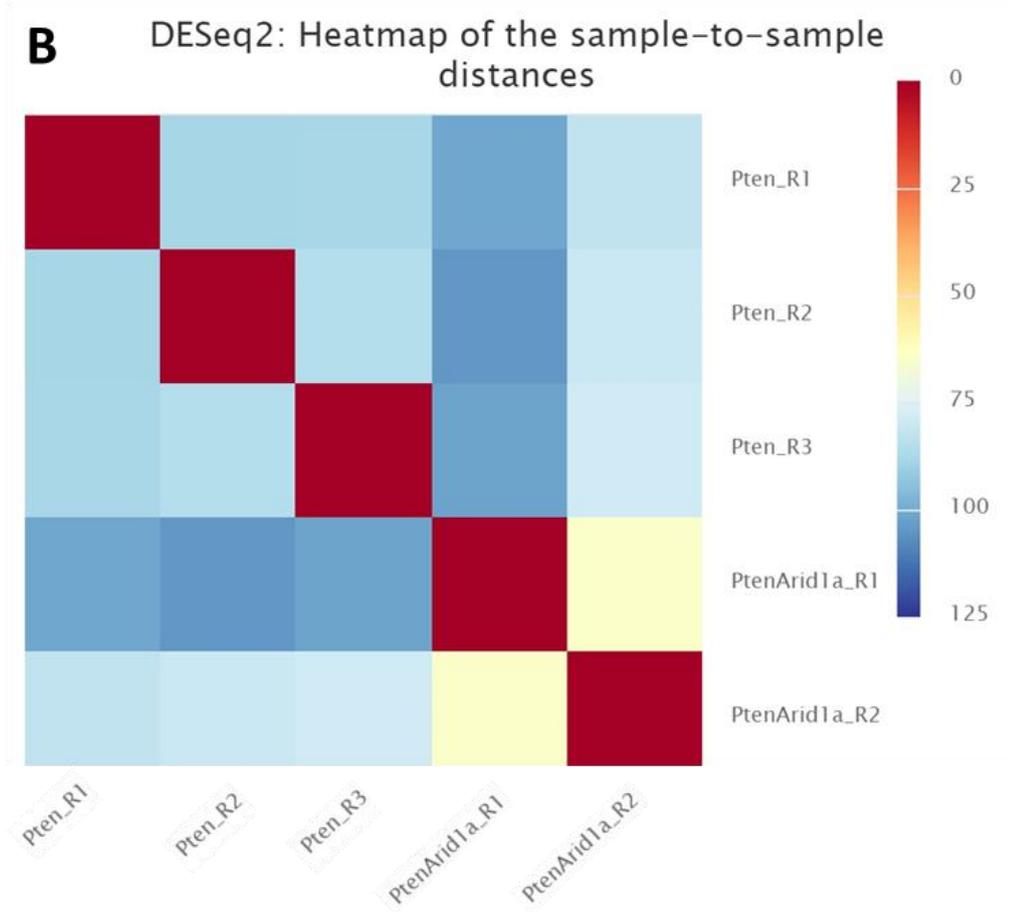


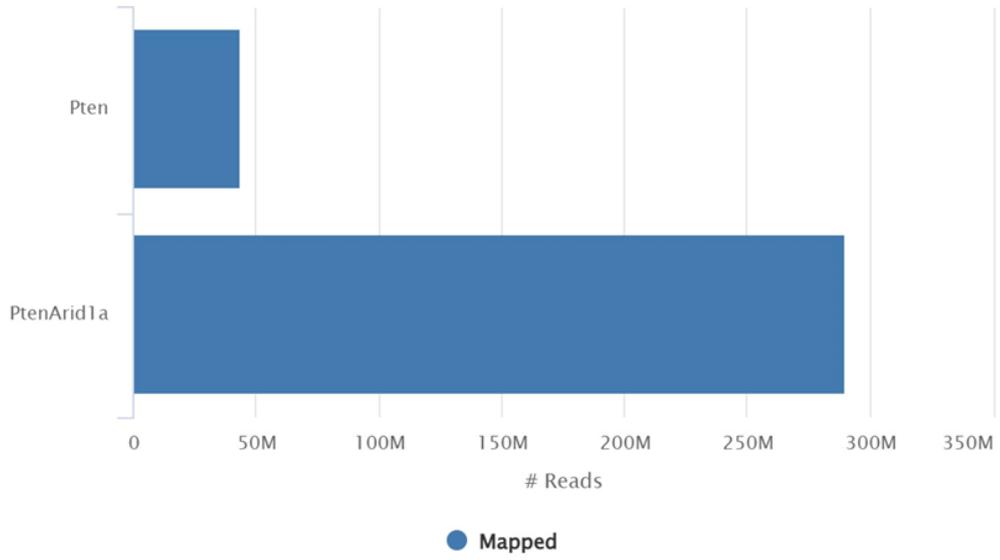
Figure 4.12 - ATAC-Seq Sample Variance

A. Principle component analysis (PCA) plot measuring the variance between three *Pb-Cre+ve Pten^{fl/fl} Arid1a^{+/+}* samples (orange) and two *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* mouse prostate tumours (green) which were processed for ATAC-Seq. **B.** DESeq2 analysis to measure variance between *Pb-Cre+ve Pten^{fl/fl} Arid1a^{+/+}* and *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* samples and to cluster based on this variance. Colour and number denote Euclidean distance between samples.

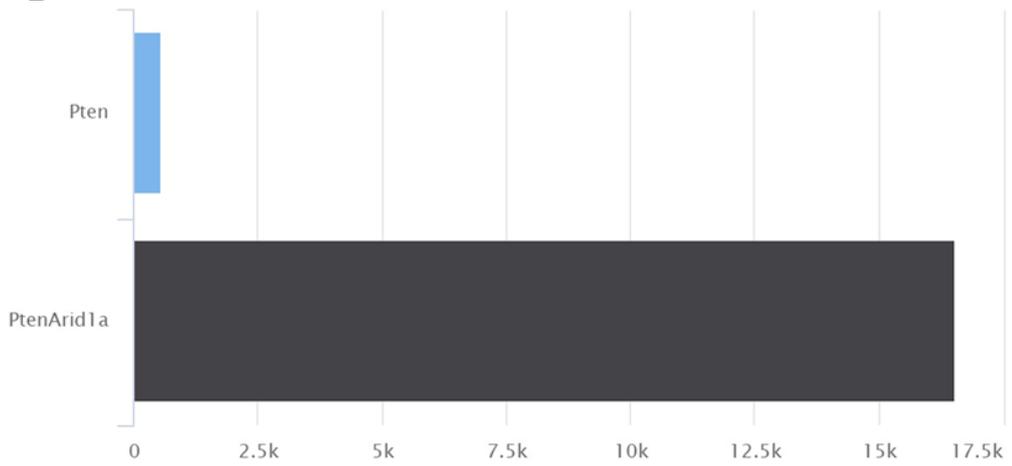
Samples were then filtered and had sequences mapped and aligned by the pipeline to understand insertion sites and of the ATAC-Seq transposon which indicates open chromatin (Figure 4.13). However, this data indicated that the ATAC-Seq may have been unsuccessful. After merging repeats into each genotype, reads were mapped to the murine genome. This showed that *Pb-Cre+ve Pten^{fl/fl} Arid1a^{+/+}* only had 48M reads compared to 280M reads in *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* cohort (Figure 4.13A). This indicated that many of the reads could not be mapped to the mouse genome or many of the reads were duplicated to the same areas. The mapped reads were then processed by MACS2 to call peaks and assign the location of the reads (Figure 4.13B). This showed only 547 reads in the *Pb-Cre+ve Pten^{fl/fl} Arid1a^{+/+}* cohort could be called to a peak compared to >16,000 for *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* cohort. Typically, the minimum number of peaks needed for further ATAC-Seq analysis would require >10,000 peaks. Of the peaks that were assigned, this showed that reads mapped mostly to intergenic regions of *Pb-Cre+ve Pten^{fl/fl} Arid1a^{+/+}* samples with a large enrichment of promoter-TSS insertions in *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* (Figure 4.13C). This may have indicated more active global transcription following *Arid1a* loss, however this cannot be validated as the *Pb-Cre+ve Pten^{fl/fl} Arid1a^{+/+}* cohort was unsuccessful.

As the *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* samples had no control with which to compare, the ATAC-Seq dataset could not be used for further analysis.

A Samtools stats: Alignment Scores



B Total peak count



C Peak to feature proportion

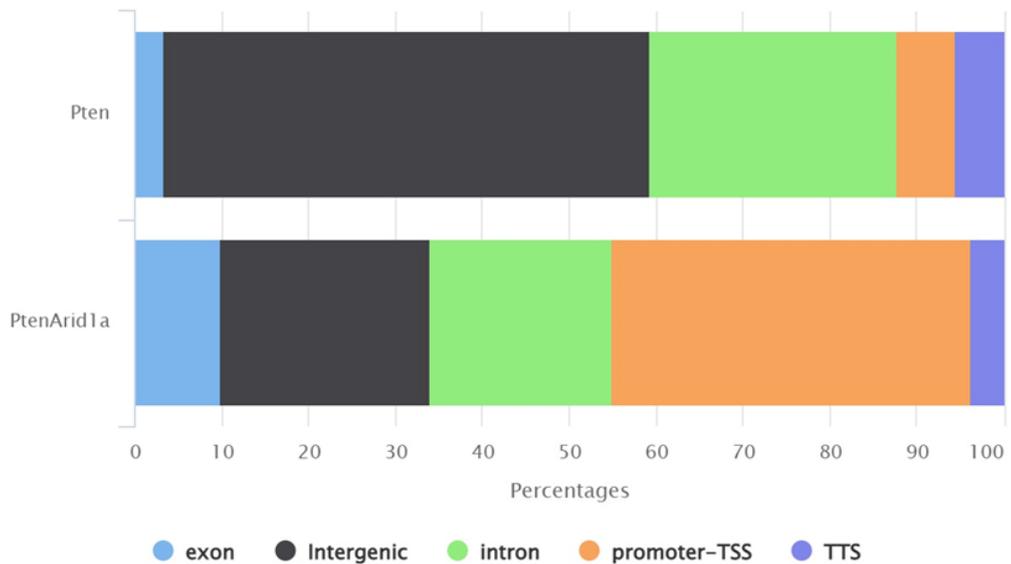


Figure 4.13 - ATAC-Seq fragment alignment and peak-calling

A. Total number of reads for *Pb-Cre+ve Pten^{fl/fl} Arid1a^{+/+}* and *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* samples aligned and mapped to mouse genome by SAMTOOLS. **B.** Aligned reads from SAMTOOLS had peaks called by MACS2. **C.** Peaks from B. called by MACS2 were ordered by feature type depending on ATAC transposon insertion site and represented as percentage of features relative to total peak count.

4.3.8 *Arid1a* loss may bypass P53-dependent senescence in *Pten* null mice

Pb-Cre+ve Pten^{fl/fl} Arid1a^{+/+} mice have been shown to be limited in growth by P53-dependent senescence signalling (Chen, Trotman et al. 2005). Loss of *Arid1a* in this model causes rapid tumour development and elevated cell cycle signalling which may suggest this senescence signalling has been bypassed. To investigate this, *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* tumours were stained for Ki67, a marker of proliferation, and P21 (CIP1/WAF1) and H2A histone family member X (γ -H2AX) which are markers of senescence. P21 is a target gene of P53 and controls G1-S phase transition. Higher levels of P21 indicate G1 arrest (Niculescu, Chen et al. 1998).

Staining for Ki67 confirms that *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* tumours have a large increase in proliferation evident by a more rapid tumour development as well as elevated cell cycle signalling (Figure 4.14). This suggests that the growth suppressive effects of P53 have been bypassed.

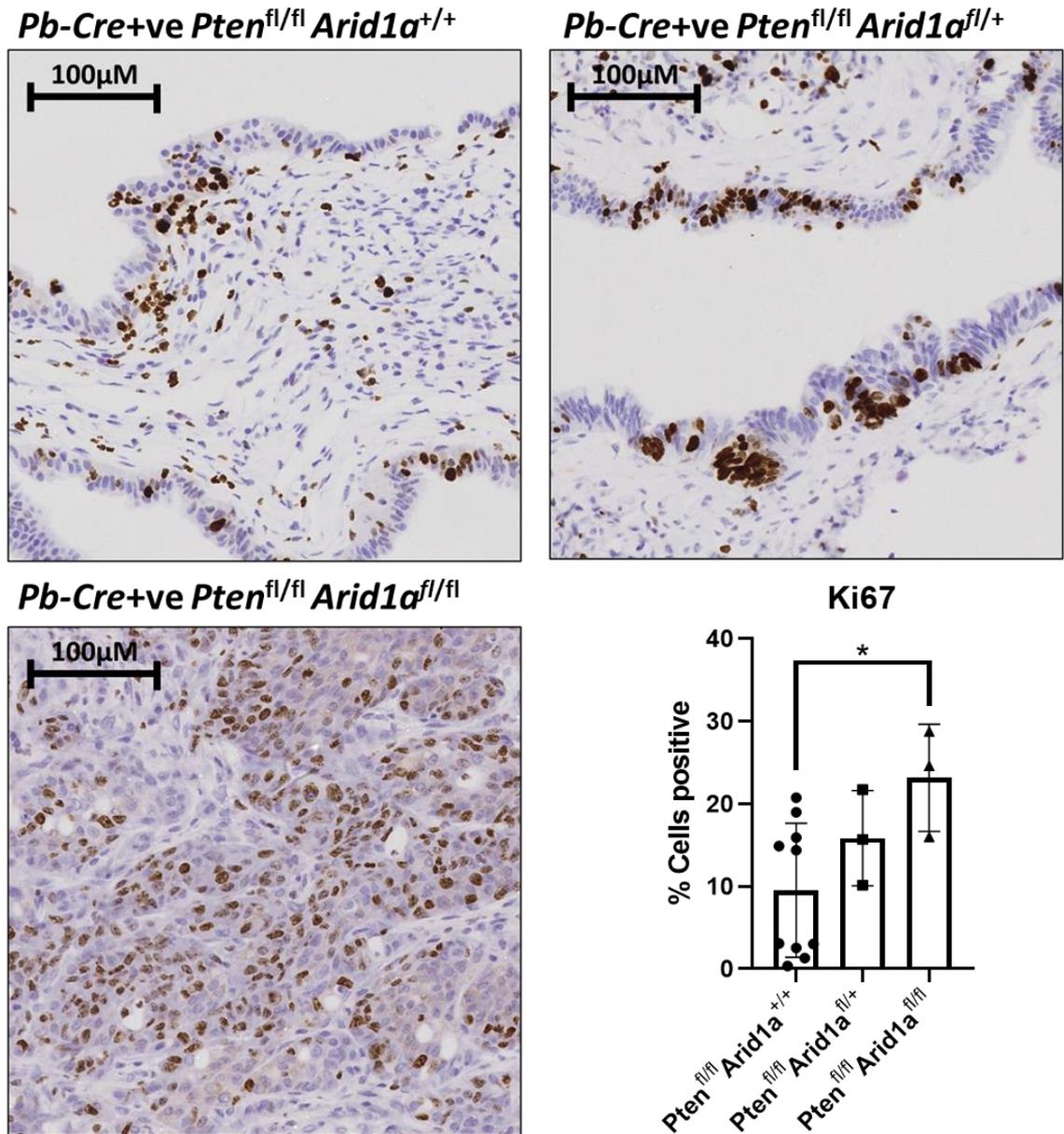


Figure 4.14 - Loss of *Arid1a* increases cell proliferation in *Pten* null prostate tumours

Prostate tumours of indicated genotype were stained for Ki67, images show represented tumours of each genotype (shown above the slide). Stain intensity was quantified for Ki67 based on a positive or negative signal using Halo Image analysis software. Each point represents one mouse. P-value < 0.05 (*) determined by one-way ANOVA.

Staining for P21 showed that *Pb-Cre+ve Pten*^{fl/fl} *Arid1a*^{+/+} have similar protein levels to *Pb-Cre+ve Pten*^{fl/fl} *Arid1a*^{fl/+} (Figure 4.15). Surprisingly, *Pb-Cre+ve*

Pten^{fl/fl} Arid1a^{fl/fl} mice had the highest protein level with strong nuclear and cytoplasmic levels, around 3-fold higher than *Pb-Cre+ve Pten^{fl/fl} Arid1a^{+/+}*. These high levels would therefore indicate senescence and a G1/S phase arrest.

However, the rapid tumour growth and high Ki67 (Figure 4.14) strongly suggest that these cells are highly proliferative. This may therefore indicate that the high P21 is unable to arrest the cell. As the RNA-Seq data shows that *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* tumours alter cell cycle signalling particularly around G2/M transition high P21 may still arrest cell cycle around G1/S phase.

P21 has also been shown to function as an oncogene by inhibiting apoptosis (Ohkoshi, Yano *et al.* 2015). This tumour suppressor or oncogenic role of P21 however is related to its localisation, with nuclear P21 inhibiting proliferation and cytoplasmic P21 inhibiting apoptosis. In this staining however cytoplasmic levels of P21 remain similar between all tumours while only nuclear P21 levels increased in *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* mice.

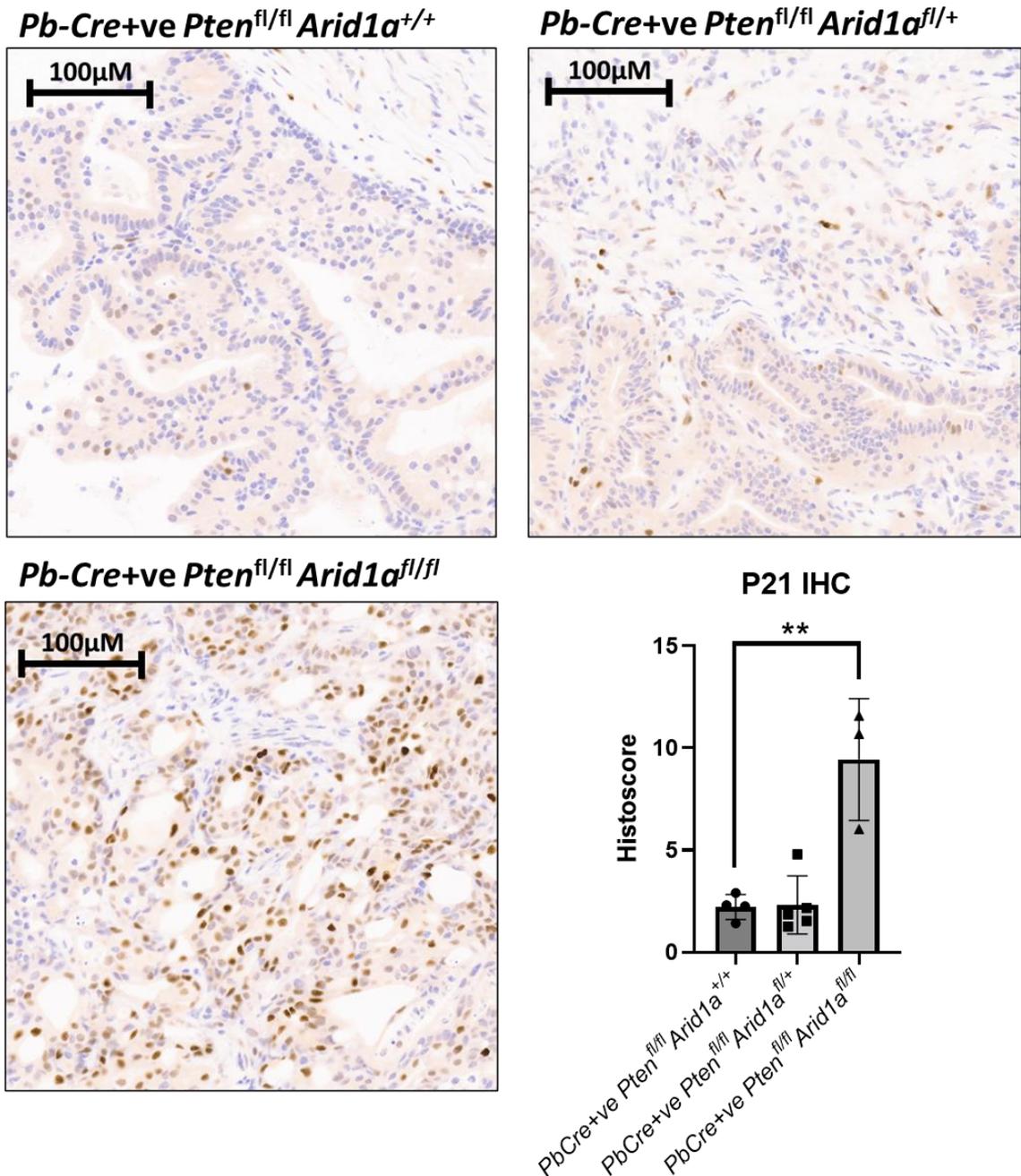


Figure 4.15 - Loss of *Arid1a* upregulates P21 in *Pten* null mice

Prostate tumours of indicated genotype were stained for P21, images show represented slides of each genotype (given above each slide). Stain intensity was quantified for P21 into a histoscore using Halo Image analysis software. Each datapoint represents a single mouse. P-value <0.005(**) determined by one-way ANOVA. Histoscore determined by formula: (% cells low intensity)+2(%cells medium intensity)+3(%cells high intensity)=Histoscore.

To characterise other senescence markers, tissues were also stained for γ -H2AX, a marker of DNA-damage which can be upregulated following replication stress,

one of the drivers of *Pten* loss induced senescence. This staining shows that *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/+}* mice had the highest levels of γ -H2AX staining, whereas *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* had the lowest compared to *Pb-Cre+ve Pten^{fl/fl} Arid1a^{+/+}* mice (Figure 4.16). This would suggest that while loss of one copy of *Arid1a* from a *Pb-Cre+ve Pten^{fl/fl}* did not strongly impact tumour growth or survival it may have introduced replication stress or DNA-damage. This could explain why *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/+}* had higher variability in their survival. This also ties in with ARID1A roles in resolving replication stress and mediating the DNA-damage response (Shen, Peng *et al.* 2015, Tsai, Fournier *et al.* 2021). However, this is not explained by *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* having the lowest levels of γ -H2AX suggesting they have the lowest levels of DNA-damage. Loss of ARID1A has also been shown to reduce γ -H2AX deposition following DNA-damage (Shen, Peng *et al.* 2015). This may indicate that *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* will have more replication stress and unresolved DNA-damage, but this is not confirmed.

