TARGETING ANDROGEN RECEPTOR AND PIP5K1A IN PROSTATE CANCER CELLS USING CRISPR-CAS9 GENOME EDITING TOOL

MIGUEL HERNANDEZ-QUILES

THESIS SUBMITTED TO THE UNIVERSITY OF NOTTINGHAM

FOR THE DEGREE OF MASTER OF RESEARCH

September 2016

ACKNOWLEDGEMENTS

First I would like to thanks to my supervisor Professor David M. Heery for his support and advice during the last months. Also Dr. Hilary Collins for her enormous patience. Thanks to Jonathan Whitchurch, Bismoy Mazumder and Cintia Monteiro, you taught me, helped me and all of you made every day better. Also thanks to all GRRB group. Specially thanks to my partner in this adventure Maria Pilar Clemente. And last but not least to my parents and sister, because all I have done is to make them proud.

ABSTRACT

According to the World Cancer Research statistics, Prostate Cancer (PCa) accounts for 15% of all new cancer cases in men every year. More than 30% of treated patients will suffer disease recurrence and will develop Castration Resistant Prostate Cancer (CRPC), which no longer responds to androgen deprivation therapies. The mechanisms underlying the development of CRPC remain unclear. Androgen Receptor (AR) is highly expressed in prostate cancer cells, and plays an essential role in growth, survival and differentiation of prostate tumours. The aberrant activation of AR and AR-variants lacking the ligand binding domain (LBD), which is the therapeutic target of most PCa treatments, are characteristics of CRPC. In addition, recent studies suggest that PIP5K1A, a lipid kinase responsible for the synthesis of PIP2, might play an important role in PCa tumourogenesis and AR activation. In this study we used CRISPR-Cas9 genome editing technology to target AR and PIP5K1A in LNCaP C4-2 prostate cancer cells. Following cell sorting of GFP+ cells and clonal expansion, we successfully generated 9 mutant cell lines for PIP5K1A and 3 for AR. PCR Genotyping and DNA sequencing confirmed the presence of indels in targeted alleles in all of these clones, and Western Blots indicated that disruption of the targeted gene expression was achieved in several clones. Preliminary data showed AR depletion also reduces PIP5K1A expression, suggesting a regulatory pathway involving these genes. The AR and PIP5K1A mutant clones will be useful cell models to explore mechanisms involved in CRPC, and thus may aid the search of new therapies.

List of Figures

List of abbreviations

Ab	Antibody
APS	Ammonium Persulphate solution
AR	Androgen Receptor
bp	Base Pair
BSA	Bovine Serum Albumin
$CaCl_2$	Calcium chloride
Cas	CRISPR-associated
cDNA	Complementary DNA
CO_2	Carbon Dioxide
CRISPR	Clustered Regularly Interspaced Palindromic Repeats
crRNA	CRISPR RNA sequence
CRPC	Castration Resistant Prostate Cancer
DBD	DNA Binding Domain
DiH_2O	Double distilled water
DMEM	Dulbecco's Modified Eagle Medium
DMSO	Dimethyl Sulphoxide
dNTP	Deoxyribonucleotide triphosphate
DSB	Double Strand Break
ECL	Enhanced Chemiluminiscence
EDTA	Ethylenediaminetetraacetic Acid
FACS	Fluorescence Activated Cell Sorting
FBS	Foetal Bovine Serum
GFP	Green Fluorescent Protein
HBS	HEPES Buffered Saline solution

HDR	Homology Directed Repair
HEK-293	Human Embryonic Kidney 293 cell line
KDa	Kilo Dalton
LB	Luria-Bertani medium
LBD	Ligand Binding Domain
M	Molar
mg	Milligram
$MgCl_2$	Magnesium chloride
MIT	Massachusetts Institute of Technology
mL	Millilitre
mM	Millimolar
NaCl	Sodium chloride
NaOH	Sodium hydroxide
NCBI	National Centre for Biotechnology Information
NHEJ	Non-Homologous End Joining
NR	Nuclear Receptor
ºC	Degree Centigrade
PAGE	Poly-Acrylamide Gel Electrophoresis
PAM	Protospacer Adjacent Motif
PBS-T	Phosphate Buffer Saline – Tween
PBS	Phosphate Buffer Saline
PCR	Polymerase Chain Reaction
Pfu	*Pyrococcus furiosus*
PIP5K1A	Phosphatidylinositol-4 phosphate 5-kinase type 1 Alpha
PMSF	Phenylmethylsulfonyl Fluoride