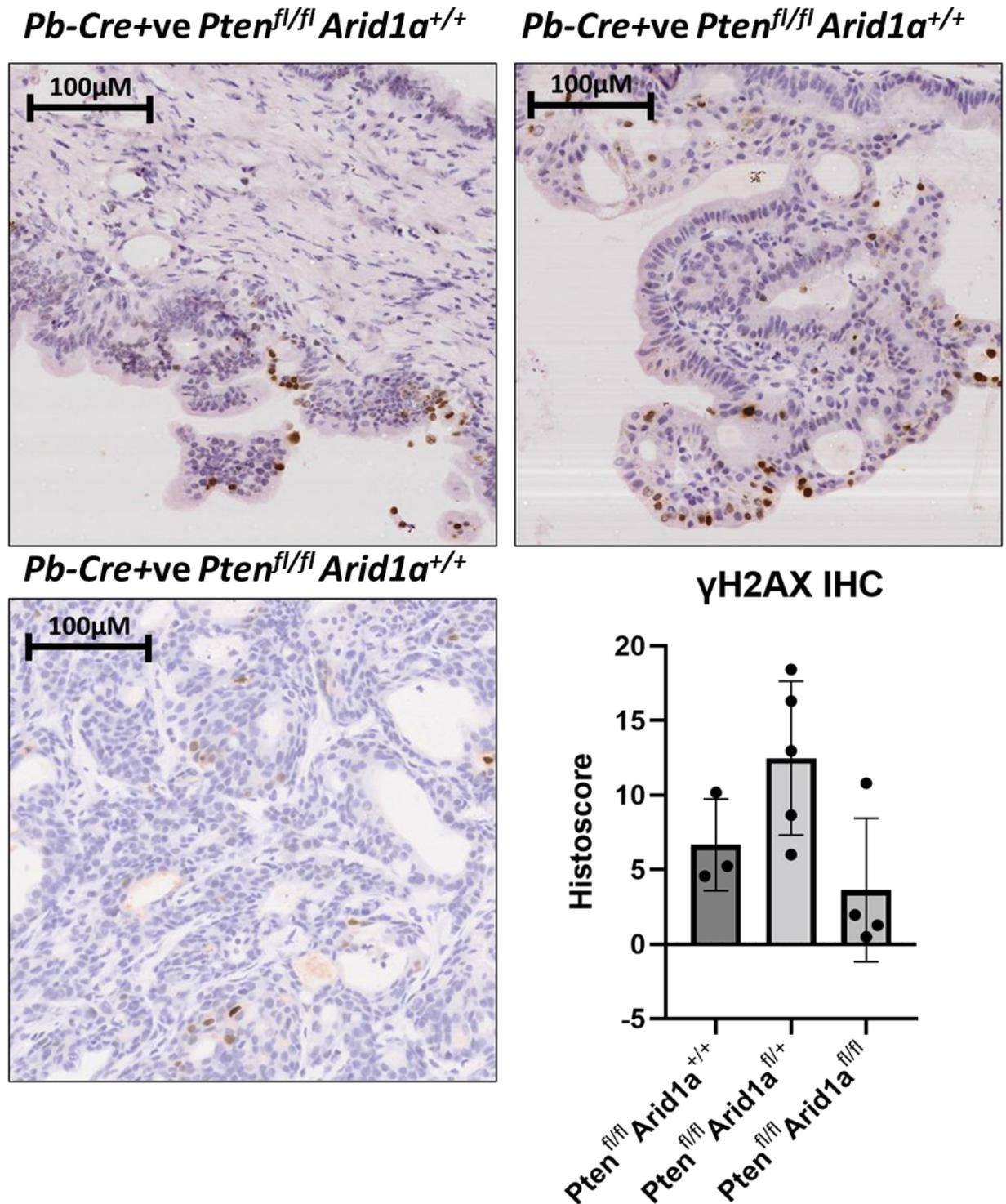


Figure 4.16 – Loss of *Arid1a* does not alter gamma-H2AX in *Pten* null mice

Prostate tumours of indicated genotype were stained for gamma-H2AX, images show representative slides of each genotype (given above each picture). Stain intensity was quantified for gamma-H2AX into a histoscore using Halo Image analysis software. Each datapoint represents a single mouse. Histoscore determined by formula: (% cells low intensity)+2(%cells medium intensity)+3(%cells high intensity)=Histoscore.

High DNA-damage and high P21 may also suggest high levels of P53 activity despite its inability to limit growth. To investigate this, the tumours were stained for P53 (Figure 4.17). This staining showed that P53 levels correlated strongly with P21 levels with *Pb-Cre+ve Pten^{fl/fl} Arid1a^{+/+}* having similar protein levels to *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/+}*, and *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* having 3-fold higher P53 levels. This indicates that either in response to DNA-damage or oncogenic signalling, P53 levels are upregulated to attempt to inhibit proliferation. However, as the tumours still proliferate rapidly, this suggests P53 is unable to arrest their growth. Supporting this idea, we performed gene set enrichment analysis (GSEA) with the RNA-Seq datasets which shows that DNA repair and P53 pathways gene sets are upregulated in *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* tumours compared to *Pb-Cre+ve Pten^{fl/fl} Arid1a^{+/+}* tumours (Figure 4-18). This may suggest that DNA repair pathways are being upregulated in response to replication stress which in turn activates P53 activity. Altogether, these findings may then indicate that loss of *Arid1a* in the *Pten* null mouse is able to bypass or overcome this senescence induced growth arrest as cells continue to proliferate despite clear upregulations in P53 activity leading to elevation in P21.

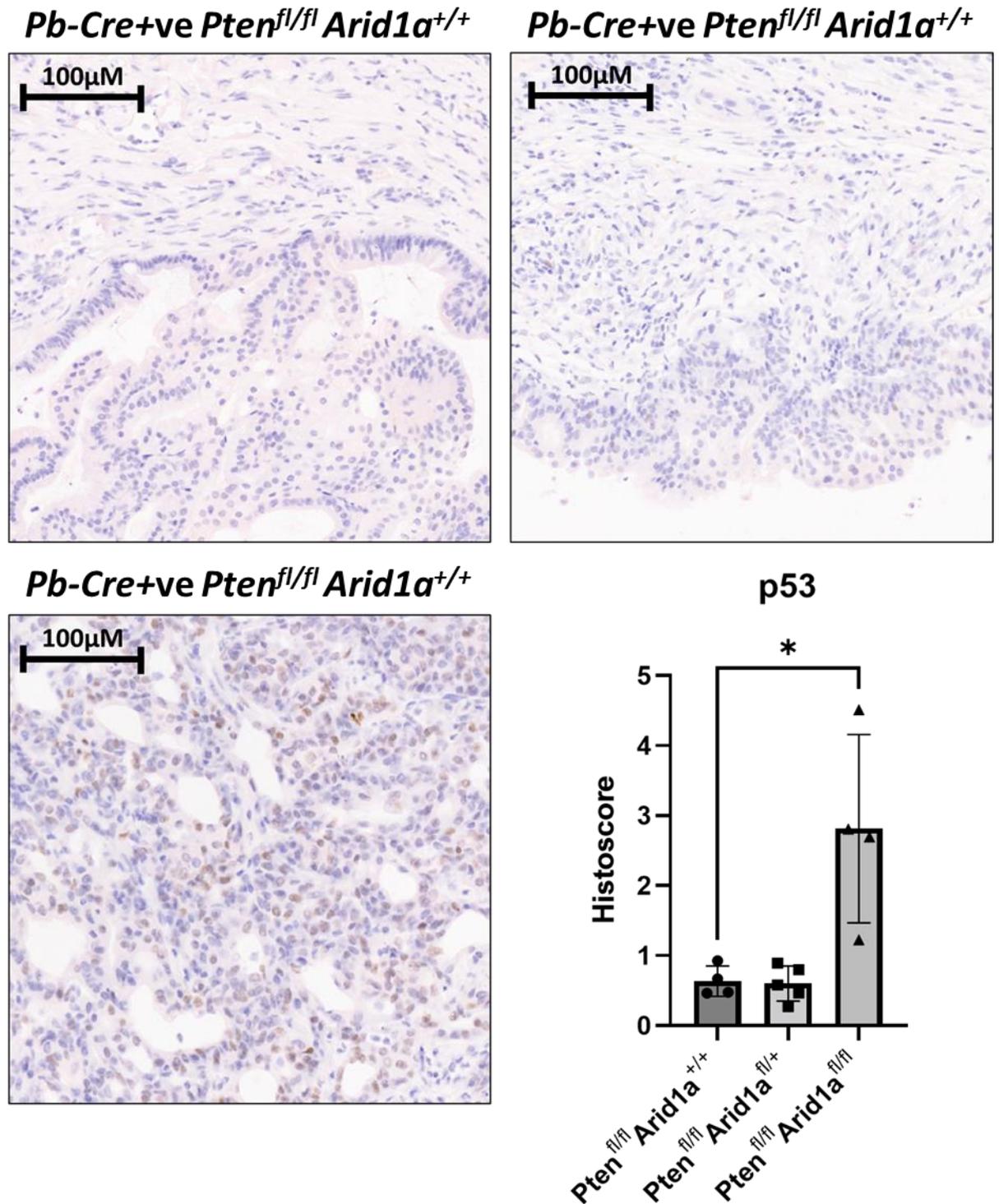


Figure 4.17 – Loss of *Arid1a* upregulates P53

Tissue of indicated genotype were stained for P53, images show represented images of each genotype. Stain intensity was quantified for p53 into a histoscore using Halo Image analysis software. Each datapoint represents a single mouse. P-value < 0.05 (*) determined by one-way ANOVA. Histoscore determined by formula: (% cells low intensity)+2(%cells medium intensity)+3(%cells high intensity)=Histoscore.

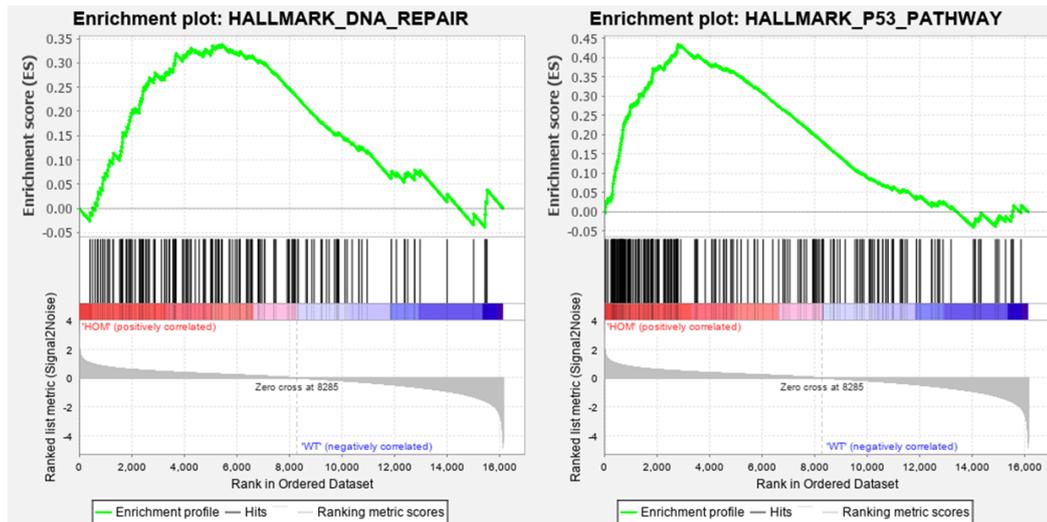


Figure 4.18 - GSEA DNA repair and P53

Significantly altered genes in *Pb-Cre+ve Pten^{fl/fl} Arid1a^{+/+}* and *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* tumours were investigated to show enriched hallmark datasets with geneset enrichment analysis (GSEA). Hallmark_DNA_Repair had a NES 1.22 (P=0.276) and Hallmark_P53_Pathway had a normalised enrichment score of 1.26 (0.181).

4.4 Discussion

4.4.1 ARID1A as a regulator of cell cycle in cancer

ARID1A's most obvious tumour suppressor roles are observed in its role in cell cycle. In murine embryonic stem cells, ARID1A protein levels are highest at G1/G0 and lower, but not depleted, in all other phases (Flores-Alcantar, Gonzalez-Sandoval *et al.* 2011).

ARID1A interacts with many other tumour suppressors to induce G1 cell cycle arrest. By interacting with HIC1, a well characterized tumour suppressor, ARID1A helps to repress E2F1 expression (Van Rechem, Boulay *et al.* 2009). ARID1A also interacts with P53 to upregulated P21 expression (Guan, Wang *et al.* 2011). This functional interaction with bona fide tumour suppressor genes is strongly

indicative of ARID1A's potential role in inhibiting tumorigenesis by mediating cell cycle arrest.

Depletion of ARID1A in an osteoblast precursor cell line (MC3T3-E1) resulted in downregulation of P21, retained expression of cyclin A, B2, and C, and continued DNA synthesis (Nagl, Patsialou *et al.* 2005). However, a later study in the same cell lines showed that ARID1A deficient cell lines in fact did eventually arrest at G1/S although the response was delayed (Nagl, Wang *et al.* 2007). These cells also reinitiated cell cycle faster than parental cells. This indicates that while ARID1A is clearly important in cell cycle arrest, it is not required. Furthermore, cells without ARID1A are more readily proliferative given faster re-entry to cell cycle. Targeting ARID1A deficient ovarian cancer cell lines with dasatinib (a drug primarily used in treating CML/ALL patients with Ph⁺ disease), has been shown to be effective. The suggested mechanism being due to upregulation of P21 triggering G1/S phase arrest (Miller, Brough *et al.* 2016). This is an interesting finding if ARID1A deficiency can be exploited by restoring normal G1 arrest functions. However, the mechanism of synthetic lethality of dasatinib is not completely clear and was also suggested to target YES1 kinase. The relevance of YES1 in ARID1A deficient cancers has not been explored and may imply an effect of dasatinib independent of ARID1A's function on cell cycle.

These studies demonstrate ARID1A as a positive regulator of P21, however in our *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* tumours we see a large upregulation of P21. This is coupled with an elevation in P53 activity. In these other settings, loss of ARID1A impaired P53 regulation of P21, though this is clearly not this case in *Arid1a* null prostate tumours (Guan, Wang *et al.* 2011).

Deficiency in ARID1A has also been shown to impair G2/M arrest in response to DNA damage (Shen, Peng *et al.* 2015). However, as ARID1A is important in mediating an efficient DDR, particularly by interacting with ATR, we cannot be sure if the inability to arrest at G2/M is due to defective DNA damage checkpoint signalling, or an ARID1A-specific role in G2/M arrest. However, a later study showed that ARID1A does have a specific role in G2/M arrest (Wu, Lyu *et al.* 2018). ARID1A has been shown to suppress AURKA, with loss of ARID1A causing upregulation of AURKA. Higher levels of AURKA leads to persistent CDC25C activation driving G2/M transition. Subsequent inhibition of AURKA triggers G2/M arrest and apoptosis, indicating another viable synthetic lethality instance in ARID1A null cancers.

In contrast, ARID1B has been shown to have pro-proliferative functions. Loss of ARID1B results in depleted levels of C-Myc and delayed entry to cell cycle (Nagl, Wang *et al.* 2007). Restoring C-Myc expression restored normal cell cycle functions. This paper also showed that ARID1B primarily interacts with E2F1 which is considered the pro-proliferative E2F transcription factor. It is interesting that ARID1A-HIC1 repress E2F1 expression, perhaps to limit C-Myc expression via ARID1B. ARID1A has also shown to be present at c-Myc promoter, perhaps also actively repressing c-Myc expression alongside repressing E2F1 (Nagl, Zweitzig *et al.* 2006).

ARID1A deficient cancer could progress through cell cycle more readily due to this upregulation of c-MYC, E2F1 in G1/S phase, and AURKA at G2/M. In this model, we see the *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* tumours upregulate the genesets of both CMYC and E2F1 as wells as AURKA itself. This phenotype is therefore consistent with what is known about ARID1A in the literature with it repressing

cell cycle signalling. The role of ARID1A in cell cycle is summarised in Figure 4.19.

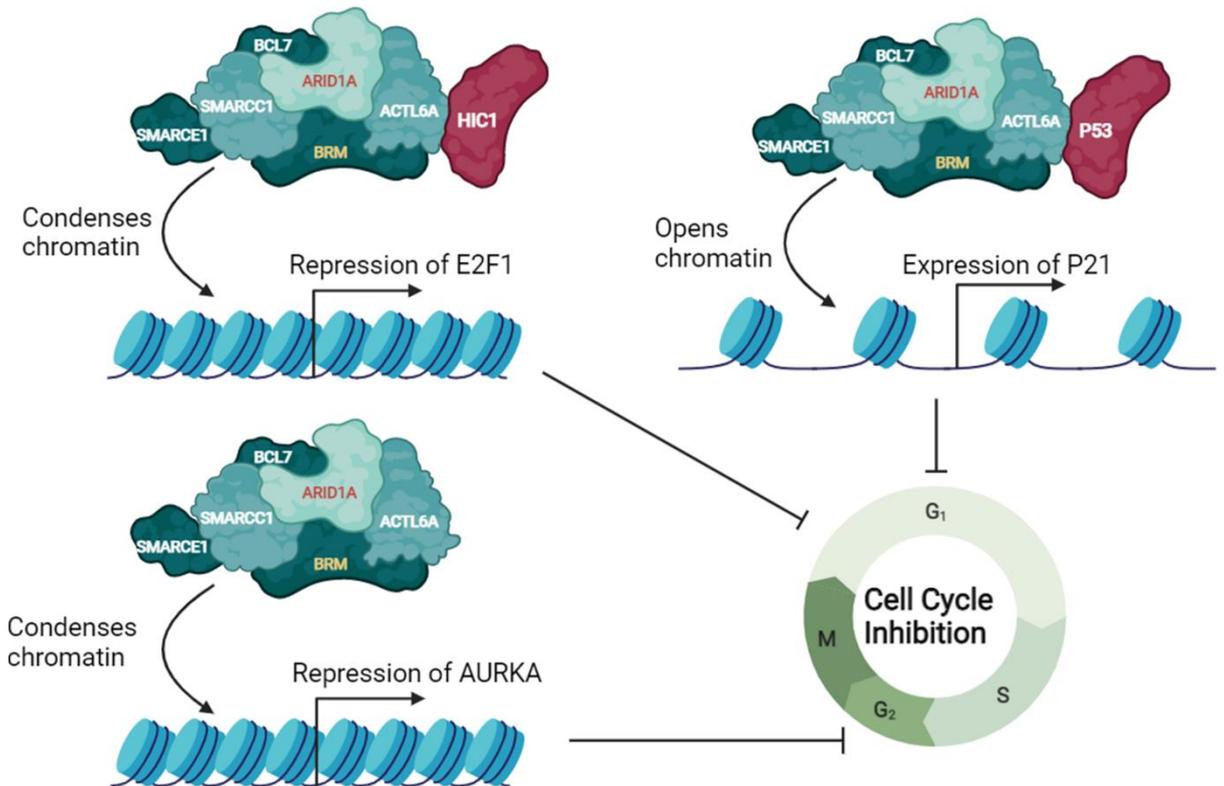


Figure 4.19 - ARID1A can control cell cycle by altering accessibility to cell cycle regulators

By interacting with HIC1, ARID1A has been shown to repress E2F1. By interacting with P53, ARID1A has been shown to upregulate P21 expression. ARID1A can directly bind and repress AURKA expression. Created with BioRender.

However, the ChIP-Seq data showed that ARID1A did not directly bind any major modulators of cell cycle signalling, and primarily bound genes involved in migration and invasion. Following loss of *Arid1a*, RNA-Seq identified 390 upregulated genes, which when analysed by Metacore strongly suggested cell cycle pathway involvement. Of those 390 genes though, 116 were directly bound by ARID1A and showed no enrichment in cell cycle signalling pathways. This indicates that the upregulation in cell cycle signalling is likely due to the

cascades in signalling after loss of ARID1A rather than the direct repression of the genes by ARID1A.

An explanation for this could be related to the BAF complex itself, with loss of ARID1A causing a retargeting to ARID1B favoured loci, which in this instance may indicate cell cycle related genes.

4.4.2 Neuroendocrine Prostate Cancer

Neuroendocrine PC (NEPC) is a relatively rare subtype of prostate cancer accounting for between 0.5 - 1% of PC cases which is characterised by neuroendocrine markers such as synaptophysin and chromogranin A (Kaarijärvi, Kaljunen *et al.* 2021). NEPC is also shown to be highly proliferative with commonly upregulated cell cycle markers including AURKA and PLK1. This subtype often appears following prolonged androgen deprivation and upregulates neural pathways, and although *de novo* NEPC has also been observed, it is phenotypically distinct from treatment-induced NEPC (Yamada and Beltran 2021). NEPC is highly heterogenous and is not well characterised among the PC subtypes.

Epigenetic remodelling has been shown to play an important role in the neuroendocrine phenotype. NEPC has been shown to express BAF complex subunits BAF53B and BAF45B which are neuronal SWI/SNF subunits which were previously thought to only be expressed in cells of a neuronal lineage (Cyrta, Augspach *et al.* 2020). While the exact function of these subunits in NEPC has not been characterised it is believed they control these neural pathways.

The *Pb-Cre +ve Pten^{fl/fl} Arid1a^{fl/fl}* genotype develops with intact androgen pathways and downregulates neural pathways, indicating that *Arid1a* may be important in mediating a neuroendocrine phenotype. Similarly, NEPC markers synaptophysin and chromogranin A are both 3-fold downregulated in the *Pb-Cre +ve Pten^{fl/fl} Arid1a^{fl/fl}* tumours.

However, ARID1A loss has been observed in some NEPC tumour organoids (Beshiri, Capaldo *et al.* 2022). These organoids had also lost ARID1B, indicating a BAF complex with neither ARID subunit. Loss of both ARID subunits was previously shown to be lethal to cells, however loss of both ARID subunits has also been observed and is capable of driving aggressive tumorigenesis (Helming, Wang *et al.* 2014, Wang, Chen *et al.* 2020). The ARID1A and ARID1B null organoids maintained a stem-like phenotype which gave rise to multiple progenitor cells with a full range of different lineage markers. These types of NEPC have also been observed to have lower levels of typical NEPC markers though are very stemlike and maintain expression of stemness markers such as SOX2 (Beshiri, Capaldo *et al.* 2022).

The *Pb-Cre +ve Pten^{fl/fl} Arid1a^{fl/fl}* tumours upregulate SOX2 2-fold and SOX2 target genesets are enriched in the RNA-Seq genes that were identified as bound by ARID1A from the ChIP-Seq. This potentially indicates that the *Pb-Cre +ve Pten^{fl/fl} Arid1a^{fl/fl}* tumours have adopted a stemness phenotype following loss of ARID1A similar to certain NEPC tumours.

4.4.3 Methodological changes and future directions

There are still many questions left to answer in understanding the role of *Arid1a* in these *Pten* null mouse tumours. While we understand which genes are directly

bound by ARID1A and are dysregulated following its loss, the cascades which altered the other pathways such as cell cycle signalling are not understood. Understanding how loss of ARID1A alters the localisation of the BAF complex may help answer this question. Performing CHIP-Seq on *Pb-Cre +ve Pten^{fl/fl} Arid1a^{+/+}* vs *Pb-Cre +ve Pten^{fl/fl} Arid1a^{fl/fl}* tumours to look for other BAF subunits ARID1B, SMARCA2 and SMARCA4 would answer this question on the re-targeting of the complex. Similarly, understanding the global changes in chromatin architecture with ATAC-Seq would allow us to understand changes in transcription factor activity following *Arid1a* loss by performing motif analysis. Many of these experiments have been limited by tissue availability due to the limited breeding of the *Pb-Cre +ve Pten^{fl/fl} Arid1a^{fl/fl}* genotype caused by their aggressive phenotype.

4.4.4 Conclusions

In this chapter, I have demonstrated that loss of *Arid1a* in combination with loss of *Pten* in prostate cancer upregulates proliferation and cell cycle signalling with an elevation in E2F1 and MYC genesets. *Arid1a* also directly binds and regulates genes involved in migration and invasion which are subsequently upregulated following *Arid1a* loss. Additional work will focus on understanding ARID1A in human prostate cancer cells by generating and investigating CRISPR-Cas9 ARID1A knockout cell lines.

5 Investigating ARID1A in Human Prostate Cancer Cell Lines

5.1 Introduction

5.1.1 ARID1A and Prostate Cancer

As previously stated, while the entirety BAF complex is highly dysregulated in PC, ARID1A is the most mutated with missense and truncation (Abida, Cyrta *et al.* 2019). Despite this, ARID1A has not been well studied in PC, and most of the current research either examines its role in NEPC or as a coactivator of AR (Sandoval, Pulice *et al.* 2018, Stelloo, Nevedomskaya *et al.* 2018, Xu, Chhangawala *et al.* 2020).

In a cell line VCaP, transmembrane protease serine 2 (TMPRSS2) and ETS-related gene (ERG) gene fusion, part of the erythroblast transformation specific (ETS) family of transcription factors, binds and targets the BAF complex to specific loci in order to remodel chromatin around ERG target genes (Sandoval, Pulice *et al.* 2018). This process and the interaction between ERG and the BAF complex were shown to be ARID1A-dependent. In this instance, it was reported that ARID1A can function as a mediator of an oncogenic transcription factor. Knockdown of ARID1A reduces the growth of VCaP due to loss of ERG driven growth signals. Interestingly, loss of ARID1B does not elicit the same response thereby highlighting ERG as one of the ARID1A specific transcription factors in this context.

Another study identified that mucin 1 cell surface associated (MUC1), a known driver of NEPC, upregulates ARID1A, BRG1, BAF60a, BAF155, and BAF170 in CRPC and NEPC cells (Hagiwara, Yasumizu *et al.* 2021). This combination of BAF

subunits makes up a BAF subtype called the embryonic stem cell BAF (esBAF) that has been suggested to be the primary type of BAF complex in stem cells which contributes to plasticity (Ho, Ronan *et al.* 2009). The esBAF has been shown to upregulate NOTCH1 and NANOG both of which contribute to cancer stemness. This lineage plasticity has been shown to be a mechanism of resistance in 10 - 20% of patients, with the esBAF being the primary driver (Cyrta, Augspach *et al.* 2020).

ARID1A has importantly been shown to be an important mediator of AR signalling. As previously shown in AR positive breast cancer cell lines, loss of ARID1A resulted in cells being unresponsive to androgen with a reduction in AR responsive genes following androgen stimulation (Xu, Chhangawala *et al.* 2020).

This has also been well demonstrated in PC. In LNCaP cells, AR has been shown to bind the BAF complex following stimulation by R1881, a synthetic androgen. AR interacted most strongly with ARID1A more than any other BAF subunit and showed no interactions with ARID1B. As LNCaP cells are AR dependent, knockdown of ARID1A reduced LNCaP viability as it disrupted AR signalling (Stelloo, Nevedomskaya *et al.* 2018).

In all, these studies of ARID1A demonstrate that it mediates key oncogenic drivers of PC such as ERG and AR. To date, no study in PC has demonstrated any tumour suppressor effects of ARID1A.

5.2 Hypothesis and Aims:

As shown in Chapters 3 and 4, loss of *Arid1a* cooperates with loss of *Pten* to drive tumorigenesis through an elevation in cell cycle signalling. Therefore, we can hypothesise that ARID1A loss in human cell lines which do not express PTEN may similarly cause an elevation in growth. If ARID1A can cooperate with other drivers besides PTEN, ARID1A loss may also increase growth of PTEN positive cell lines also. ARID1A has also be shown in PC to be able to facilitate some oncogenic signalling in PC such as ERG or AR, therefore loss of ARID1A may also impair growth in some settings.

The aims of this chapter are as follows:

- Investigate the levels of ARID1A in PC cell lines and characterise the effect of its knockdown on cell growth.
- Generate ARID1A knockout clones of ARID1A in PC cell lines and perform RNA-Seq to investigate signalling alterations following loss of ARID1A.
- Perform phenotypic analysis of ARID1A knockout clones to understand how loss of ARID1A effects tumorigenic properties such as migration, invasion, and survival.
- Compare and contrast RNA-Seq datasets between ARID1A knockout clones and GEM model to understand shared overlap in signalling following ARID1A loss.

5.3 Results

5.3.1 ARID1A mutations in human prostate cancer

Using CBioportal, patients can be visualised with mutations, copy number alterations, and mRNA dysregulation. From this the alteration rates of ARID1A among these different cohorts can be seen (Figure 5.1). This data shows that ARID1A is mutated between 1.5 - 3% of cases across each of the different cohorts with the highest alteration rates being found in metastatic PC cohorts indicating that loss of ARID1A is perhaps more important in advanced disease.

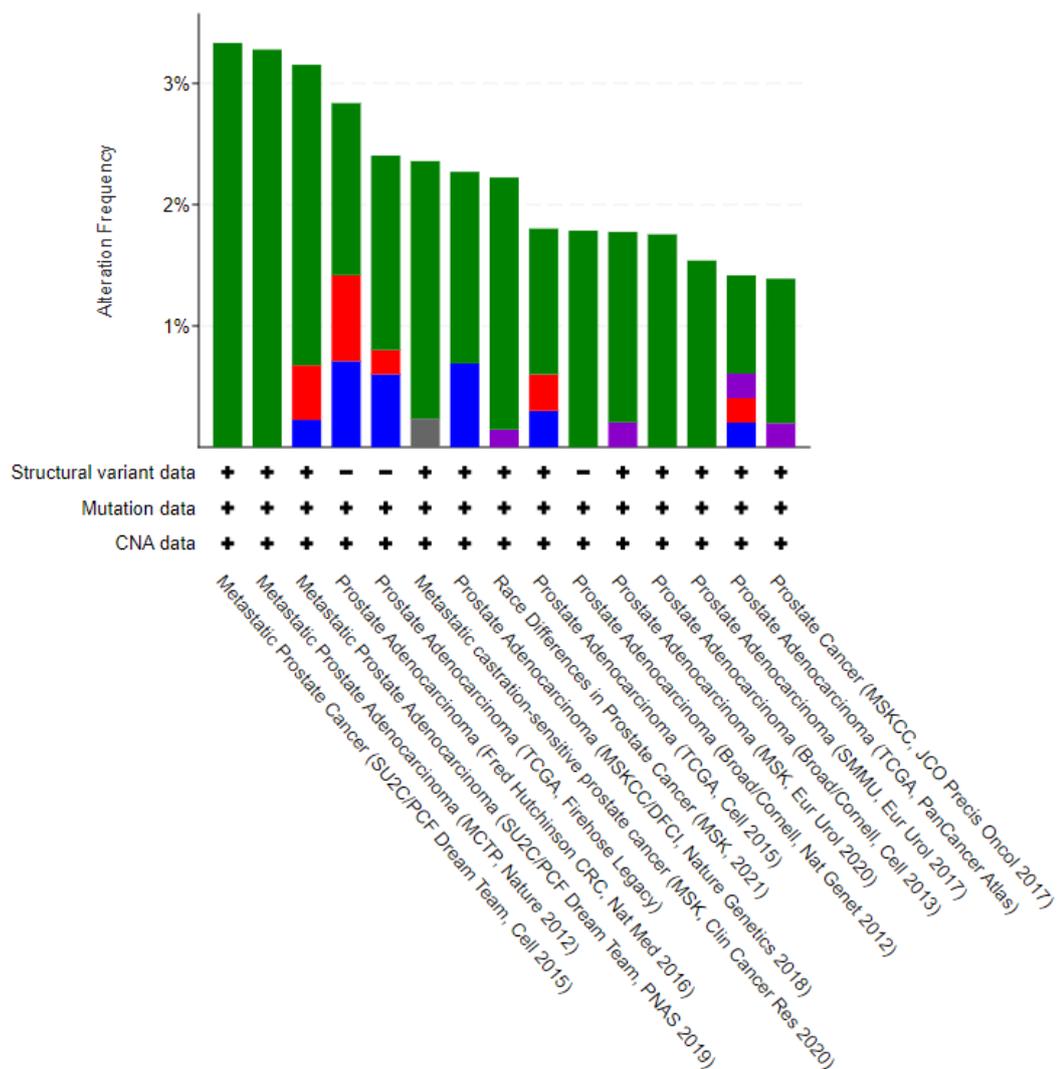


Figure 5.1 – ARID1A mutations in Prostate Cancer Cohorts

Rates of mutations shown for ARID1A gene visualised on CBioportal with indicated cohorts show with mutational characterising data indicated. Green indicates mutation by missense or truncation,

red indicates copy number amplification, blue indicates deep deletion of gene. Purple indicates gene fusion. Each study reference is given in the figure.

It is possible to map these mutations to the ARID1A gene to investigate if there are any hotspots for mutation (Figure 5.2). These mutations however are spread out throughout the ARID1A gene and most often result in truncation and missense mutation which will likely lead to degradation of the ARID1A transcript. As these mutations likely do not cause changes in the protein structure at a particular part of ARID1A it may indicate that the complete loss of ARID1A is required. This may be due to ARID1A's role as a binding hub with no single structural alteration completely affecting this role.

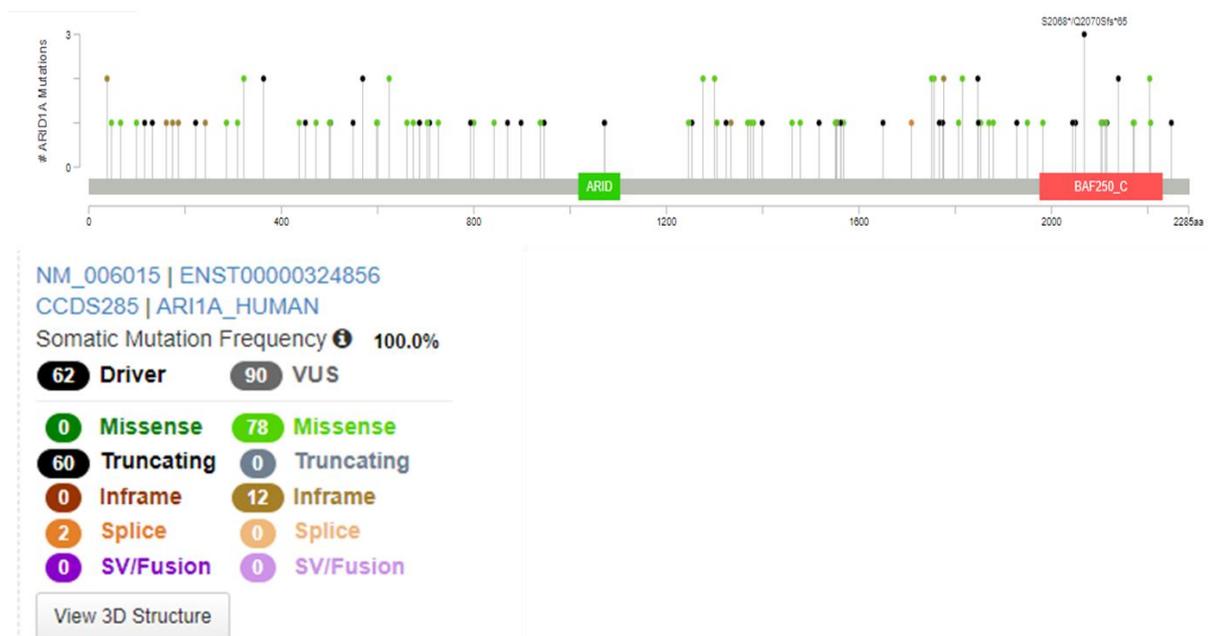


Figure 5.2 - Hotspot mutations in ARID1A gene

Mutational data from all prostate cancer cohorts on CBioPortal with mutations mapped to ARID1A gene. Visualised in CBioPortal.

The entirety of the BAF complex is highly mutated in metastatic PC, with 60-70% of cases having an alteration in at least one subunit of the complex. These

alterations likely alter the composition of the BAF complex to facilitate remodelling of chromatin around oncogenes. Of these subunits, ARID1A remains the most mutated either by missense or truncation mutations with the lowest instance of amplification (Figure 1.11).

To understand the molecular drivers that are enriched in ARID1A mutant PC, we can use cBioPortal to filter cases by presence of ARID1A mutation and look at the frequency of alterations in classical drivers of PC (Table 5.1). This data shows that in ARID1A mutant primary PC has an enrichment of PTEN deletions, TP53 mutations, and MYC amplifications. Meanwhile ARID1A mutant metastatic PC only has a MYC enrichment. Interestingly, only 20% of ARID1A mutant metastatic PC cases have PTEN deletion. As the murine model in chapter 3 indicated, *Pten* loss was a prerequisite to *Arid1a* loss being tumorigenic. However, this data may suggest that ARID1A alone is simply insufficient, and it needs to be combined with another driver, not PTEN specifically. Both primary and metastatic PC had lower instances of TMPRSS2-ERG fusions in ARID1A mutant PC. This is likely because ARID1A can function as a coactivator of ERG so the two mutations do not synergise (Sandoval, Pulice *et al.* 2018).

	Gene	Alteration rate in ARID1A mutant PC	Alteration rate in ARID1A WT PC	Co-occurrence or exclusivity with ARID1A
Primary PC	PTEN	15.2% Deletion	10.1% Deletion	N/A
	AR	0% Amplification	1.1% Amplification	N/A
	MYC	17.1% Amplification	4.4% Amplification	Co-occurrence (P<0.001)
	TP53	24.5% Mutation	15.1% Mutation	N/A
	TMPRSS-ERG	15.6% Fusion	30.9% Fusion	Mutual Exclusive (P=0.003)
Metastatic PC	PTEN	20% Deletion	22.5% Deletion	N/A
	AR	24% Amplification	30.9% Amplification	N/A
	MYC	16% Amplification	11.8% Amplification	N/A
	TP53	29.1% Mutation	35.8% Mutation	N/A
	TMPRSS-ERG	22.0% Fusion	26% Fusion	N/A

Table 5.1 - Alteration frequency of oncogenic drivers of prostate cancer with or without ARID1A mutation

Table derived from CBioPortal using all available prostate cancer datasets filtered by metastatic disease vs primary disease with and without ARID1A missense or truncation mutations.

5.3.2 Cell seeding and lysis buffer optimisation for western blotting

To characterise ARID1A levels in PC cell lines, western blotting for ARID1A protein was performed. However, there was inconsistency in detecting ARID1A protein between lysates of the same cell line. To investigate this, the lysis method was optimised in the PC3 cell line which appeared to have the most consistent ARID1A levels.

This optimisation test shows that radioimmunoprecipitation assay buffer (RIPA) buffer did not allow as much detection of the ARID1A protein compared to 1% SDS (Figure 5.3). However lower bands were equally detected in both lysates. This may indicate that 1% SDS liberated nuclear, and chromatin bound proteins better than RIPA buffer.

It was also shown that lysing pellets of trypsinised cells did not allow detection of ARID1A. With RIPA buffer, there was no detectable ARID1A protein, and with 1% SDS, we saw a greatly reduced level of ARID1A and an enrichment of degradation products. This indicates that in response to trypsinisation, ARID1A is likely being degraded as part of the stress response. Suspending cells by physically disrupting their attachment with PBS and lysing the resultant pellet still allows ARID1A detection by RIPA buffer. This indicates that the stress of suspension alone did not cause ARID1A degradation.

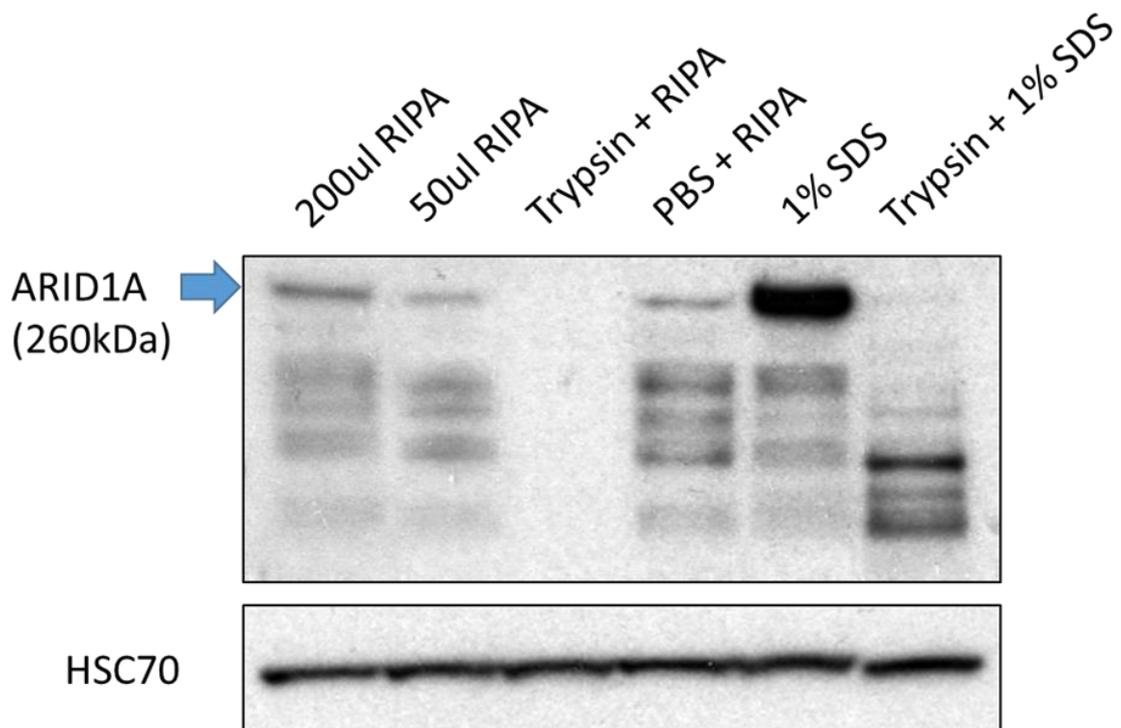


Figure 5.3 - Lysis buffer optimisation

Immunoblotting for ARID1A in PC3. 2×10^5 cells were seeded and lysed after 48h with the indicated lysis buffer. HSC70 used as a loading control. Representative blot based on three independent experimental replicates.

Optimisation of seeding densities for detecting ARID1A also showed changes in the band pattern of detecting ARID1A. To investigate this, the band pattern and

ARID1A levels were compared as confluence changed in PC3 (Figure 5.4A-C). Quantifying these bands shows that as confluency increased (Figure 5.4C), the 260kDa ARID1A band accumulates up to 72h and slightly decreased at 96h (Figure 5.4B). The degradation products likewise accumulated over time peaking in 96h. To validate if these lower bands were indeed degradation products, at 72h cells were incubated with MG132, a proteasome inhibitor, for 30 and 60 minutes (Figure 5.4D). Quantifying these bands showed a reduction in degradation products and an increase in ARID1A at 60 minutes (Figure 5.4E). This indicates that these lower bands were degradation products which accumulated at higher confluence.

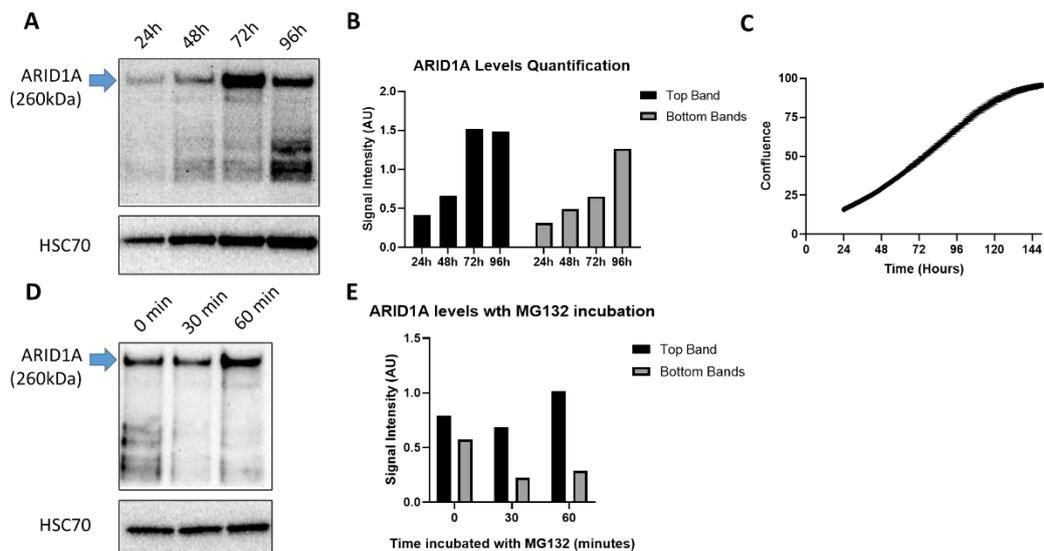


Figure 5.4 - Optimisation of seeding densities

A. Immunoblotting PC3 lysates for ARID1A at indicated timepoints of 24, 48, 72, and 96 hours. Representative blot of three independent experimental replicates. **B.** Quantified ARID1A band intensity of blot A. relative to HSC70 loading control. Quantified bands based on representative blot. **C.** PC3 cells were seeded and cell confluency measured by Incucyte Zoom Cell Imager every hour for 140h. **D.** Immunoblotting PC3 lysates after incubating cells with MG132 for indicated timepoints of 0, 30, and 60 minutes. Representative blot of three independent experimental replicates. **E.** Quantifying ARID1A band intensity of blot D. relative to HSC70 loading control. Quantified bands based on representative blot.

5.3.3 Initial characterising ARID1A knockdown in PC cell lines

Western blotting for ARID1A and ARID1B in a range of PC cell lines to understand the endogenous levels of these proteins was carried out (Figure 5.5). The amount of these proteins could be compared to RWPE-1, an immortalised prostate epithelial cell which represents benign tissue. From this, no correlation between ARID1A levels regarding PTEN or AR status was observed. RWPE and C4-2 had the highest levels of ARID1A, and only CWR22 and 22Rv1, which are both PTEN positive and AR positive, showed loss of ARID1A with all other cell lines having some detectable ARID1A. ARID1B levels did not seem to correlate with ARID1A levels either, however CWR22 showed a loss of ARID1B. This is quite unexpected as loss of both ARID1A and ARID1B has been shown to globally dysregulate chromatin organisation, although a dysfunctional BAF previously been reported in other cancers and cell lines (Kelso, Porter *et al.* 2017).