qPCR	Quantitative Polymerase Chain Reaction
Rpm	Rotation per minute
RPMI	Roswell Park Memorial Institute-1640 Medium
SDS	Sodium Dodecyl Sulphate
sgRNA	Synthetic Guide RNA
SNP	Single Nucleotide Polymorphism
SSB	Single Strand Break
TALEN	Transcription Activator-Like Effector Nuclease
Taq	*Thermus aquaticus*
TBE	Tris-Borate-EDTA
TEMED	Tetra Methyl Ethylene Diamine
TGS	Tris-Glycine-SDS buffer
tracrRNA	Trans-activating crRNA
UV	Ultraviolet
V	Volts
μL	Microlitre
μM	Micromolar

CHAPTER 1

INTRODUCTION

1. INTRODUCTION

1.1. Prostate Cancer

Prostate cancer is the most common non-cutaneous cancer malignancy in men, and major cause of cancer deaths in most developed countries (is the cause of 1-2% of deaths in men) (1). Men have a small gland, part of the male reproductive system called the prostate gland; it is an exocrine gland located below the bladder. The prostate surrounds the urethra, which carries urine and sex fluids. The prostate gland is divided into 2 lobes, to the left and the right of a central groove. The prostate gland produces compounds that are important for the normal density of the semen. Testosterone regulates the normal growth and function of the prostate. Testosterone is a male sex hormone and is produced in the testes.

The early stage of the prostate cancer is normally asymptomatic therefore this disease often implies local extension or metastasis (2). Thus, it is critical that detection and diagnosis of the disease is made as early as possible. To detect prostate cancer the most common techniques are digital rectal exam, the transrectal ultrasound and the detection of biomarkers in blood such as PSA (prostate specific antigen) (3).

Prostate cancer can spread through three different ways; Tissue invasion (tumour grows into the nearby areas), lymphatic system (cancer cells enter the lymph vessels to spread to other parts of the body), and blood (cancer cells spread through vascular system) (4).

Among all different types of prostate cancer cases approximately 98% are glandular in origin (5). The prostate cancer is multifocal, and has an average of at least two geographically distinct foci, and different histological pattern. The Gleason score is the most commonly used grading protocol test to classify this type of cancer. This involves scoring multiple needle biopsies which are examined for their degree of abnormality/ cancer grade. The peripheral portion of the gland is where most of the malignancies start, and it is away from the urethra. When the cancer begins to affect the urethra and the bladder is when the first symptoms appear, and usually it has been spread to others parts of the body through the blood or the lymph system. Another cause of delay in the diagnosis is the wrong association with the benign prostatic hyperplasemia (BPH), due the similarity in the symptoms.

1.2. Types of Prostate Cancer

There are several types of prostate cancer according to the Cancer Research Society. The precursor to prostate cancer is known as prostatic intraepithelial neoplasia, and it can be also found in different parts within the prostate.

- Benign prostatic hyperplasia (BPH): This is a benign type of tumour of the prostate. The prostate grows and can affect the urethra, by preventing the normal flow of urine.

- Prostatic adenocarcinoma: This type of tumour is the most frequent form of prostate cancer and accounts for 90 to 95% of prostate cancer. Many of these cancers grow extremely slowly and are not likely to

spread. However some of the cancer cells may spread from the prostate to other parts of the body, particularly the bones and lymph nodes. Prostatic adenocarcinoma affects the cells lining the glandular organ of the prostate; this area is responsible for the production of compounds presents on the seminal fluid.

- Small cell carcinoma: This type of tumour is neuroendocrine in origin; it is a very aggressive form of prostate cancer that does not lead to a change in prostate specific antigen (PSA). It is more difficult to detect than adenocarcinoma.

- Squamous cell carcinoma: Is a non-glandular type of prostate cancer that starts from the cells that cover the prostate gland and does not lead to an increase in prostate specific antigens (PSA), which makes it more difficult to detect. Squamous cell carcinoma is a very aggressive form of cancer.

- Prostatic sarcomas: This is a rare type of prostate cancer and account for less than 0.1% of prostate cancer. Prostatic sarcoma occurs in young men (35-50 years old). The tumour starts from muscle cells and are capable of developing into connective tissues, lymphatic vessels and blood vessels.

- Transitional cell carcinomas: This type of tumour derives from primary tumours in the bladder or urethra, and rarely affects the prostate.

Some of the risk factors associated with prostate cancer are age and race. For men younger than 40 the chances of having prostate cancer are rare, but the probabilities rise rapidly for men older than 50.

PCa occurs more in African American men, who are more than twice as likely to die of Prostate cancer as white men. Also Asian and Hispanic are less likely to have the disease than the non-hispanic white men (data from American Cancer Society). The reasons behind this data remain unclear.

Aside from age and race, the only established risk factor for prostate cancer is a family history of the disease. The risk for men with a first degree relative with prostate cancer is about twice that for men in general population, this familial risk is more than four times higher than that for the general population if the first degree relative that is diagnosed with prostate cancer is younger than 60 years old. The risk is 50% higher in monozygotic twins than in dizygotic twin. All this evidence supports the genetic factor as an important determinant of the variation in risk at the population level (1). Genetic predisposition for prostate cancer is related with rare highly penetrant mutations, genetic variants conferring lower risk, or a combination of these two. Only the homeo box gene HOXB13 has been identified as a prostate cancer predisposition gene.

1.3. Androgens and Androgen Receptor

Androgens exerts their biological effects through the Androgen Receptor, and both are required for prostate development and normal prostate function (6). Androgens were first discovered in 1936. The major circulating androgen is testosterone, the main function is inducing virilisation of internal sex organs and also plays and important role in

the hypothalamus-pituitary-testis loop to maintain proper amounts of sex steroids in the body.

Testosterone production (figure 1) is regulated by luteinizing hormone (LH) and Luteinizing hormone-releasing hormone (LHRH). The hypothalamus releases LHRH, which stimulates the pituitary gland to secrete LH. The testosterone is produced by the testes through the action of LH on specific testes cells. Testosterone and other androgens (normally produced by the adrenal gland) are taken up by the prostate cells, where they bind to the AR or are converted to DHT (dihydroxy testosterone), which has a greater binding affinity for the androgen receptor than testosterone (7).

DHT and androstenedione are important androgens required for male development. DHT is produced by the enzyme 5' reductase, which transforms Testosterone in DHT inside the cell (8), DHT promotes virilization of the external genitalia and the development of the secondary sexual characteristics. Although androgens are considered male sex hormones, males and females have them to different degrees. In 1960 the hypothesis that a receptor might be needed to mediate the biological effects of androgens was developed. And finally, the full-length human AR cDNA was isolated by Chang *et al.* 1989 and Lubahn *et al.* 1988 based on its sequence homology to other steroid receptors by using DNA oligonucleotide probe hybridisation.