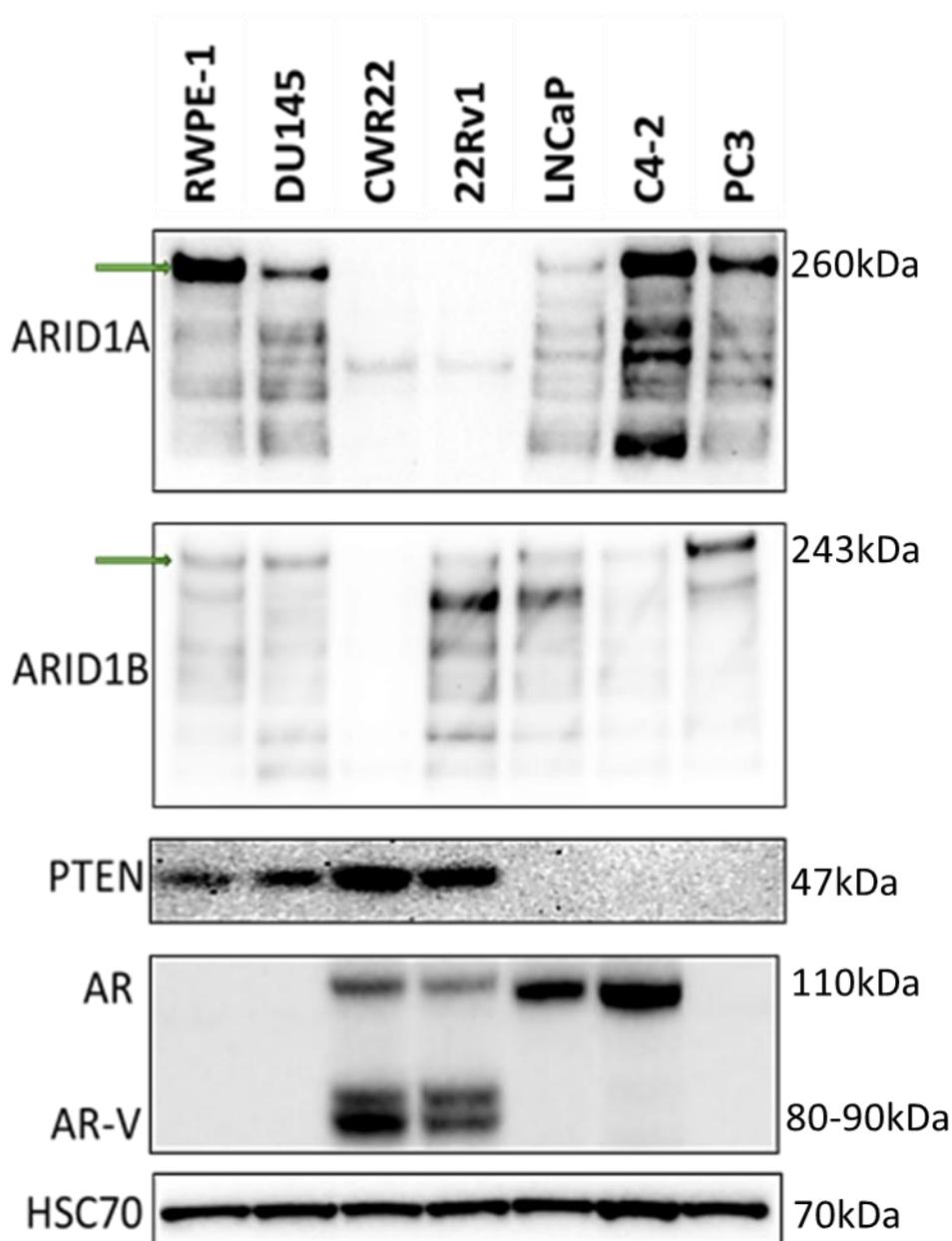


Figure 5.5 - Characterising ARID protein levels in prostate cancer cell lines

Indicated cell lines were seeded at 2×10^5 for 24h before lysis. Lysates were blotted for indicated proteins to determine protein levels relative to a HSC70 loading control. Representative blot from three independent experimental replicates.

From here, we wanted to investigate the effects of ARID1A knockdown on these cell lines (Figure 5.6). We chose DU145 (PTEN positive, AR negative), PC3 (PTEN negative, AR negative), and LNCaP (PTEN negative, AR positive) to investigate as they show a range of the well characterised drivers of PC and are all well studied cell lines in the field. Knockdown of ARID1A by siRNA resulted in a good

knockdown which removed all detectable ARID1A protein (Figure 5.6). This knockdown was also validated by QRT-PCR which also demonstrated a good knockdown in each cell line. This also shows endogenous ARID1A expression is highest in PC3, and of similar levels in both DU145 and LNCaP.

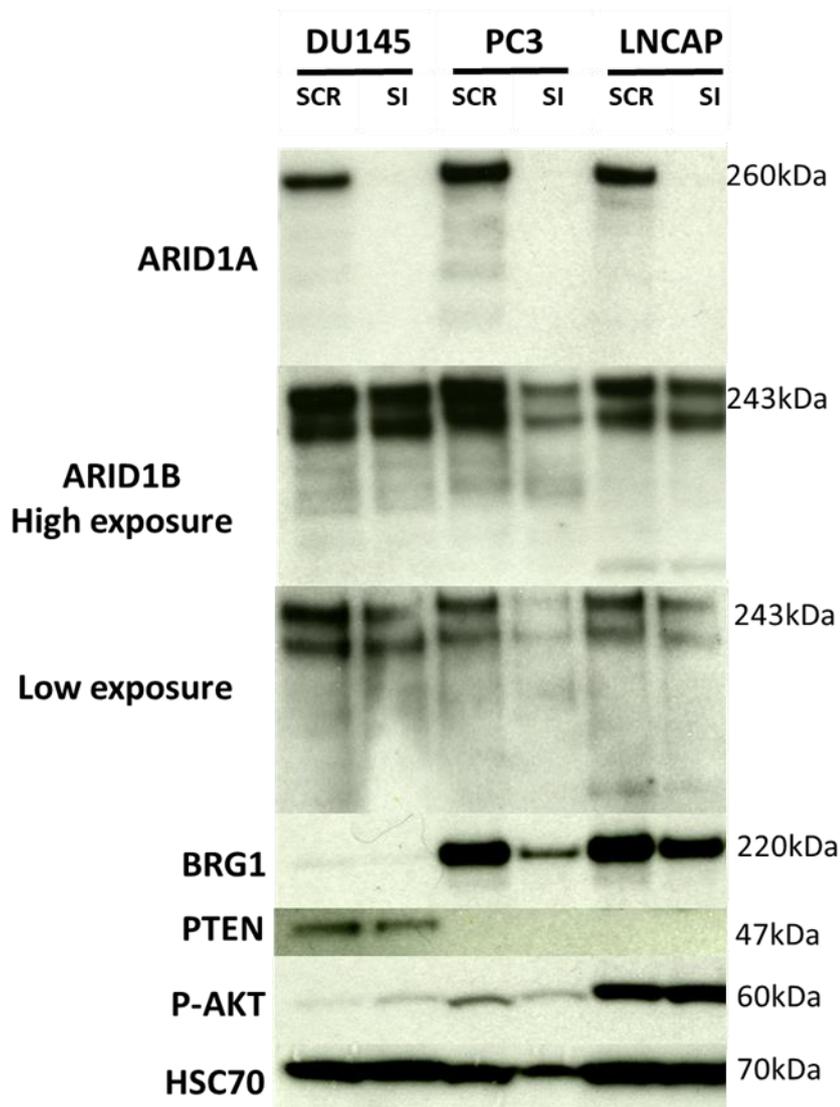


Figure 5.6 - Validating ARID1A protein knockdown from transfection

Indicated cell lines were seeded at 2×10^5 and reverse transfected with either a non-targeting (SCR) or siARID1A (SI) for 72h before lysis. Lysates were blotted for indicated proteins to determine protein levels relative to a HSC70 loading control. Representative blot based on three independent experimental replicates.

As we expect these cells to become dependent on ARID1B after ARID1A

knockdown, we blotted for ARID1B and saw no changes in its levels or banding

pattern following ARID1A knockdown. This indicated that the cells did not have to compensate for the loss of ARID1A and the endogenous ARID1B was sufficient. Interestingly, the protein levels of BRG1, one of the catalytic subunits of the BAF complex, are reduced following ARID1A knockdown. A previous study showed that BRG1 is stabilised by AKT activity following PTEN loss in PC (Ding, Li *et al.* 2019). This explains why DU145 do not show high levels of BRG1 as it is degraded. As BRG1 is reduced following ARID1A knockdown, we blotted for Ser473 phosphorylation of AKT to indicate a change in AKT activity. This showed no change in AKT phosphorylation following ARID1A loss indicating that BRG1's stability was being altered independent of this. ARID1A has also been shown to elevate AKT phosphorylation in breast and ovarian cancers by upregulating ANXA1 and downregulating PI3KIP1 respectively. However, as stated, AKT phosphorylation is not altered following ARID1A knockdown in these PC cell lines.

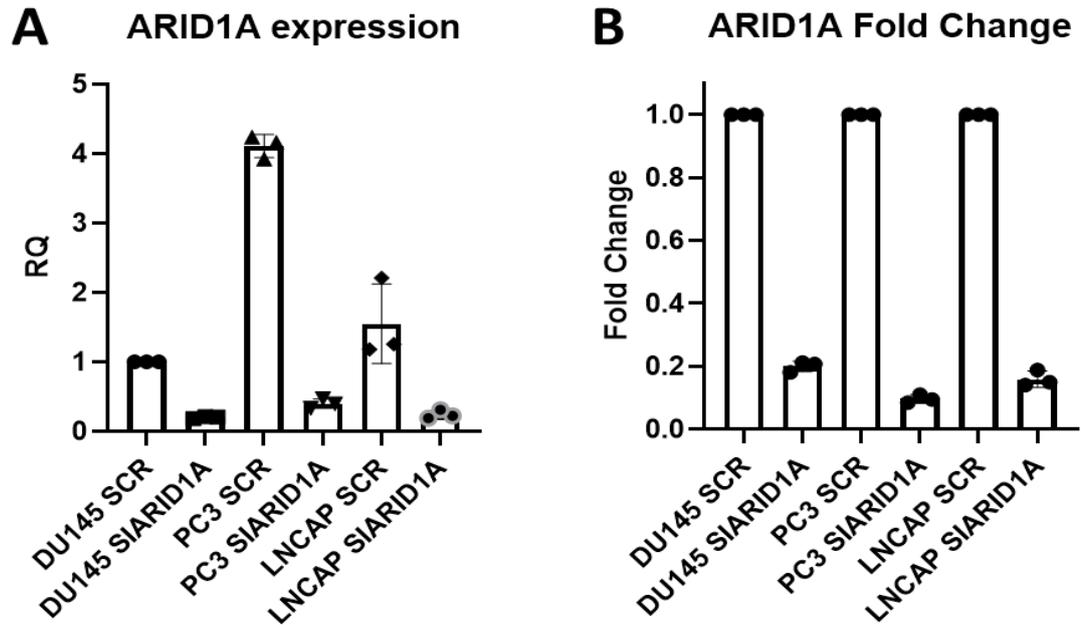


Figure 5.7 - ARID1A expression in PC cell lines

A. Indicated cell lines seeded at 1×10^5 per well of a 6 well plate followed by reverse transfection with either non-targeting control (SCR) or siARID1A (SI). Cells left for 72h prior to RNA extraction. Endogenous ARID1A levels quantified by QRT-PCR to validate knockdown (N=3). **B.** ARID1A levels quantified by QRT-PCR normalised to non-targeting control for each cell line as indicated. Each point represents one experimental replicate made up of three technical replicates.

5.3.4 Phenotypic characterisation of ARID1A knockdown in PC cell lines

For the initial phenotypic characterisation, ARID1A was knocked down with siRNA and the growth of these cells investigated by cell count (Figure 5.8).

Surprisingly, each cell line saw reduced cell numbers after 72h with the siARID1A indicating that the reduction of ARID1A attenuated growth. We also performed a colony forming assay which similarly also saw reduced colony number and colony size following knockdown of ARID1A (Figure 5.9). This data indicated that loss of ARID1A reduced the growth and survival of each of these cells irrespective of PTEN or AR status.

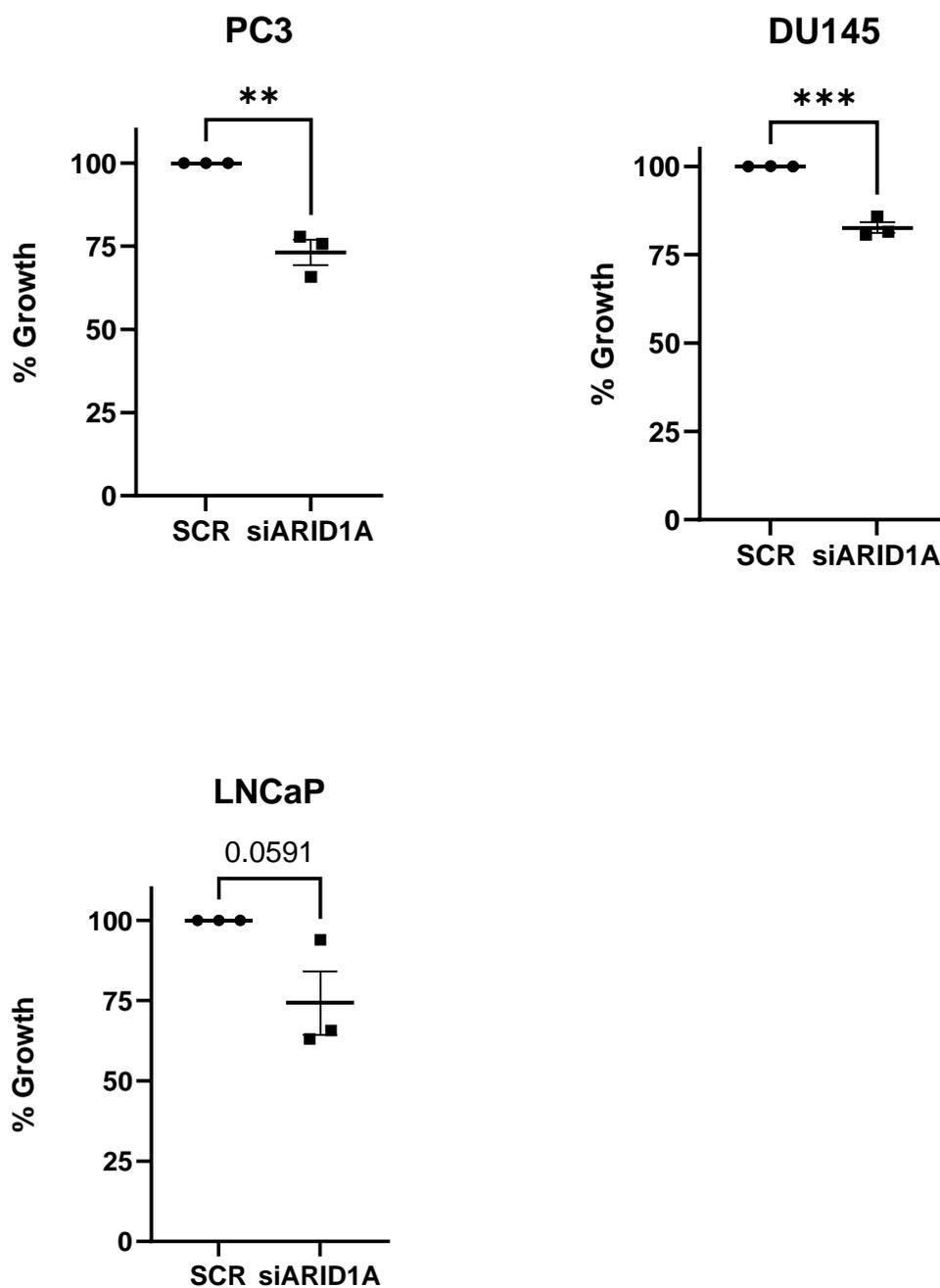


Figure 5.8 - Transient of knockdown of ARID1A reduces growth in PC cell lines

Cells seeded at 5×10^4 per well in three technical replicates and reverse transfected with either non-targeting control (SCR) or siARID1A (SI). After 72h cells were counted and cell counts were normalised to SCR control. Each dot represents one independent experimental replicates each made up of three technical replicates. ** = $P < 0.01$, *** = $P < 0.005$ as determined by Mann-Whitney test.

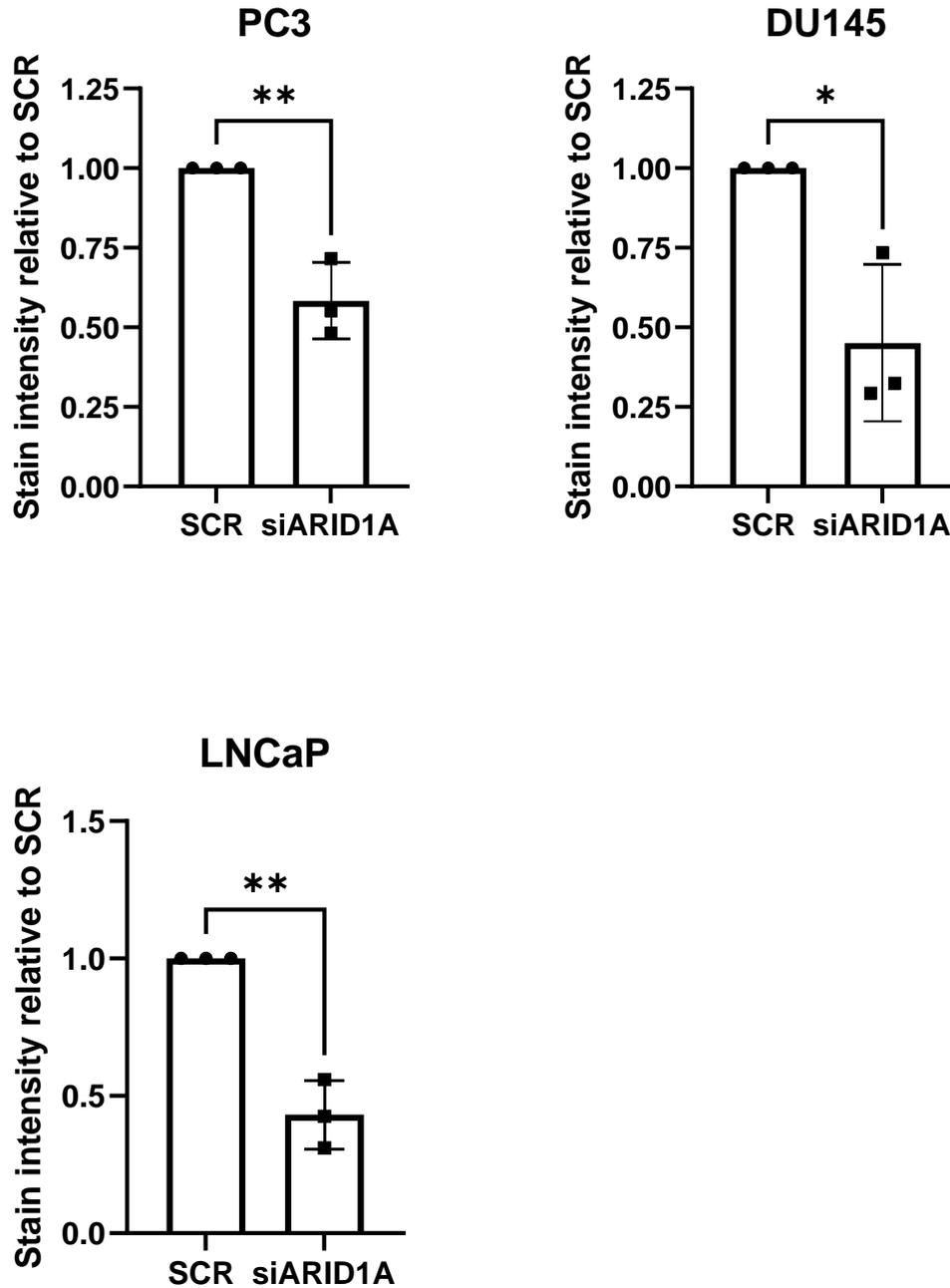


Figure 5.9 - Transient knockdown of ARID1A reduces colony forming assays in PC cell lines

Cell lines have ARID1A knocked down and seeded as single cells. After colony growth cells stained with crystal violet and quantified by LI-COR. Approximately 2 weeks for LNCaP and 8-10 days for DU145 and PC3. Each point represents one independent experimental replicate made up of three technical replicates. *= $P < 0.05$, **= $P < 0.01$ as determined by Mann-Whitney test.

5.3.5 Generating ARID1A knockout clones of human cell lines

Knocking out ARID1A completely by CRISPR-Cas9 may be able to produce a stable cell which gains the advantages of ARID1A loss without the stress of transient knockdown disrupting the chromatin landscape. The cells can be transfected with the ARID1A CRISPR-Cas9 knockout plasmid, and a HDR plasmid which provides an ARID1A template for homologous recombination which disrupts ARID1A expression and incorporates puromycin resistance. To allow for selection of knockout (KO) clones by puromycin, a kill curve was made to select the correct dose of puromycin to kill cells which did not incorporate the plasmid (Figure 5.10). From these puromycin kill curves, we chose puromycin at 2ug/ml to allow selection of puromycin resistance.

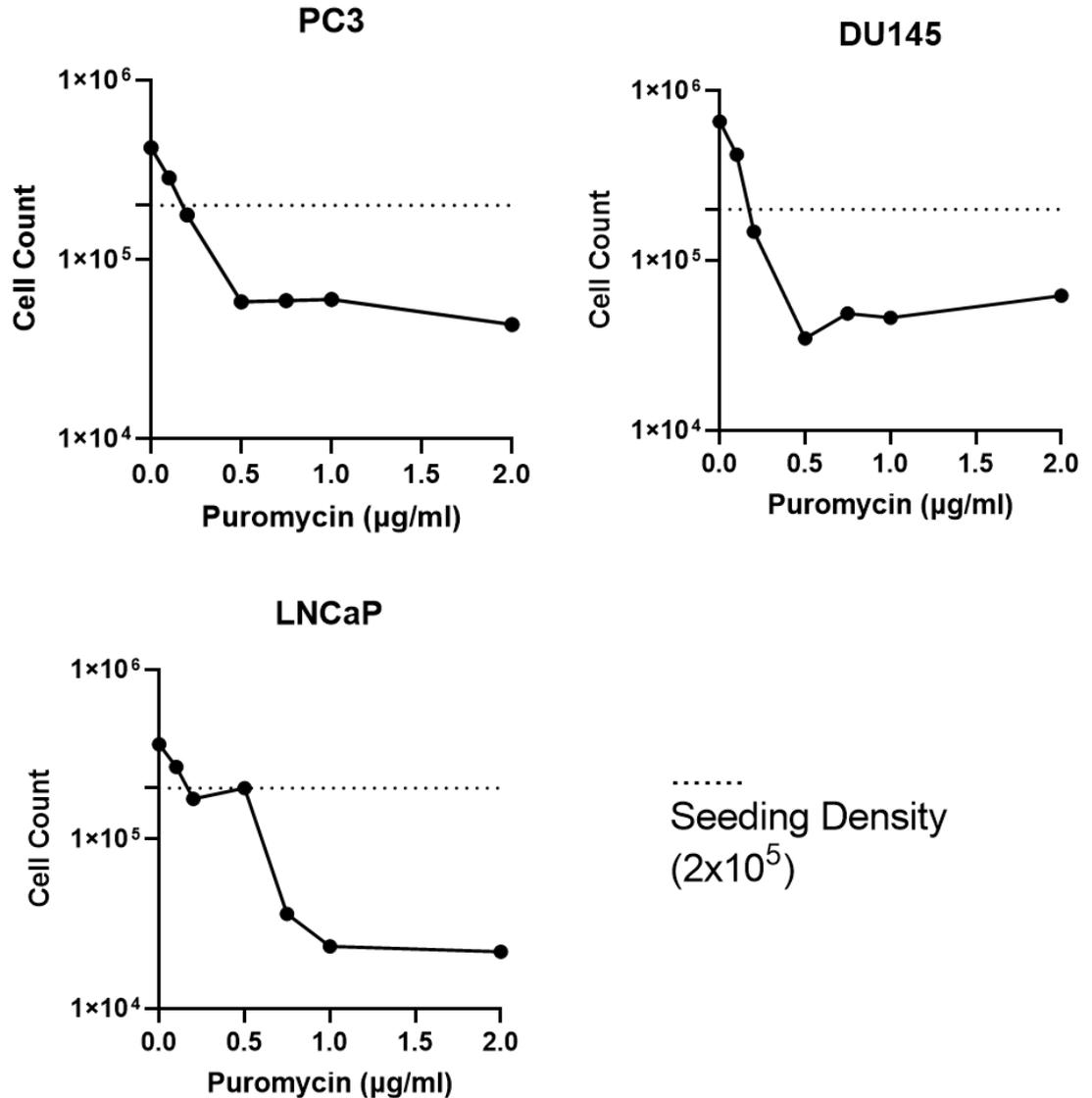


Figure 5.10 - Optimisation of puromycin kill curve

Cells seeded at 2×10^5 and incubated with indicated concentration of puromycin. Cells were then trypsinised and counted in solution by CASY-Counter after 48 hours.

Cells were then transfected by nucleofactor electrophoresis with either ARID1A KO CRISPR-Cas9 plasmid and ARID1A HDR plasmid, or non-targeting CRISPR-Cas9 and IRFP plasmid. However, after selection by puromycin, only DU145 produced colonies indicating puromycin resistance. To validate the transfection procedure worked for PC3 and LNCaP, we imaged them to look for GFP fluorescence as both the ARID1A KO CRISPR-Cas9 and the non-targeting CRISPR-Cas9 plasmid

express GFP (Figure 5.11). This showed that 24h after transfection, both LNCaP and PC3 were able to begin expression of the CRISPR-Cas9 plasmids. However, after 72h as the cells underwent puromycin selection, all GFP signal disappeared in both LNCaP and PC3 while remaining in DU145 only. This indicates that neither PC3 nor LNCaP were stable following ARID1A knockdown, unlike DU145.

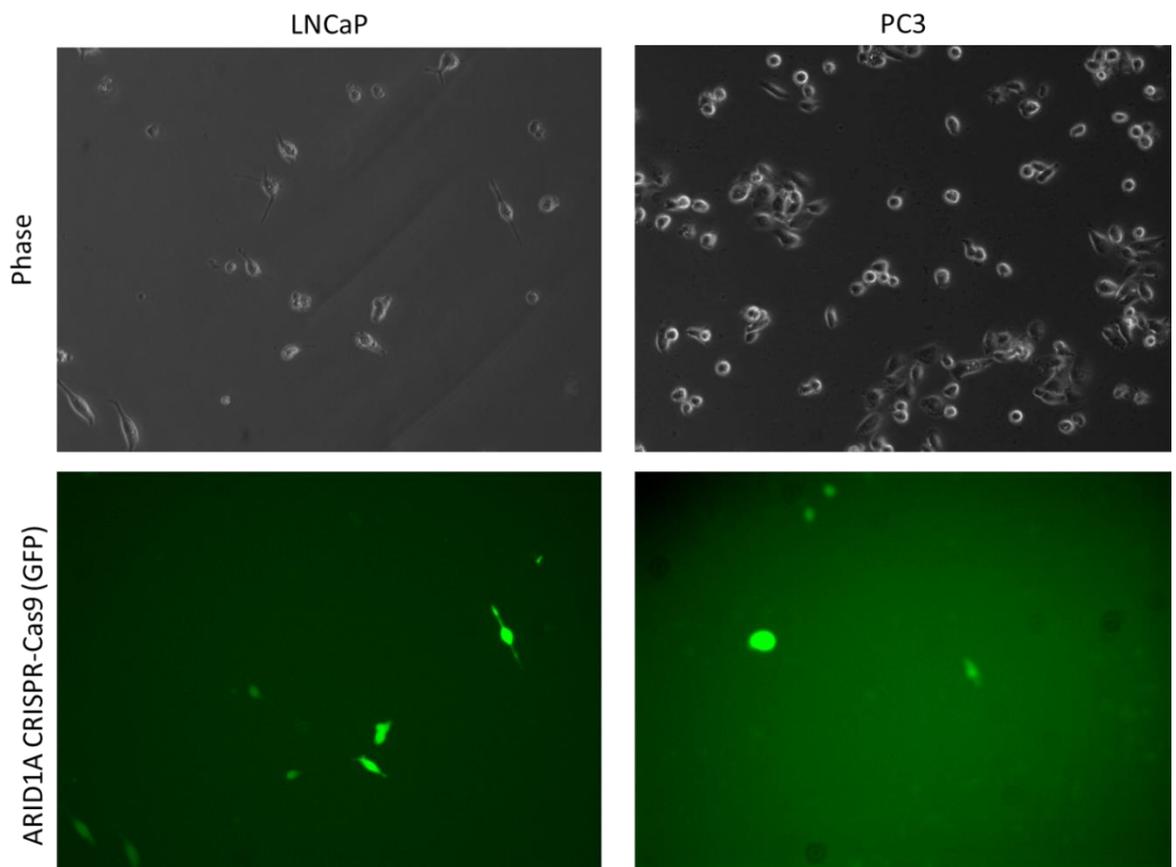


Figure 5.11 - ARID1A CRISPR-Cas9 plasmid expresses in LNCaP and PC3

Indicated cell lines transfected with ARID1A CRISPR-CAS9 and visualised with GFP signal.

Next, individual DU145 clones were selected for ARID1A by western blotting for loss of ARID1A protein, and by confirming reduction in ARID1A expression by qRT-PCR. Empty-vector (EV) clones were also selected as controls which retain wildtype ARID1A protein levels and expression while gaining puromycin resistance (Figure 5.12). From here, we got 2 empty vector clones (EV1 and

EV2), and 2 KO clones (KO2 and KO4). Any KO clones which showed a reduction, but not a complete loss, of ARID1A were pooled together to form the DU145 KO Pool.

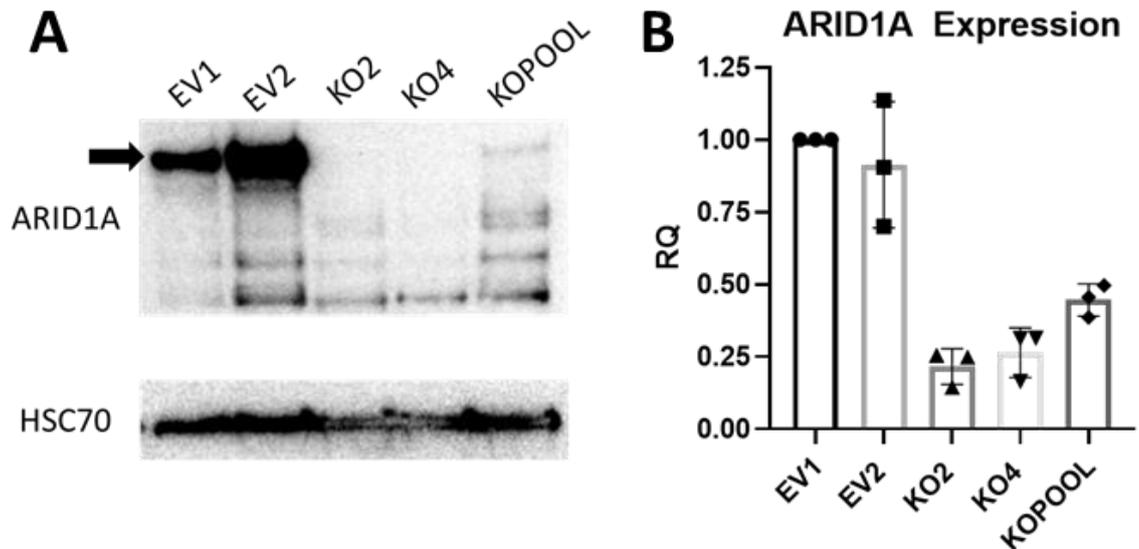


Figure 5.12 - Validating ARID1A knockout of DU145 clones

A. Immunoblotting for ARID1A knockout clones for DU145 cells transfected with CRISPR-Cas9 empty vector (EV) or CRISPR Cas9 ARID1A (KO). HSC70 used as loading control. Representative blot of three independent experimental replicates. **B.** QRT-PCR for ARID1A levels for DU145 cells transfected with CRISPR-Cas9 empty vector (EV) or CRISPR Cas9 ARID1A (KO). Each point represents one independent experimental replicate made up of three technical replicates.

5.3.6 Characterising DU145 ARID1A KO clones by RNA-Seq

To understand which pathways and transcriptional networks had changed following loss of ARID1A, we performed RNA-Seq. RNA was extracted from DU145 EV1, DU145 ARID1A KO2, and DU145 ARID1A KO4 cell lines and processed into a library with 5 independent experimental replicates per cell line. The quality was assessed on tapestation (Figure 5.13A) to validate the library generation. This gel shows a single large band in each of the samples indicating a uniform library

was produced. These samples were then sequenced, and variance analysed by PCA (Figure 5.13B). Analysis was performed by Mr Robin Shaw (Informatician, Beatson Institute for Cancer Research). The PCA analysis shows a clear distinction between each of the 3 cell lines, though the largest difference is between the WT ARID1A and KO ARID1A. There are 2241 significantly altered shared genes between KO2 vs EV1 and KO4 vs EV1. KO2 has 1010 unique significantly altered genes, and KO4 has 656 unique significantly altered genes (Figure 5.13C).

When plotting the genes significantly up or downregulated it was found that 10% more genes are downregulated compared to upregulated in KO2 vs EV1, while 40% more genes are upregulated compared to downregulated in KO4 vs EV1 (Figure 5.13D & Figure 5.13E). This shows there are subtle differences between these clones. This is interesting compared to the murine samples which downregulated 300% more genes than upregulated.

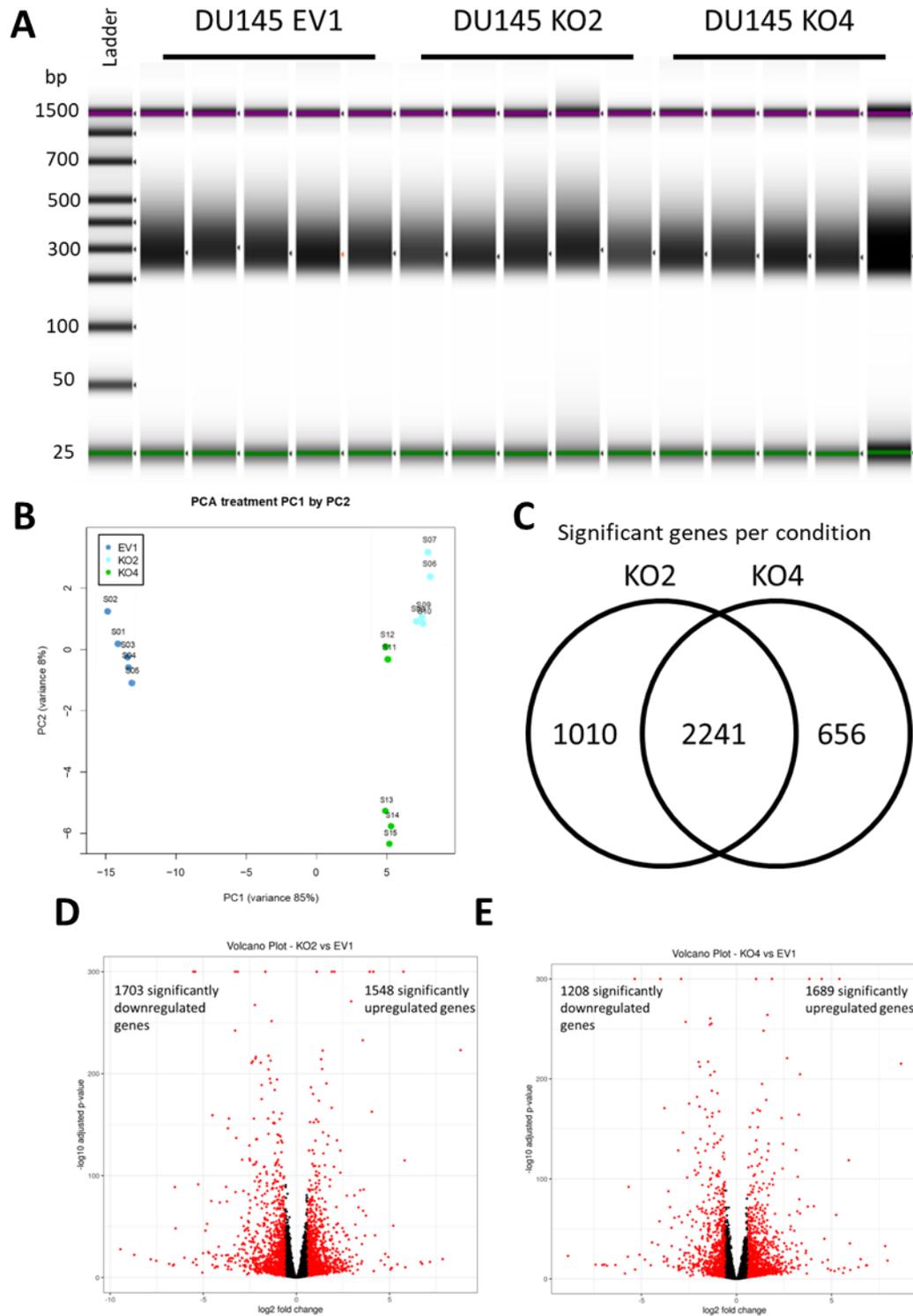


Figure 5.13 – RNA-Seq library preparation and characterisation

A. RNA libraries prepared for indicated DU145 clones with either wildtype ARID1A expression (EV1) or ARID1A knockout (KO) and ran on tapestation to validate library size. **B.** Principal component analysis (PCA) plot showing variation between biological replicates for each DU145 clone. **C.** Number of significantly enriched genes in each DU145 ARID1A knockout clone relative to EV1. **D.** Volcano plot showing significantly enriched genes in red from KO2 vs EV1. Genes with a base mean below 10 counts, absolute fold change below 0.5, and/or P-Value above 0.8 have been removed from the plot for clarity. **E.** Volcano plot showing significantly enriched genes in red from

KO4 vs EV1. Genes with a base mean below 10 counts, absolute fold change below 0.5, and/or P-Value above 0.8 have been removed from the plot for clarity.

The significantly enriched genes from the RNA-Sequencing were imported into Metacore, a data analysis software, to investigate whether the genes were part of similar pathways or signalling networks (Figure 5.14).

- DU145 KO2 vs DU145 EV1
- DU145 KO4 vs DU145 EV1

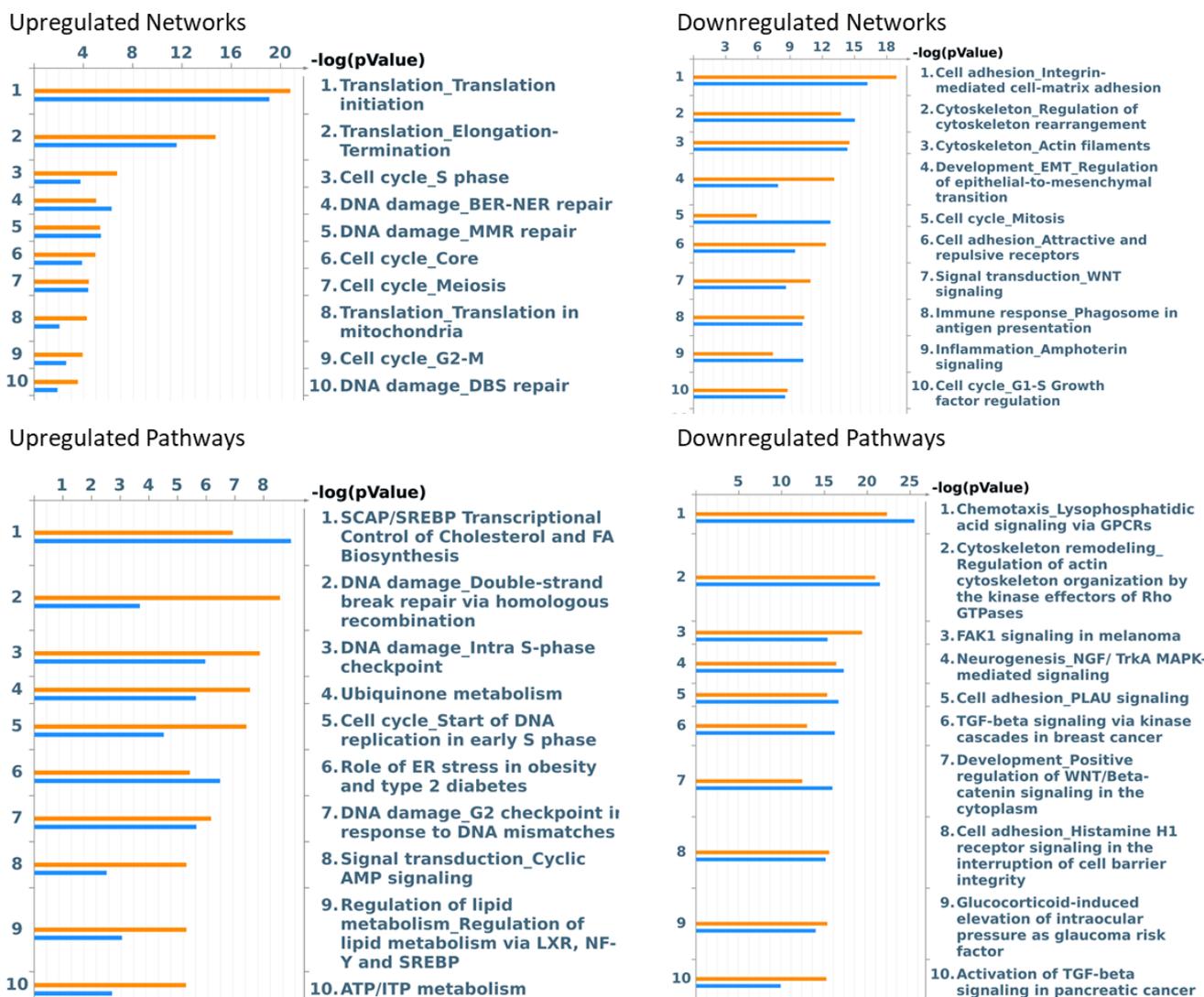


Figure 5.14 - Pathway analysis for DU145 KO clones

Significantly altered genes between DU145 EV1 and DU145 ARID1A KO2 and KO4 were uploaded to Metacore and pathway and network analysis was performed. Genes were filtered by threshold of 1.5 and P-value <0.05.

The most upregulated networks in DU145 following loss of ARID1A appeared to be cell cycle signalling, similar to the murine model. This also included elevated translation initiation and elongation suggesting global protein synthesis is upregulated. When we look at specific pathways, we see that cell cycle is still upregulated, but specifically fatty acid and lipid metabolism via sterol regulatory element binding protein 1 and 2 (SREBP1/2) are also upregulated. These pathways show key regulators of fatty acid synthesis such as fatty acid synthase (FASN), acyl-coa synthase (ACYL), and ELVOVL fatty acid elongase 6 (ELOVL6) are all upregulated by 30-40%. Similarly, other growth pathways such as ubiquinone and ATP/ITP metabolism are upregulated which suggest elevation in mitochondrial biogenesis. All together these pathways suggest a strong elevation in growth signalling which supports the upregulation in cell cycle signalling networks as well as the increased growth phenotype.

The downregulated networks and pathways clearly indicate a downregulation of certain cytoskeleton remodelling and cell adhesion signalling. Some pathways suggest extracellular matrix remodelling is altered to promote migration. This is shown with collagen IV A1-4 all being downregulated 2-4-fold, while matrix metalloproteases (MMPs) 1, 9, and 13 are all upregulated 2-13-fold. Meanwhile tissue inhibitor metalloproteases (TIMPs) which inhibit MMP activity are down 2-fold. These indicate that less collagen IV, which is a crucial component in basement membranes, is secreted while more is degraded by MMP activity. Meanwhile, focal adhesion kinase (FAK1) and many of its upstream integrins are downregulated which would suggest a downregulation in migration. Other regulators of cytoskeleton regulation such as RhoA and Rac1 are downregulated

10 -20% indicating a dysregulation of actin remodelling which would normally promote cell migration.

5.3.7 Phenotypic analysis of ARID1A KO DU145 clones

To understand the effect that ARID1A loss has had on the cells the growth, survival, and migration of the ARID1A KO DU145 cells was characterised (Figure 5.15). Growth rate was investigated by cell count which showed that loss of ARID1A from the DU145 elevated the growth rate of these cells (Figure 5.15A). Survival by colony forming assay showed that loss of ARID1A formed more colonies and larger colonies (Figure 5.15B) while migration by wound scratch assay showed wound closure was faster following loss of ARID1A (Figure 5.15C). The velocity of wound closure at the edge of the wound showed elevated migration rate in ARID1A KO DU145 cell lines (Figure 5.15D).