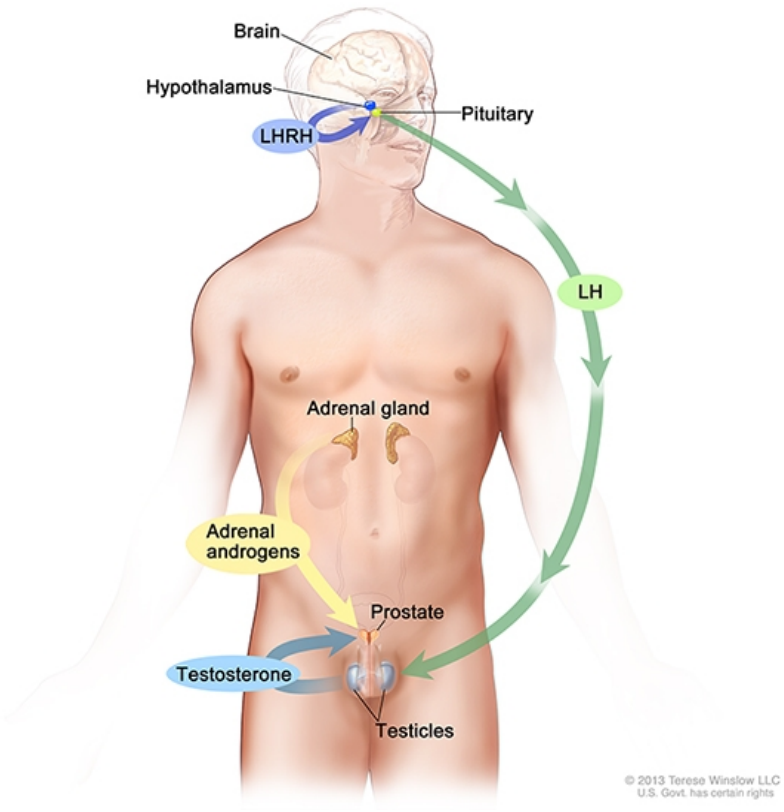

Figure 1

Hypothalamus-Pituitary-Testes loop controls the amount of androgens in the body. Figure from the National Cancer Institute.

1.3.1. AR as a member of nuclear receptor class 1 family

The androgen receptor (AR or NR3C4, nuclear receptor subfamily 3, group C, gene 4) (9) is a member of the steroid hormone group of nuclear receptors. Others nuclear receptors that belong to this family are; estrogen receptor (ER), glucocorticoid receptor (GR), progesterone receptor (PR) and mineralocorticoid receptor (MR) (10)(11). The nuclear receptor family describes a related but diverse group of transcriptional factors. Proteins of nuclear receptor superfamily are single polypeptide chains with three major domains. The DNA-binding

domain (DBD), this domain binds in highly specific manner to DNA sequences called response elements (RE) (12) and it is highly conserved. The ligand binding domain (LBD), the less conserved part between the nuclear receptors, recognizes specific hormonal and non-hormonal ligands, and it is directly related with their biological response. The LBD is contained in the C-terminal half of the receptor. Finally the variable N-terminal domain (NTD), which serves as modulatory domain (13).

The nuclear receptor family can be classified in three classes:

- Class 1: Steroid receptor family and includes the progesterone receptor (PR), the estrogen receptor (ER), the glucocorticoid receptor (GR), the androgen receptor (AR) and the mineralocorticoid receptor (MR).
- Class 2: the thyroid/retinoid family includes the thyroid receptors (TRs), vitamin D receptor (VDR), the retinoic acid receptors (RARs) and the peroxisome proliferator-activated receptors (PPARs).
- Class 3: This class of nuclear receptor is known as the orphan receptor family. It is a group of proteins that presents the nuclear receptor structure but their ligands have not yet been identified.

1.3.2. AR: Gene and protein structure

AR gene is more than 90 kb long and it is located on the chromosome X (Figure 2), the gene codes for a protein of 919 amino acids. The AR gene has 8 exons, including a large exon 1 (1586 bp) encoding for the

N-terminal domain. Exons 2 (152 pb) and exon 3 (117 bp) encode the DNA-binding domain (14). And the ligand binding domain (LBD) is encoded by the remaining 5 exons. There is also a small Hinge region between the LBD and the DBD that is required for a normal 3D conformational structure.

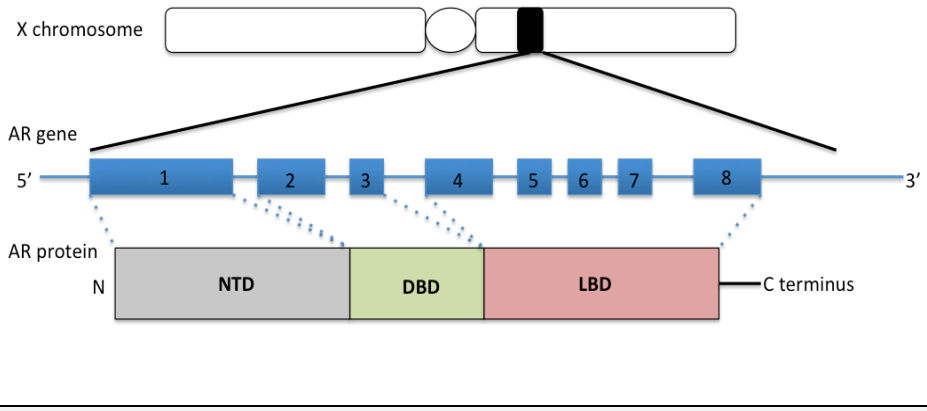

Figure 2

AR gene and protein structure.