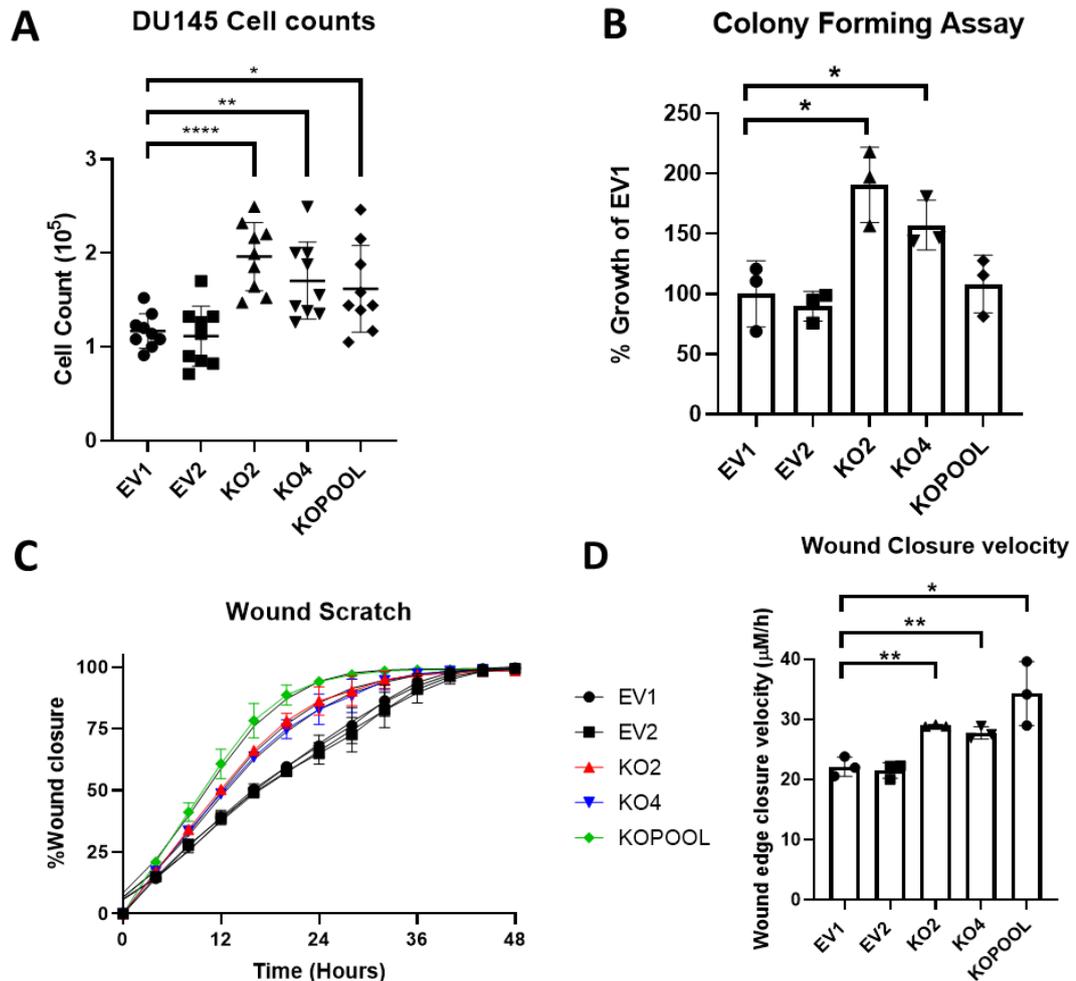


Figure 5.15 - ARID1A knockout increases growth, survival, and wound healing in DU145 cells

A. Graph showing cell counts after 72h. N=3 independent experimental replicates and N=9 technical replicates of ARID1A knockout clones for DU145 cells transfected with CRISPR-Cas9 empty vector (EV1 and EV2) or CRISPR Cas9 ARID1A (KO2, KO4, and KOPOOL). Each data point shows technical replicates. **B.** Colony forming assay showing crystal violet-stained colonies quantified by LI-COR, data represented as percentage growth relative to EV1. Each data point represents one independent experimental replicate made up two technical replicates. **C.** Wound healing assay showing wound closure over time with 100% wound closure representing a monolayer of cells. Each datapoint represents one biological replicate made up of eight technical replicates. **D.** Wound closure velocity calculated by time taken to close the wound. Velocity at fastest rate was used, between 0 and 16 hours. *= $P < 0.05$, **= $P < 0.01$, ****= $P < 0.0001$, as determined by one-way ANOVA.

This data suggests that loss of ARID1A upregulates growth, survival, and migration in DU145 cells. As transient ARID1A knockdown did not have this effect, this may suggest that the instability of the cells following knockdown did not allow for the tumorigenic advantages of ARID1A loss.

However, RNA-Seq analysis indicated a reduction in expression of genes involved in migration however, this phenotypic data demonstrated that loss of ARID1A elevated migration in wound healing assay. This potentially indicates that the apparent elevation in migration was caused by the elevated growth rate of the cells allowing them to close the wound faster. To test this, a transwell migration assay which would allow the cells to move through a membrane towards a chemoattractant was performed. This data indicates that loss of ARID1A does impair the migration ability of DU145 as fewer cells migrated through the membrane (Figure 5.16).

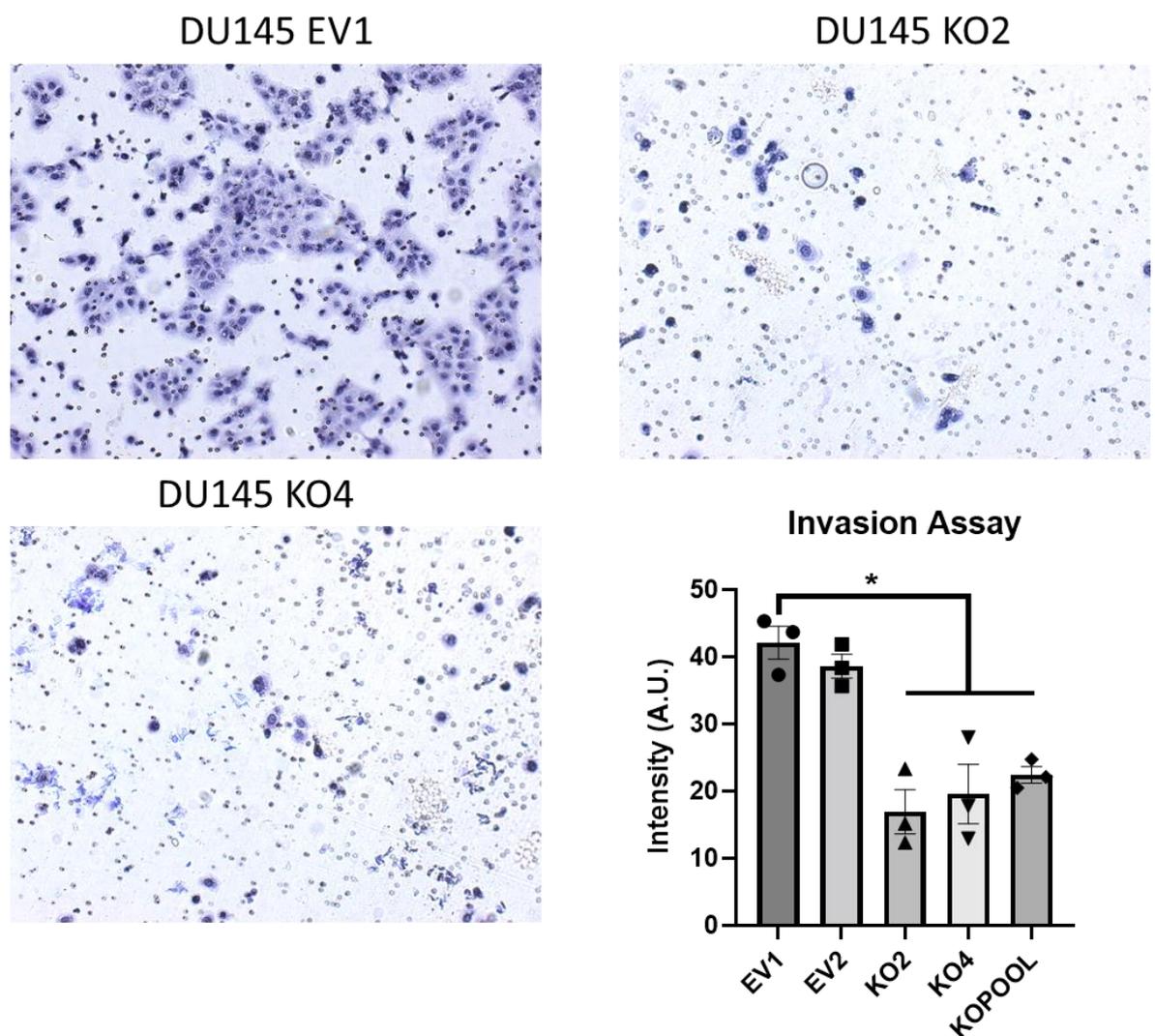


Figure 5.16 - Loss of ARID1A impairs chemotaxis

Representative images of ARID1A knockout clones for DU145 cells transfected with CRISPR-Cas9 empty vector (EV1 and EV2) or CRISPR Cas9 ARID1A (KO2, KO4, and KOPOOL) which were allowed to invade transmembrane and were stained by haemotoxilin. Larger dark stain shows DU145 cells which were capable of invading and attaching to the membrane. Small dark stained areas show membrane itself. Each point represents 5 images which were taken of each membrane per independent experimental replicate for each clone and intensity of image quantified by ImageJ for a total of independent experimental replicates. *= $P < 0.05$ as determined by one-way ANOVA.

5.3.8 Comparing human and murine datasets following ARID1A loss

Comparing RNA-Seq datasets showing dysregulated genes following loss of ARID1A in DU145 cells (this Chapter) and the GEM model (Chapter 4), allowed investigation of shared signalling trends between each of these models. This is done by looking at pathway maps which group pathways by their broad category. From this the two most upregulated pathway maps are relating to DDR and cell cycle (Figure 5.17A). The upregulation in cell cycle signalling in both models is consistent with phenotypic data in each case, with loss of ARID1A increasing cell growth.

The increase in DDR however is quite interesting and was not strongly suggested when analysing these datasets separately. From these pathways, we see these models tend to increase expression of parts of the DDR perhaps due to compensate for the increased replicative stress induced by the higher growth rate. Alternatively, ARID1A has been shown to help mediate an effective DDR so this elevation in DDR machinery could also be compensation for this effect (Shen, Peng et al. 2015).

Interestingly, ARID1A mutant PC has a higher instance of microsatellite instability and higher mutation count when compared to ARID1A wildtype PC (Figure 5.17B and 5.17C). This suggests that this increase in DDR pathways is

likely consistent with clinical findings with ARID1A mutant tumours being more mutagenic potentially as a consequence of ARID1A loss. This aspect of ARID1A has been highlighted in other cancers as a therapeutic target by inhibiting DDR in ARID1A mutant tumours (Shen, Peng et al. 2015, Williamson, Miller et al. 2016). This may suggest that therapeutics inhibiting DDR may be efficacious in treating ARID1A mutant PC.

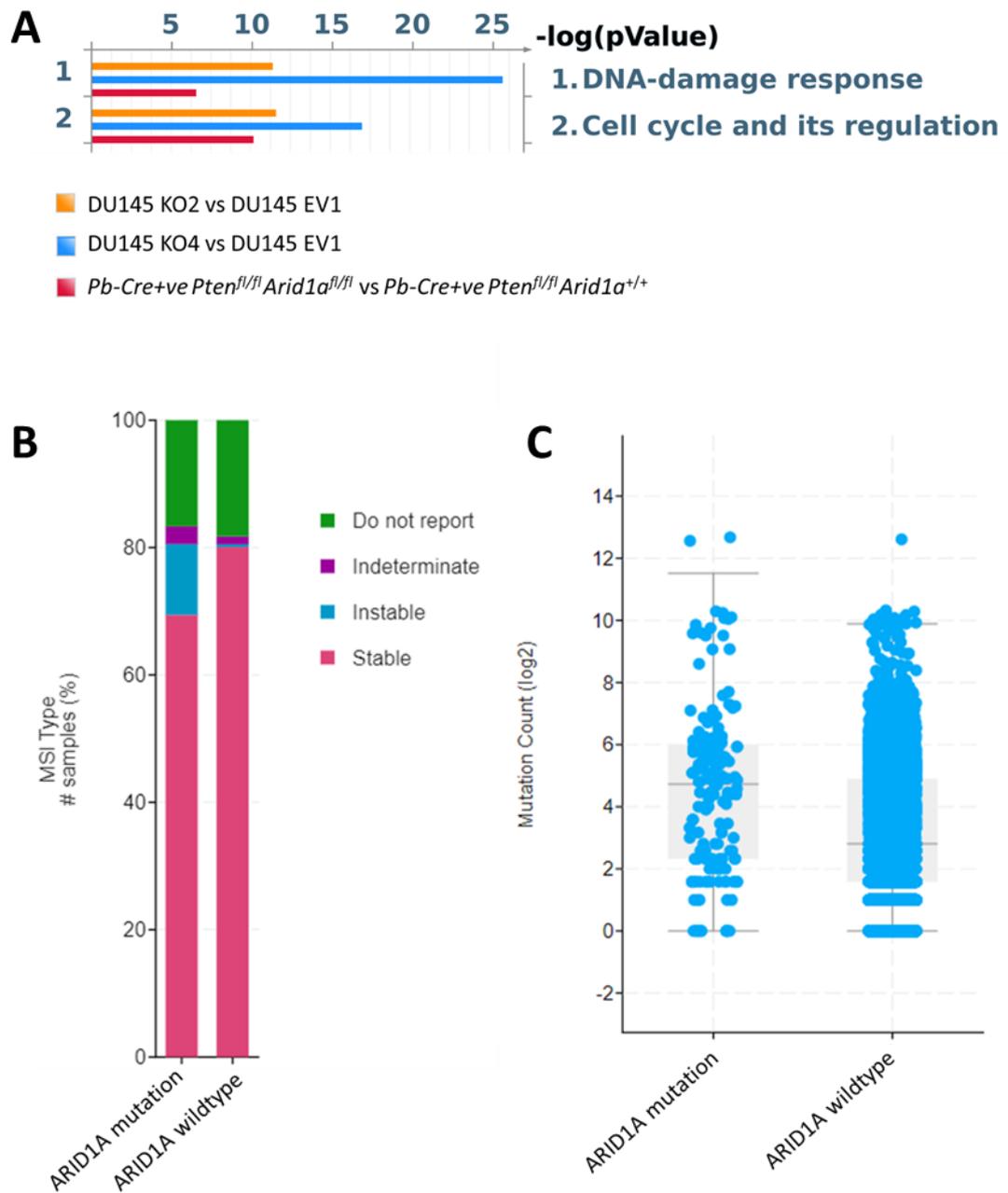


Figure 5.17 - Loss of ARID1A upregulates DDR and cell cycle pathways

A. Comparing three datasets on Metacore up of DU145 KO2 vs EV1, DU145 KO4 vs EV2, and *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* vs *Pb-Cre+ve Pten^{fl/fl} Arid1a^{+/+}* to show enrichment of groups of pathways. **B.** Comparing all prostate cancer datasets on CBioPortal with or without ARID1A mutation for alterations in microsatellite instability. **C.** Comparing all prostate cancer datasets on CBioPortal with or without ARID1A mutation for mutation count

5.4 Discussion

5.4.1 LNCaP and PC3 dependency on ARID1A

ARID1A has previously been investigated in some human PC cell lines which each suggested that ARID1A can mediate certain oncogenic signalling pathways such as AR, TMPRSS2-ERG, and β -catenin (Sandoval, Pulice *et al.* 2018, Stelloo, Nevedomskaya *et al.* 2018, Hiramatsu, Fukuda *et al.* 2019). This is likely the reason that LNCaP and PC3 ARID1A knockout clones could not be generated using CRISPR-Cas9 as they were dependent on one or more of these ARID1A-mediated pathways.

As the mouse model demonstrated that loss of ARID1A was tumorigenic, this data was surprising, particularly as both LNCaP and PC3 are PTEN negative cell lines. One study shows that acute and chronic loss of ARID1A alter distinctly different pathways (Blümli, Wiechens *et al.* 2021). One hypothesis to explain this is that loss of ARID1A destabilises immediate chromatin architecture and resultant changes further altered the chromatin landscape over time. A chronic loss of ARID1A will potentially come to an equilibrium which is less dynamic and permissive for growth. The initial changes that occur with a dynamically changing chromatin landscape are likely not beneficial to growth and if any ARID1A remains in the system it will further impact cell growth. This is a potential explanation for why DU145, PC3, and LNCaP all did not tolerate ARID1A

knockdown transiently. With chronic loss of ARID1A also not being tolerated by LNCaP and PC3.

DU145 however, a PTEN positive cell line, could have ARID1A knockout clones generated suggesting they are either not dependent on these pathways or are capable of being driven by other growth pathways that emerge after ARID1A loss. These cells still did not respond well to transient ARID1A knockdown suggesting there may be this acute vs chronic effect of ARID1A loss.

Interestingly, CWR22 and 22Rv1 both have endogenously lost ARID1A. As these cell lines are dependent on AR signalling, either through full-length or variant AR, this indicates ARID1A loss does not prevent AR activity (Dagvadorj, Tan *et al.* 2008). This is surprising as AR signalling in LNCaP has been shown to require ARID1A. This therefore raises the question why AR is dependent on ARID1A in LNCaP and not CWR22 and 22Rv1. Molecular differences between LNCaP and CWR22/22Rv1 cell lines are summarised in table 5.2.

Cell Feature	LNCaP	CWR22 and 22Rv1
PTEN Status	Negative	Positive
AR status	Full-length only	Full-length and variant

P53 status	WT/WT	WT/Q33R1 (truncation mutation)
Prostate Cancer stage	Castrate-sensitive	Castrate-resistant

Table 5.2 – Molecular drivers and characteristics of LNCaP vs CWR22

AR-variants are not likely to be independent of BAF remodelling. AR-variants largely overlap with full-length AR target genes, and while AR-variants do have some unique target genes, the overlapped pathways are retained in CWR22 (Lu, Lonergan *et al.* 2015). Another hypothesis relates to PTEN status, as both CWR22, 22Rv1 and DU145 have lost or can tolerate loss of ARID1A and have PTEN, while LNCaP, PC3, and VCaP have lost PTEN and cannot tolerate ARID1A loss. However, CBioPortal does not suggest a mutual exclusivity between PTEN and ARID1A and in fact suggests co-occurrence. Similarly, our mouse model clearly demonstrates synergism between the loss of *Pten* and *Arid1a*. Common PC oncogenic pathways such as AR, TMPRSS2-ERG, and β -catenin have shown ARID1A dependency, despite none of them being dependent on PTEN status. This likely suggests that PTEN loss may not be the deciding factor in ARID1A dependency.

5.4.2 DU145 ARID1A KO cells compared to the *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* mouse model

When comparing these two models there are obvious similarities and differences between them, both in phenotype, and their RNA-Seq datasets. Interestingly, both models upregulate growth and cell cycle related pathways, though the mouse model appears to favour alterations around G2/M transition, whereas the

DU145 favour alterations around G1/S phase transition. This pro-growth signalling is reflected in phenotype also, with both models having elevated proliferation compared to their wildtype ARID1A counterparts.

Motility and migration however seem different when comparing these models, with DU145 KO ARID1A downregulating migration signalling and invading less phenotypically, while the murine model upregulates migration and indeed has an invasive phenotype. However, while DU145 mechanically do migrate less, they have also largely upregulated MMPs, TIMPs, and collagen IV, suggesting they may be capable of priming the cells environment for migration by degrading and remodelling the extracellular matrix. This signalling may facilitate migration *in vivo* of other more motile cancer cells in a heterogenous tumour. The reason for this discrepancy could be attributed to many other factors, as there are other differences between these models, namely PTEN status, AR status, and *in vitro* vs *in vivo* modelling.

Despite this, ARID1A clearly functions as an inhibitor of growth signalling in each of these models. While some cell lines have shown a dependency on ARID1A, DU145 and *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* both clearly treat ARID1A as a tumour suppressor, highlighting a complex context-dependent function of ARID1A in prostate cancer.

5.4.3 Methodological changes and future directions

From this work, we now have clear understanding that ARID1A can function as a tumour suppressor in prostate cancer, predominantly through upregulation in cell cycle signaling and subsequent elevations in cell growth. However, this function is clearly context-dependent, with some other cell lines showing a

dependency on ARID1A. While CBioPortal does not suggest mutual exclusivity of ARID1A mutations with any common oncogenic pathways, experimental data suggests otherwise. Exploring the interactions of ARID1A with these pathways, and understanding the contexts in which ARID1A is lost in patient derived tissues, may yield a better understanding of when ARID1A is mutated in advanced disease and the roles it may play. Further experiments to explore ARID1A in the human cell lines would also be beneficial in understanding the function of ARID1A, such as re-expressing ARID1A in CWR22/22Rv1 or exploring DU145 KO ARID1A *in vivo*.

5.4.4 Conclusions

In this chapter, I have demonstrated that ARID1A can have a context-dependent role in prostate cancer cell lines with some cells showing a dependency on ARID1A. However, upon stably knocking out ARID1A from DU145, these cells also upregulate growth and cell cycle signalling similar to the *Pb-Cre+ve Pten^{fl/fl} Arid1a^{fl/fl}* clearly demonstrating a tumour suppressor effect of ARID1A.

6 Discussion and future directions

6.1 *Arid1a* in a murine GEM model

Arid1a was identified by the *Sleeping Beauty* screen as a novel driver of metastasis in combination with loss of *Pten*. To investigate this, a mouse model was generated to understand the genetic interaction of *Arid1a* and *Pten*. This showed that complete loss of both *Arid1a* and *Pten* caused the development of aggressive and locally invasive PC. This aggressive phenotype only emerged when

both copies of *Arid1a* and *Pten* were lost suggesting that *Pten* loss is required for *Arid1a* loss to be tumorigenic at least in the context of this mouse model.

Following this, I wanted to understand the mechanism by which loss of *Arid1a* drove tumorigenesis following loss of *Pten*. I performed RNA-Seq and CHIP-Seq on the murine tumours and identified that loss of *Arid1a* upregulates cell cycle and growth signalling, particularly around the G2/M phase checkpoint. I also identified that *Arid1a* directly binds and represses the expression of genes involved in migration and invasion. These findings suggest that loss of *Arid1a* upregulates genes involved in cell cycle progression, migration, and invasion. These findings clearly show that *Arid1a* can function as a bone fide tumour suppressor in PC.

6.2 ARID1A in Human Prostate Cancer

To investigate the role of ARID1A in human PC, I developed ARID1A knockout clones in DU145 cells using CRISPR-Cas9. With these cells, I performed RNA-Seq to compare gene expression with and without ARID1A expression. This showed that loss of ARID1A upregulates growth signalling such as fatty acid metabolism, and cell cycle signalling around G1/S phase transition. Interestingly, loss of ARID1A seemed to downregulate migration and invasion signalling. Phenotypic analysis confirmed that loss of ARID1A elevates proliferation and colony forming in DU145 cells, but reduced invasion.

Comparing the GEM model and DU145 ARID1A KO clones shows that loss of ARID1A upregulates growth and cell cycle signalling in both models. While the GEM model shows local invasion, the DU145 shows decreases invasion but may be capable of ECM remodelling and behave differently *in vivo*. These data

demonstrate that ARID1A can function as a tumour suppressor in PC cancer by repressing cell cycle signalling.

6.3 Relevance for the Clinic

As shown in Figure 5.1, ARID1A mutations are most commonly found in metastatic PC, though still comprise a fairly uncommon population at around 3% of instances. However, our murine model and DU145 cell line show that a more aggressive disease can develop following loss of ARID1A. In the case of the murine model, this is shown to be a highly proliferative, locally invasive disease which would likely progress rapidly if untreated. For this reason, ARID1A loss could represent an uncommon and aggressive form of PC which should be identified at diagnosis allowing for a more radical treatment regime.

As LNCaP and PC3 do not well tolerate ARID1A loss, and given the heterogenous nature of PC, only a small population of a tumour may lose ARID1A. Therefore, identifying ARID1A loss at diagnosis may not be sufficient. Similarly, patients may go on to lose ARID1A as their cancer evolves through treatment resistance or may simply present at clinic with advanced disease. These patients could still benefit from targeted therapies which exploit ARID1A loss. Such therapeutic options have been explored in other cancers, most notably the targeting of the DNA-damage response machinery such as PARP or ATR inhibition (Shen, Peng et al. 2015, Williamson, Miller et al. 2016). Such therapeutics would need to be first validated for their efficacy in treating ARID1A mutant PC. However, as DDR signalling is upregulated in both our GEM and DU145 models, we may see effective use of these therapies in PC too.

6.4 Future directions

Future work to improve our understanding of ARID1A in PC could look at the functionality of the BAF complex following the loss of ARID1A. Following loss of ARID1A, only the ARID1B-BAF complex will remain which will have a distinct role compared to the ARID1A-BAF. We have demonstrated that loss of ARID1A upregulates cell cycle signalling, however how the localisation of the BAF complex changes when ARID1B is incorporated will improve our understanding of this mechanism. As ARID1A was not seen to directly repress genes involved in cell cycle, this may indicate that the relative increase in ARID1B-BAF is what causes this increase in cell cycle signalling. To investigate this, ChIP-Seq of other BAF subunits such as ARID1B, BRG1, and BRM could be performed which would encompass all other BAF subtypes. By looking at ARID1B localisation which overlaps with BRG1 and BRM, we can understand which genes ARID1B regulates. BRG1 and BRM localisation changes following loss of ARID1A can then show the relative retargeting of the BAF complex itself.

Our ATAC-Seq performed in section 4.3.7 was limited in its analysis by a lack of satisfactory control. Refining and optimising this protocol further would allow us to investigate how chromatin accessibility is altered following ARID1A loss and improve our understanding of ARID1A loss on cancer development. This can be combined with transcription factor motif analysis to understand which transcription factors have their accessibility altered which would help inform us about what drives the phenotype. As we see an increase in cell cycle signalling and E2F1 and Myc genesets following loss of ARID1A, we may then expect to see an increase in chromatin accessibility around the target genes of these transcription factors. Interestingly, we could also investigate those transcription

factors which are coregulated by both ARID1A-BAF and ARID1B-BAF. This would allow us to understand which BAF complex positively or negatively regulates its chromatin accessibility. Alternatively, only some of its target genes may change, thus showing unique ARID1A or ARID1B regulated geneset from a single transcription factor. These aspects of the BAF complex are neither understood or well-studied, so experiments in this area would improve our overall understanding of the BAF complex in cancer.

This ChIP-Seq and ATAC-Seq analysis could also be performed on the ARID1A knockout DU145 clones compared to empty-vector DU145 to compare and contrast the signalling that occurs in each of these models. This cell line could also be modelled *in vivo* by orthotopically injecting these cells into the murine prostate. This would allow us to understand how ARID1A KO DU145 interact with the tumour microenvironment and allow us to probe whether the upregulation of MMPs has a meaningful impact on metastasis.

6.5 Concluding Summary

My PhD thesis has identified that *Arid1a* can function as a tumour suppressor in a GEM model combined with loss of *Pten* which causes the development of locally invasive and proliferative cancer (Chapter 3). Loss of *Arid1a* in these tumours upregulates cell cycle, migration and invasion signalling which may be a result of escaping P53-dependent senescence (Chapter 4). Knockout of ARID1A in DU145 also upregulates cell cycle though downregulates migration. These cells also phenotypically form more colonies and proliferate more, though have reduced migration (Chapter 5).

References

- Abdulkadir, S. A.**, J. A. Magee, T. J. Peters, Z. Kaleem, C. K. Naughton, P. A. Humphrey and J. Milbrandt (2002). "Conditional loss of Nkx3.1 in adult mice induces prostatic intraepithelial neoplasia." Mol Cell Biol **22**(5): 1495-1503.
- Abida, W.**, J. Cyrta, G. Heller, D. Prandi, J. Armenia, I. Coleman, M. Cieslik, M. Benelli, D. Robinson, E. M. Van Allen, A. Sboner, T. Fedrizzi, J. M. Mosquera, B. D. Robinson, N. De Sarkar, L. P. Kunju, S. Tomlins, Y. M. Wu, D. Nava Rodrigues, M. Loda, A. Gopalan, V. E. Reuter, C. C. Pritchard, J. Mateo, D. Bianchini, S. Miranda, S. Carreira, P. Rescigno, J. Filipenko, J. Vinson, R. B. Montgomery, H. Beltran, E. I. Heath, H. I. Scher, P. W. Kantoff, M. E. Taplin, N. Schultz, J. S. deBono, F. Demichelis, P. S. Nelson, M. A. Rubin, A. M.
- Chinnaiyan and C. L. Sawyers** (2019). "Genomic correlates of clinical outcome in advanced prostate cancer." Proc Natl Acad Sci U S A **116**(23): 11428-11436.
- Adhyam, M. and A. K. Gupta (2012). "A Review on the Clinical Utility of PSA in Cancer Prostate." Indian journal of surgical oncology **3**(2): 120-129.
- Ahmad, I.**, E. Mui, L. Galbraith, R. Patel, E. H. Tan, M. Salji, A. G. Rust, P. Repiscak, A. Hedley, E. Markert, C. Loveridge, L. van der Weyden, J. Edwards, O. J. Sansom, D. J. Adams and H. Y. Leung (2016). "Sleeping Beauty screen reveals Pparg activation in metastatic prostate cancer." Proceedings of the National Academy of Sciences **113**(29): 8290.
- Ahmad, I.**, R. Patel, L. B. Singh, C. Nixon, M. Seywright, R. J. Barnetson, V. G. Brunton, W. J. Muller, J. Edwards, O. J. Sansom and H. Y. Leung (2011). "HER2 overcomes PTEN (loss)-induced senescence to cause aggressive prostate cancer." Proceedings of the National Academy of Sciences of the United States of America **108**(39): 16392-16397.
- Alessi, D. R.**, M. Andjelkovic, B. Caudwell, P. Cron, N. Morrice, P. Cohen and B. A. Hemmings (1996). "Mechanism of activation of protein kinase B by insulin and IGF-1." The EMBO journal **15**(23): 6541-6551.
- Alessi, D. R.**, M. Deak, A. Casamayor, F. Barry Caudwell, N. Morrice, D. G. Norman, P. Gaffney, C. B. Reese, C. N. MacDougall, D. Harbison, A. Ashworth and M. Bownes (1997). "3-Phosphoinositide-dependent protein kinase-1 (PDK1): structural and functional homology with the Drosophila DSTPK61 kinase." Current Biology **7**(10): 776-789.
- Alfert, A., N. Moreno and K. Kerl (2019). "The BAF complex in development and disease." Epigenetics & Chromatin **12**(1): 19.
- Auerbach, C.** (1949). "Chemical mutagenesis" Biological Reviews **24**(3): 355-391.
- Auger, K. R.**, L. A. Serunian, S. P. Soltoff, P. Libby and L. C. Cantley (1989). "PDGF-dependent tyrosine phosphorylation stimulates production of novel polyphosphoinositides in intact cells." Cell **57**(1): 167-175.
- Aytes, A.**, A. Mitrofanova, C. W. Kinkade, C. Lefebvre, M. Lei, V. Phelan, H. C. LeKaye, J. A. Koutcher, R. D. Cardiff, A. Califano, M. M. Shen and C. Abate-Shen (2013). "ETV4 promotes metastasis in response to activation of PI3-kinase and Ras signaling in a mouse model of advanced prostate cancer." Proc Natl Acad Sci U S A **110**(37): E3506-3515.
- Baig, F. A.**, A. Hamid, T. Mirza and S. Syed (2015). "Ductal and Acinar Adenocarcinoma of Prostate: Morphological and Immunohistochemical Characterization." Oman medical journal **30**(3): 162-166.

Bao, Y., V. Vinciotti, E. Wit and P. A. C. t Hoen (2013). "Accounting for immunoprecipitation efficiencies in the statistical analysis of ChIP-seq data." BMC bioinformatics **14**: 169-169.

Berns, K., A. Sonnenblick, A. Gennissen, S. Brohee, E. M. Hijmans, B. Evers, D. Fumagalli, C. Desmedt, S. Loibl, C. Denkert, P. Neven, W. Guo, F. Zhang, T. A. Knijnenburg, T. Bosse, M. S. van der Heijden, S. Hindriksen, W. Nijkamp, L. F. Wessels, H. Joensuu, G. B. Mills, R. L. Beijersbergen, C. Sotiriou and R. Bernards (2016). "Loss of ARID1A Activates ANXA1, which Serves as a Predictive Biomarker for Trastuzumab Resistance." Clin Cancer Res **22**(21): 5238-5248.

Beshiri, M. L., B. J. Capaldo, R. Lake, A. T. Ku, D. Burner, C. M. Tice, C. Tran, J. Kostas, A. N. Aililin, J. Yin, S. Agarwal, S. A. Morris, F. H. Karzai, T. L. Lotan, W. L. Dahut, A. G. Sowalsky and K. Kelly (2022). "A cancer stem cell population underlies a multi-lineage phenotype and drug resistance in prostate cancer." bioRxiv: 2022.2003.2024.484651.

Bitler, B. G., K. M. Aird, A. Garipov, H. Li, M. Amatangelo, A. V. Kossenkov, D. C. Schultz, Q. Liu, I.-M. Shih, J. R. Conejo-Garcia, D. W. Speicher and R. Zhang (2015). "Synthetic lethality by targeting EZH2 methyltransferase activity in ARID1A-mutated cancers." Nature medicine **21**(3): 231-238.

Blüml, S., N. Wiechens, M.-Y. Wu, V. Singh, M. Gierlinski, G. Schweikert, N. Gilbert, C. Naughton, R. Sundaramoorthy, J. Varghese, R. Gourlay, R. Soares, D. Clark and T. Owen-Hughes (2021). "Acute depletion of the ARID1A subunit of SWI/SNF complexes reveals distinct pathways for activation and repression of transcription." Cell reports **37**(5): 109943-109943.

Bosse, T., N. T. ter Haar, L. M. Seeber, P. J. v Diest, F. J. Hes, H. F. Vasen, R. A. Nout, C. L. Creutzberg, H. Morreau and V. T. Smit (2013). "Loss of ARID1A expression and its relationship with PI3K-Akt pathway alterations, TP53 and microsatellite instability in endometrial cancer." Mod Pathol **26**(11): 1525-1535.

Brammied, J. S., M. Petljak, I. Martincorena, S. P. Williams, L. G. Alonso, A. Dalmases, B. Bellosillo, C. D. Robles-Espinoza, S. Price, S. Barthorpe, P. Tarpey, C. Alifrangis, G. Bignell, J. Vidal, J. Young, L. Stebbings, K. Beal, M. R. Stratton, J. Saez-Rodriguez, M. Garnett, C. Montagut, F. Iorio and U. McDermott (2017). "Genome-wide chemical mutagenesis screens allow unbiased saturation of the cancer genome and identification of drug resistance mutations." Genome research **27**(4): 613-625.

Buenrostro, J. D., P. G. Giresi, L. C. Zaba, H. Y. Chang and W. J. Greenleaf (2013). "Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position." Nature Methods **10**(12): 1213-1218.

Cancer Genome Atlas Research, N., C. Kandoth, N. Schultz, A. D. Cherniack, R. Akbani, Y. Liu, H. Shen, A. G. Robertson, I. Pashtan, R. Shen, C. C. Benz, C. Yau, P. W. Laird, L. Ding, W. Zhang, G. B. Mills, R. Kucherlapati, E. R. Mardis and D. A. Levine (2013). "Integrated genomic characterization of endometrial carcinoma." Nature **497**(7447): 67-73.

Castellano, E. and J. Downward (2011). "RAS Interaction with PI3K: More Than Just Another Effector Pathway." Genes & cancer **2**(3): 261-274.

Chang, F., J. T. Lee, P. M. Navolanic, L. S. Steelman, J. G. Shelton, W. L. Blalock, R. A. Franklin and J. A. McCubrey (2003). "Involvement of PI3K/Akt pathway in cell cycle progression, apoptosis, and neoplastic transformation: a target for cancer chemotherapy." Leukemia **17**(3): 590-603.

Chen, N. and Q. Zhou (2016). "The evolving Gleason grading system." Chinese journal of cancer research = Chung-kuo yen cheng yen chiu **28**(1): 58-64.

Chen, Z., L. C. Trotman, D. Shaffer, H.-K. Lin, Z. A. Dotan, M. Niki, J. A. Koutcher, H. I. Scher, T. Ludwig, W. Gerald, C. Cordon-Cardo and P. P. Pandolfi (2005). "Crucial role of p53-dependent cellular senescence in suppression of Pten-deficient tumorigenesis." Nature **436**(7051): 725-730.

Clegg, N. J., S. S. Couto, J. Wongvipat, H. Hieronymus, B. S. Carver, B. S. Taylor, K. Ellwood-Yen, W. L. Gerald, C. Sander and C. L. Sawyers (2011). "MYC cooperates with AKT in prostate tumorigenesis and alters sensitivity to mTOR inhibitors." PLoS One **6**(3): e17449.

CRUK-Incidence. ([Accessed July 2022]). "PC Incidence UK." Retrieved 10/2020, from <https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/prostate-cancer/incidence>.

CRUK-Mortality. ([Accessed July 2022]). "PC Mortality UK." from <https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/prostate-cancer/mortality>.

Cunha, G. R., A. A. Donjacour, P. S. Cooke, S. Mee, R. M. Bigsby, S. J. Higgins and Y. Sugimura (1987). "The endocrinology and developmental biology of the prostate." Endocr Rev **8**(3): 338-362.

Cyrta, J., A. Augspach, M. R. De Filippo, D. Prandi, P. Thienger, M. Benelli, V. Cooley, R. Bareja, D. Wilkes, S.-S. Chae, P. Cavaliere, N. Dephore, A.-C. Uldry, S. B. Lagache, L. Roma, S. Cohen, M. Jaquet, L. P. Brandt, M. Alshalalfa, L. Puca, A. Sboner, F. Feng, S. Wang, H. Beltran, T. Lotan, M. Spahn, M. Kruithof-de Julio, Y. Chen, K. V. Ballman, F. Demichelis, S. Piscuoglio and M. A. Rubin (2020). "Role of specialized composition of SWI/SNF complexes in prostate cancer lineage plasticity." Nature Communications **11**(1): 5549.

Dagvadorj, A., S.-H. Tan, Z. Liao, L. R. Cavalli, B. R. Haddad and M. T. Nevalainen (2008). "Androgen-regulated and highly tumorigenic human prostate cancer cell line established from a transplantable primary CWR22 tumor." Clinical cancer research : an official journal of the American Association for Cancer Research **14**(19): 6062-6072.

DeNicola, G. M., F. A. Karreth, D. J. Adams and C. C. Wong (2015). "The utility of transposon mutagenesis for cancer studies in the era of genome editing." Genome Biology **16**(1): 229.

Denmeade SR, I. J. (2003). Cellular Organisation of the Normal Prostate

Ding, Y., N. Li, B. Dong, W. Guo, H. Wei, Q. Chen, H. Yuan, Y. Han, H. Chang, S. Kan, X. Wang, Q. Pan, P. Wu, C. Peng, T. Qiu, Q. Li, D. Gao, W. Xue and J. Qin (2019). "Chromatin remodeling ATPase BRG1 and PTEN are synthetic lethal in prostate cancer." J Clin Invest **129**(2): 759-773.

Ding, Z., C.-J. Wu, G. C. Chu, Y. Xiao, D. Ho, J. Zhang, S. R. Perry, E. S. Labrot, X. Wu, R. Lis, Y. Hoshida, D. Hiller, B. Hu, S. Jiang, H. Zheng, A. H. Stegh, K. L. Scott, S. Signoretti, N. Bardeesy, Y. A. Wang, D. E. Hill, T. R. Golub, M. J. Stampfer, W. H. Wong, M. Loda, L. Mucci, L. Chin and R. A. DePinho (2011). "SMAD4-dependent barrier constrains prostate cancer growth and metastatic progression." Nature **470**(7333): 269-273.

Flores-Alcantar, A., A. Gonzalez-Sandoval, D. Escalante-Alcalde and H. Lomeli (2011). "Dynamics of expression of ARID1A and ARID1B subunits in mouse embryos and in cells during the cell cycle." Cell and Tissue Research **345**(1): 137-148.

- Fujita, K.** and N. Nonomura (2019). "Role of Androgen Receptor in Prostate Cancer: A Review." The world journal of men's health **37**(3): 288-295.
- Fund, W. C. R. "PC incidence worldwide." from <https://www.wcrf.org/dietandcancer/cancer-trends/prostate-cancer-statistics>.
- Gao, X.**, P. Tate, P. Hu, R. Tjian, W. C. Skarnes and Z. Wang (2008). "ES cell pluripotency and germ-layer formation require the SWI/SNF chromatin remodeling component BAF250a." Proc Natl Acad Sci U S A **105**(18): 6656-6661.
- Guan, B.**, Y. S. Rahmanto, R.-C. Wu, Y. Wang, Z. Wang, T.-L. Wang and I.-M. Shih (2014). "Roles of deletion of Arid1a, a tumor suppressor, in mouse ovarian tumorigenesis." Journal of the National Cancer Institute **106**(7): dju146.
- Guan, B.**, T. L. Wang and M. Shih (2011). "ARID1A, a factor that promotes formation of SWI/SNF-mediated chromatin remodeling, is a tumor suppressor in gynecologic cancers." Cancer Res **71**(21): 6718-6727.
- Gunner, C.**, A. Gulamhusein and D. J. Rosario (2016). "The modern role of androgen deprivation therapy in the management of localised and locally advanced prostate cancer." Journal of clinical urology **9**(2 Suppl): 24-29.
- Hagiwara, M.**, Y. Yasumizu, N. Yamashita, H. Rajabi, A. Fushimi, M. D. Long, W. Li, A. Bhattacharya, R. Ahmad, M. Oya, S. Liu and D. Kufe (2021). "MUC1-C Activates the BAF (mSWI/SNF) Complex in Prostate Cancer Stem Cells." Cancer Research **81**(4): 1111.
- Hanada, M.**, J. Feng and B. A. Hemmings (2004). "Structure, regulation and function of PKB/AKT--a major therapeutic target." Biochim Biophys Acta **1697**(1-2): 3-16.
- Hare, L. M., T. J. Pheasant, P. M. Waring, K. G. Montgomery, K. M. Kinross, K. Mills, V.
- Roh, J. K. Heath**, R. G. Ramsay, M. Ernst and W. A. Phillips (2014). "Physiological expression of the PI3K-activating mutation Pik3ca(H1047R) combines with Apc loss to promote development of invasive intestinal adenocarcinomas in mice." Biochem J **458**(2): 251-258.
- He, X.**, Z. Zhu, C. Johnson, J. Stoops, A. E. Eaker, W. Bowen and M. C. DeFrances (2008). "PIK3IP1, a negative regulator of PI3K, suppresses the development of hepatocellular carcinoma." Cancer Res **68**(14): 5591-5598.
- Helming, K. C.**, X. Wang, B. G. Wilson, F. Vazquez, J. R. Haswell, H. E. Manchester, Y. Kim, G. V. Kryukov, M. Ghandi, A. J. Aguirre, Z. Jagani, Z. Wang, L. A. Garraway, W. C. Hahn and C. W. M. Roberts (2014). "ARID1B is a specific vulnerability in ARID1A-mutant cancers." Nature medicine **20**(3): 251-254.
- Hiramatsu, Y.**, A. Fukuda, S. Ogawa, N. Goto, K. Ikuta, M. Tsuda, Y. Matsumoto, Y. Kimura, T. Yoshioka, Y. Takada, T. Maruno, Y. Hanyu, T. Tsuruyama, Z. Wang, H. Akiyama, S. Takaishi, H. Miyoshi, M. M. Taketo, T. Chiba and H. Seno (2019). "Arid1a is essential for intestinal stem cells through Sox9 regulation." Proceedings of the National Academy of Sciences **116**(5): 1704-1713.
- Hiramatsu, Y.**, A. Fukuda, S. Ogawa, N. Goto, K. Ikuta, M. Tsuda, Y. Matsumoto, Y. Kimura, T. Yoshioka, Y. Takada, T. Maruno, Y. Hanyu, T. Tsuruyama, Z. Wang, H. Akiyama, S. Takaishi, H. Miyoshi, M. M. Taketo, T. Chiba and H. Seno (2019). "Arid1a is essential for intestinal stem cells through Sox9 regulation." Proc Natl Acad Sci U S A **116**(5): 1704-1713.
- Ho, L.**, J. L. Ronan, J. Wu, B. T. Staahl, L. Chen, A. Kuo, J. Lessard, A. I. Nesvizhskii, J. Ranish and G. R. Crabtree (2009). "An embryonic stem cell chromatin remodeling

complex, esBAF, is essential for embryonic stem cell self-renewal and pluripotency." Proc Natl Acad Sci U S A **106**(13): 5181-5186.

Holgado-Madruga, M., D. K. Moscatello, D. R. Emler, R. Dieterich and A. J. Wong (1997). "Grb2-associated binder-1 mediates phosphatidylinositol 3-kinase activation and the promotion of cell survival by nerve growth factor." Proc Natl Acad Sci U S A **94**(23): 12419-12424.

Hu, X. and Z. Zhang (2016). "Understanding the Genetic Mechanisms of Cancer Drug Resistance Using Genomic Approaches." Trends in Genetics **32**(2): 127-137.
Huang, H.-N., M.-C. Lin, W.-C. Huang, Y.-C. Chiang and K.-T. Kuo (2014). "Loss of ARID1A expression and its relationship with PI3K-Akt pathway alterations and ZNF217 amplification in ovarian clear cell carcinoma." Modern Pathology **27**(7): 983-990.

Jamaspishvili, T., D. M. Berman, A. E. Ross, H. I. Scher, A. M. De Marzo, J. A. Squire and T. L. Lotan (2018). "Clinical implications of PTEN loss in prostate cancer." Nat Rev Urol **15**(4): 222-234.

Jones, S., T.-L. Wang, I.-M. Shih, T.-L. Mao, K. Nakayama, R. Roden, R. Glas, D. Slamon, L. A. Diaz, B. Vogelstein, K. W. Kinzler, V. E. Velculescu and N. Papadopoulos (2010). "Frequent Mutations of Chromatin Remodeling Gene *ARID1A* in Ovarian Clear Cell Carcinoma." Science **330**(6001): 228-231.

Jones, S., T. L. Wang, M. Shih, T. L. Mao, K. Nakayama, R. Roden, R. Glas, D. Slamon, L. A. Diaz, Jr., B. Vogelstein, K. W. Kinzler, V. E. Velculescu and N. Papadopoulos (2010). "Frequent mutations of chromatin remodeling gene *ARID1A* in ovarian clear cell carcinoma." Science **330**(6001): 228-231.

Kaarijärvi, R., H. Kaljunen and K. Ketola (2021). "Molecular and Functional Links between Neurodevelopmental Processes and Treatment-Induced Neuroendocrine Plasticity in Prostate Cancer Progression." Cancers **13**(4): 692.

Kelso, T. W. R., D. K. Porter, M. L. Amaral, M. N. Shokhirev, C. Benner and D. C. Hargreaves (2017). "Chromatin accessibility underlies synthetic lethality of SWI/SNF subunits in *ARID1A*-mutant cancers." Elife **6**.

Kurita, T., R. T. Medina, A. A. Mills and G. R. Cunha (2004). "Role of p63 and basal cells in the prostate." Development **131**(20): 4955.

Liu, S., W. Cao, Y. Niu, J. Luo, Y. Zhao, Z. Hu and C. Zong (2021). "Single-PanIN-seq unveils that *ARID1A* deficiency promotes pancreatic tumorigenesis by attenuating KRAS-induced senescence." Elife **10**.

Lloyd, T., L. Hounsome, A. Mehay, S. Mee, J. Verne and A. Cooper (2015). "Lifetime risk of being diagnosed with, or dying from, prostate cancer by major ethnic group in England 2008–2010." BMC Medicine **13**(1): 171.

Lu, J., P. E. Lonergan, L. P. Nacusi, L. Wang, L. J. Schmidt, Z. Sun, T. Van der Steen, S. A. Boorjian, F. Kosari, G. Vasmatazis, G. G. Klee, S. P. Balk, H. Huang, C. Wang and D. J. Tindall (2015). "The cistrome and gene signature of androgen receptor splice variants in castration resistant prostate cancer cells." The Journal of urology **193**(2): 690-698.

Marker, P. C., A. A. Donjacour, R. Dahiya and G. R. Cunha (2003). "Hormonal, cellular, and molecular control of prostatic development." Developmental Biology **253**(2): 165-174.
Mathur, R., B. H. Alver, A. K. San Roman, B. G. Wilson, X. Wang, A. T. Agoston, P. J. Park, R. A. Shivdasani and C. W. M. Roberts (2017). "*ARID1A* loss impairs enhancer-

mediated gene regulation and drives colon cancer in mice." Nature genetics **49**(2): 296-302.

Miller, R. E., R. Brough, I. Bajrami, C. T. Williamson, S. McDade, J. Campbell, A. Kigozi, R. Rafiq, H. Pemberton, R. Natrajan, J. Joel, H. Astley, C. Mahoney, J. D. Moore, C. Torrance, J. D. Gordan, J. T. Webber, R. S. Levin, K. M. Shokat, S. Bandyopadhyay, C. J. Lord and A. Ashworth (2016). "Synthetic Lethal Targeting of *ARID1A*-Mutant Ovarian Clear Cell Tumors with Dasatinib." Molecular Cancer Therapeutics **15**(7): 1472-1484.

Nagl, N. G., Jr., A. Patsialou, D. S. Haines, P. B. Dallas, G. R. Beck, Jr. and E. Moran (2005). "The p270 (ARID1A/SMARCF1) subunit of mammalian SWI/SNF-related complexes is essential for normal cell cycle arrest." Cancer Res **65**(20): 9236-9244.
Nagl, N. G., Jr., X. Wang, A. Patsialou, M. Van Scoy and E. Moran (2007). "Distinct mammalian SWI/SNF chromatin remodeling complexes with opposing roles in cell-cycle control." The EMBO journal **26**(3): 752-763.

Nagl, N. G., Jr., D. R. Zweitzig, B. Thimmapaya, G. R. Beck, Jr. and E. Moran (2006). "The c-myc gene is a direct target of mammalian SWI/SNF-related complexes during differentiation-associated cell cycle arrest." Cancer Res **66**(3): 1289-1293.

Niculescu, A. B., 3rd, X. Chen, M. Smeets, L. Hengst, C. Prives and S. I. Reed (1998). "Effects of p21(Cip1/Waf1) at both the G1/S and the G2/M cell cycle transitions: pRb is a critical determinant in blocking DNA replication and in preventing endoreduplication." Mol Cell Biol **18**(1): 629-643.

Ohkoshi, S., M. Yano and Y. Matsuda (2015). "Oncogenic role of p21 in hepatocarcinogenesis suggests a new treatment strategy." World journal of gastroenterology **21**(42): 12150-12156.

Osaki, M., M. Oshimura and H. Ito (2004). "PI3K-Akt pathway: Its functions and alterations in human cancer." Apoptosis **9**(6): 667-676.

Otto, G. P., B. Rathkolb, M. A. Oestereich, C. J. Lengger, C. Moerth, K. Micklich, H. Fuchs, V. Gailus-Durner, E. Wolf and M. Hrabě de Angelis (2016). "Clinical Chemistry Reference Intervals for C57BL/6J, C57BL/6N, and C3HeB/FeJ Mice (Mus musculus)." Journal of the American Association for Laboratory Animal Science : JAALAS **55**(4): 375-386.

Parisotto, M., E. Grelet, R. El Bizri, Y. Dai, J. Terzic, D. Eckert, L. Gargowitsch, J.-M. Bornert and D. Metzger (2018). "PTEN deletion in luminal cells of mature prostate induces replication stress and senescence in vivo." Journal of Experimental Medicine **215**(6): 1749-1763.

Parisotto, M. and D. Metzger (2013). "Genetically engineered mouse models of prostate cancer." Molecular oncology **7**(2): 190-205.

Patel, R., E. A. Brzezinska, P. Repiscak, I. Ahmad, E. Mui, M. Gao, A. Blomme, V. Harle, E. H. Tan, G. Malviya, A. Mrowinska, C. J. Loveridge, L. K. Rushworth, J. Edwards, C. Ntala, C. Nixon, A. Hedley, G. Mackay, S. Tardito, O. J. Sansom and H. Y. Leung (2020). "Activation of β -Catenin Cooperates with Loss of Pten to Drive AR-Independent Castration-Resistant Prostate Cancer." Cancer Research **80**(3): 576-590.

Patsialou, A., D. Wilsker and E. Moran (2005). "DNA-binding properties of ARID family proteins." Nucleic acids research **33**(1): 66-80.
Perlmutter, M. A. and H. Lepor (2007). "Androgen deprivation therapy in the treatment of advanced prostate cancer." Reviews in urology **9 Suppl 1**(Suppl 1): S3-S8.

Pospihalj, B. (2015). "Staging of prostate cancer: a review with reference for further refinement." *Anal Quant Cytopathol Histpathol* **37**(1): 69-74.

Raab, J. R., S. Resnick and T. Magnuson (2015). "Genome-Wide Transcriptional Regulation Mediated by Biochemically Distinct SWI/SNF Complexes." *PLoS Genet* **11**(12): e1005748.

Rameh, L. E., K. F. Tolia, B. C. Duckworth and L. C. Cantley (1997). "A new pathway for synthesis of phosphatidylinositol-4,5-bisphosphate." *Nature* **390**(6656): 192-196.
Ratnacaram, C. K., M. Teletin, M. Jiang, X. Meng, P. Chambon and D. Metzger (2008). "Temporally controlled ablation of PTEN in adult mouse prostate epithelium generates a model of invasive prostatic adenocarcinoma." *Proceedings of the National Academy of Sciences* **105**(7): 2521-2526.

Robbins, C. M., W. A. Tembe, A. Baker, S. Sinari, T. Y. Moses, S. Beckstrom-Sternberg, J. Beckstrom-Sternberg, M. Barrett, J. Long, A. Chinnaiyan, J. Lowey, E. Suh, J. V. Pearson, D. W. Craig, D. B. Agus, K. J. Pienta and J. D. Carpten (2011). "Copy number and targeted mutational analysis reveals novel somatic events in metastatic prostate tumors." *Genome Res* **21**(1): 47-55.

Roobol, M. J. (2013). "The prostate-specific antigen test." *Expert Opinion on Medical Diagnostics* **7**(5): 423-426.

Rordorf-Nikolic, T., D. J. Van Horn, D. Chen, M. F. White and J. M. Backer (1995). "Regulation of Phosphatidylinositol 3'-Kinase by Tyrosyl Phosphoproteins: FULL ACTIVATION REQUIRES OCCUPANCY OF BOTH SH2 DOMAINS IN THE 85-kDa REGULATORY SUBUNIT." *Journal of Biological Chemistry* **270**(8): 3662-3666.
Ruderman, N. B., R. Kapeller, M. F. White and L. C. Cantley (1990). "Activation of phosphatidylinositol 3-kinase by insulin." *Proc Natl Acad Sci U S A* **87**(4): 1411-1415.

Saal, L. H., P. Johansson, K. Holm, S. K. Gruvberger-Saal, Q. B. She, M. Maurer, S. Koujak, A. A. Ferrando, P. Malmström, L. Memeo, J. Isola, P. O. Bendahl, N. Rosen, H. Hibshoosh, M. Ringnér, A. Borg and R. Parsons (2007). "Poor prognosis in carcinoma is associated with a gene expression signature of aberrant PTEN tumor suppressor pathway activity." *Proc Natl Acad Sci U S A* **104**(18): 7564-7569.

Sandoval, G. J., J. L. Pulice, H. Pakula, M. Schenone, D. Y. Takeda, M. Pop, G. Boulay, K. E. Williamson, M. J. McBride, J. Pan, R. St Pierre, E. Hartman, L. A. Garraway, S. A. Carr, M. N. Rivera, Z. Li, L. Ronco, W. C. Hahn and C. Kadoch (2018). "Binding of TMPRSS2-ERG to BAF Chromatin Remodeling Complexes Mediates Prostate Oncogenesis." *Mol Cell* **71**(4): 554-566.e557.

Sandoval, G. J., J. L. Pulice, H. Pakula, M. Schenone, D. Y. Takeda, M. Pop, G. Boulay, K. E. Williamson, M. J. McBride, J. Pan, R. St Pierre, E. Hartman, L. A. Garraway, S. A. Carr, M. N. Rivera, Z. Li, L. Ronco, W. C. Hahn and C. Kadoch (2018). "Binding of TMPRSS2-ERG to BAF Chromatin Remodeling Complexes Mediates Prostate Oncogenesis." *Molecular cell* **71**(4): 554-566.e557.

Sega, G. A. (1984). "A review of the genetic effects of ethyl methanesulfonate." *Mutation Research/Reviews in Genetic Toxicology* **134**(2): 113-142.

Sen, M., X. Wang, F. H. Hamdan, J. Rapp, J. Eggert, R. L. Kosinsky, F. Wegwitz, A. P. Kutschat, F. S. Younesi, J. Gaedcke, M. Grade, E. Hessmann, A. Papantonis, P. Ströbel and S. A. Johnsen (2019). "ARID1A facilitates KRAS signaling-regulated enhancer activity in an AP1-dependent manner in colorectal cancer cells." *Clinical Epigenetics* **11**(1): 92.

Shen, J., Y. Peng, L. Wei, W. Zhang, L. Yang, L. Lan, P. Kapoor, Z. Ju, Q. Mo, I.-M. Shih, I. P. Uray, X. Wu, P. H. Brown, X. Shen, G. B. Mills and G. Peng (2015). "ARID1A Deficiency Impairs the DNA Damage Checkpoint and Sensitizes Cells to PARP Inhibitors." Cancer discovery **5**(7): 752-767.

Shen, J., Y. Peng, L. Wei, W. Zhang, L. Yang, L. Lan, P. Kapoor, Z. Ju, Q. Mo, M. Shih le, I. P. Uray, X. Wu, P. H. Brown, X. Shen, G. B. Mills and G. Peng (2015). "ARID1A Deficiency Impairs the DNA Damage Checkpoint and Sensitizes Cells to PARP Inhibitors." Cancer Discov **5**(7): 752-767.

Škara, L., A. Huđek Turković, I. Pezelj, A. Vrtarić, N. Sinčić, B. Krušlin and M. Ulamec (2021). "Prostate Cancer-Focus on Cholesterol." Cancers **13**(18): 4696.

Skinder, D., I. Zacharia, J. Studin and J. Covino (2016). "Benign prostatic hyperplasia: A clinical review." Jaapa **29**(8): 19-23.

Stelloo, S., E. Nevedomskaya, Y. Kim, L. Hoekman, O. B. Bleijerveld, T. Mirza, L. F. A. Wessels, W. M. van Weerden, A. F. M. Altelaar, A. M. Bergman and W. Zwart (2018). "Endogenous androgen receptor proteomic profiling reveals genomic subcomplex involved in prostate tumorigenesis." Oncogene **37**(3): 313-322.

Stocking, J. J., M. V. Fiandalo, E. A. Pop, J. H. Wilton, G. Azabdaftari and J. L. Mohler (2016). "Characterization of Prostate Cancer in a Functional Eunuch." Journal of the National Comprehensive Cancer Network : JNCCN **14**(9): 1054-1060.

Strohner, R., M. Wachsmuth, K. Dachauer, J. Mazurkiewicz, J. Hochstatter, K. Rippe and G. Längst (2005). "A 'loop recapture' mechanism for ACF-dependent nucleosome remodeling." Nat Struct Mol Biol **12**(8): 683-690.

Sun, A., O. Tawfik, B. Gayed, J. B. Thrasher, S. Hoestje, C. Li and B. Li (2007). "Aberrant expression of SWI/SNF catalytic subunits BRG1/BRM is associated with tumor development and increased invasiveness in prostate cancers." Prostate **67**(2): 203-213.

Suryo Rahmanto, Y., W. Shen, X. Shi, X. Chen, Y. Yu, Z.-C. Yu, T. Miyamoto, M.-H. Lee, V. Singh, R. Asaka, G. Shimberg, M. I. Vitolo, S. S. Martin, D. Wirtz, R. Drapkin, J. Xuan, T.-L. Wang and I.-M. Shih (2020). "Inactivation of Arid1a in the endometrium is associated with endometrioid tumorigenesis through transcriptional reprogramming." Nature Communications **11**(1): 2717.

Taitt, H. E. (2018). "Global Trends and Prostate Cancer: A Review of Incidence, Detection, and Mortality as Influenced by Race, Ethnicity, and Geographic Location." American journal of men's health **12**(6): 1807-1823.

Tokunaga, R., J. Xiu, R. M. Goldberg, P. A. Philip, A. Seeber, F. Battaglin, H. Arai, J. H. Lo, M. Naseem, A. Puccini, M. D. Berger, S. Soni, W. Zhang, S. Chen, J. J. Hwang, A. F. Shields, J. L. Marshall, H. Baba, W. M. Korn and H.-J. Lenz (2020). "The impact of ARID1A mutation on molecular characteristics in colorectal cancer." European Journal of Cancer **140**: 119-129.

Tsai, S., L.-A. Fournier, E. Y.-c. Chang, J. P. Wells, S. W. Minaker, Y. D. Zhu, A. Y.-H. Wang, Y. Wang, D. G. Huntsman and P. C. Stirling (2021). "ARID1A regulates R-loop associated DNA replication stress." PLOS Genetics **17**(4): e1009238.

U.K., C. R. (2020). "PC incidence over time ", from <https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/prostate-cancer/incidence#ref-2>.

- Van Rechem, C.**, G. Boulay and D. Leprince (2009). "HIC1 interacts with a specific subunit of SWI/SNF complexes, ARID1A/BAF250A." Biochem Biophys Res Commun **385**(4): 586-590.
- Vasileiou, G.**, A. B. Ekici, S. Uebe, C. Zweier, J. Hoyer, H. Engels, J. Behrens, A. Reis and M. V. Hadjihannas (2015). "Chromatin-Remodeling-Factor ARID1B Represses Wnt/ β -Catenin Signaling." American journal of human genetics **97**(3): 445-456.
- Velonas, V. M.**, H. H. Woo, C. G. d. Remedios and S. J. Assinder (2013). "Current Status of Biomarkers for Prostate Cancer." International Journal of Molecular Sciences **14**(6): 11034-11060.
- Wang, C.**, Y. Feng, C. Zhang, D. Cheng, R. Wu, Y. Yang, D. Sargsyan, D. Kumar and A.-N. Kong (2020). "PTEN deletion drives aberrations of DNA methylome and transcriptome in different stages of prostate cancer." The FASEB Journal **34**(1): 1304-1318.
- Wang, S.**, J. Gao, Q. Lei, N. Rozengurt, C. Pritchard, J. Jiao, G. V. Thomas, G. Li, P. Roy-Burman, P. S. Nelson, X. Liu and H. Wu (2003). "Prostate-specific deletion of the murine Pten tumor suppressor gene leads to metastatic prostate cancer." Cancer Cell **4**(3): 209-221.
- Wang, W.**, J. Côté, Y. Xue, S. Zhou, P. A. Khavari, S. R. Biggar, C. Muchardt, G. V. Kalpana, S. P. Goff, M. Yaniv, J. L. Workman and G. R. Crabtree (1996). "Purification and biochemical heterogeneity of the mammalian SWI-SNF complex." Embo j **15**(19): 5370-5382.
- Wang, W.**, Y. Xue, S. Zhou, A. Kuo, B. R. Cairns and G. R. Crabtree (1996). "Diversity and specialization of mammalian SWI/SNF complexes." Genes Dev **10**(17): 2117-2130.
- Wang, Z., K. Chen, Y. Jia, J.-C. Chuang, X. Sun, Y.-H. Lin, C. Celen, L. Li, F. Huang, X.
- Liu, D. H. Castrillon**, T. Wang and H. Zhu (2020). "Dual ARID1A/ARID1B loss leads to rapid carcinogenesis and disruptive redistribution of BAF complexes." Nature cancer **1**(9): 909-922.
- Wiegand, K. C.**, B. T. Hennessy, S. Leung, Y. Wang, Z. Ju, M. McGahren, S. E. Kalloger, S. Finlayson, K. Stemke-Hale, Y. Lu, F. Zhang, M. S. Anglesio, B. Gilks, G. B. Mills, D. G. Huntsman and M. S. Carey (2014). "A functional proteogenomic analysis of endometrioid and clear cell carcinomas using reverse phase protein array and mutation analysis: protein expression is histotype-specific and loss of ARID1A/BAF250a is associated with AKT phosphorylation." BMC Cancer **14**: 120.
- Williamson, C. T.**, R. Miller, H. N. Pemberton, S. E. Jones, J. Campbell, A. Konde, N. Badham, R. Rafiq, R. Brough, A. Gulati, C. J. Ryan, J. Francis, P. B. Vermulen, A. R. Reynolds, P. M. Reaper, J. R. Pollard, A. Ashworth and C. J. Lord (2016). "ATR inhibitors as a synthetic lethal therapy for tumours deficient in ARID1A." Nat Commun **7**: 13837.
- Wu, C.**, J. Lyu, E. J. Yang, Y. Liu, B. Zhang and J. S. Shim (2018). "Targeting AURKA-CDC25C axis to induce synthetic lethality in ARID1A-deficient colorectal cancer cells." Nature Communications **9**(1): 3212.
- Wu, J. N.** and C. W. Roberts (2013). "ARID1A mutations in cancer: another epigenetic tumor suppressor?" Cancer Discov **3**(1): 35-43.
- Wu, X.**, J. Wu, J. Huang, W. C. Powell, J. Zhang, R. J. Matusik, F. O. Sangiorgi, R. E. Maxson, H. M. Sucov and P. Roy-Burman (2001). "Generation of a prostate epithelial cell-specific Cre transgenic mouse model for tissue-specific gene ablation." Mech Dev **101**(1-2): 61-69.

Wyld, L., I. Bellantuono, T. Tchkonja, J. Morgan, O. Turner, F. Foss, J. George, S. Danson and J. L. Kirkland (2020). "Senescence and Cancer: A Review of Clinical Implications of Senescence and Senotherapies." Cancers **12**(8): 2134.

Xie, Q., Y. Liu, T. Cai, C. Horton, J. Stefanson and Z. A. Wang (2017). "Dissecting cell-type-specific roles of androgen receptor in prostate homeostasis and regeneration through lineage tracing." Nature communications **8**: 14284-14284.

Xu, G., S. Chhangawala, E. Cocco, P. Razavi, Y. Cai, J. E. Otto, L. Ferrando, P. Selenica, E. Ladewig, C. Chan, A. Da Cruz Paula, M. Witkin, Y. Cheng, J. Park, C. Serna-Tamayo, H. Zhao, F. Wu, M. Sallaku, X. Qu, A. Zhao, C. K. Collings, A. R. D'Avino, K. Jhaveri, R. Koche, R. L. Levine, J. S. Reis-Filho, C. Kadoch, M. Scaltriti, C. S. Leslie, J. Baselga and E. Toska (2020). "ARID1A determines luminal identity and therapeutic response in estrogen-receptor-positive breast cancer." Nature Genetics **52**(2): 198-207.

Yamada, Y. and H. Beltran (2021). "Clinical and Biological Features of Neuroendocrine Prostate Cancer." Curr Oncol Rep **23**(2): 15.

Yan, H.-B., X.-F. Wang, Q. Zhang, Z.-Q. Tang, Y.-H. Jiang, H.-Z. Fan, Y.-h. Sun, P.-Y. Yang and F. Liu (2013). "Reduced expression of the chromatin remodeling gene ARID1A enhances gastric cancer cell migration and invasion via downregulation of E-cadherin transcription." Carcinogenesis **35**(4): 867-876.

Yang, J. and Y. Yin (2020). "PTEN in Chromatin Remodeling." Cold Spring Harb Perspect Med **10**(2).

Ye, J., Y. Zhou, M. R. Weiser, M. Gonen, L. Zhang, T. Samdani, R. Bacares, D. DeLair, S. Ivelja, E. Vakiani, D. S. Klimstra, R. A. Soslow and J. Shia (2014). "Immunohistochemical detection of ARID1A in colorectal carcinoma: loss of staining is associated with sporadic microsatellite unstable tumors with medullary histology and high TNM stage." Hum Pathol **45**(12): 2430-2436.

Zhai, Y., R. Kuick, C. Tipton, R. Wu, M. Sessine, Z. Wang, S. J. Baker, E. R. Fearon and K. R. Cho (2016). "Arid1a inactivation in an Apc- and Pten-defective mouse ovarian cancer model enhances epithelial differentiation and prolongs survival." The Journal of pathology **238**(1): 21-30.