All steroid receptors have a similar organization at protein level. The N-terminal domain is the least conserved between the different steroid receptors (less than 25% identity), the highest identity (59-82%) corresponds to the DBD, to their common need to bind the DNA. The LBD shows sequence identity ranging from 22% to 55% (15), due to the receptor specificity for different hormones.

1.3.3. AR function

AR is a transcription factor that controls the expression of specific genes, the AR is ligand-dependent, and in the absence of its cognate

ligands DHT and Testosterone, the AR remains inactive in the cytoplasm. Unbound AR in associated with a complex of heat shock proteins (HSPs). However upon binding to a ligand AR is activated, after the interaction between the LBD and the ligand AR goes through a series of conformational changes that result in the dissociation of the HSPs, dimerization, phosphorylation and finally is translocated to the nucleus (16).

Once in the nucleus translocated AR binds to androgen response elements (ARE), which is characterized by six nucleotide half-site consensus sequence 5”-TGTTCT-3” followed by three random nucleotides and it is located in the promoter or enhancer region of AR gene targets. Recruitment of the others transcriptional co-regulators and transcriptional machinery initiates the gene expression. To summarize, the ligand-induced conformational changes in the LBD between the androgen and the AR initiates male sexual development and differentiation. A model of action of androgens and AR is shown in Figure 3.

The transcriptional activation of AR is regulated by the interaction of AR with coregulators and by phosphorylation in response to growth factors (6)(17)(18). AR and the modulators of AR remain important in prostate cancer, the majority of treatments for prostate cancer is directed towards the reduction of serum androgens and the inhibition of AR (19). However, experimental data suggest that prostate cancer progression occurs by dysregulation of AR activity through signal transduction cascade, an alteration in the expression of AR co-

regulators and mutations in AR LBD that produce an AR capable to be activated in the absence of androgens, or even in response to antiandrogen treatments (antagonist ligands) or other mechanisms (7). This is why the androgen ablation therapy fails in the majority of patients within 2 years, and prostate cancer progresses to a hormone refractory state.

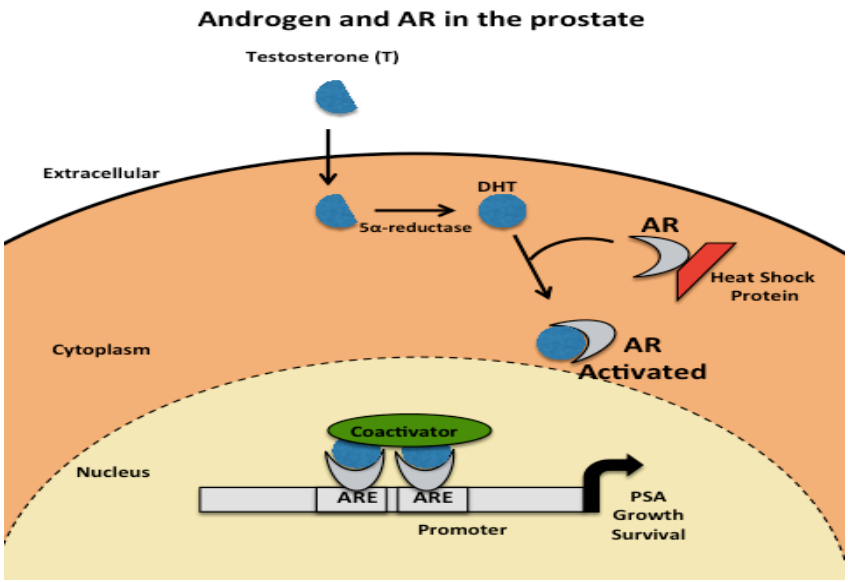

Figure 3

Androgens activate AR and both are translocated to the nucleus. Adapted from Tan *et al.* 2014.

1.4. Prostate Cancer therapies

Men with localised disease may have different prognoses and different treatment options. The treatment chosen is based on the risk assessments, which normally is a combination of patient age, clinical tumour stage, serum PSA, PSA density, Gleason score, number of positive prostate biopsies and amount of malignant tissue. Radical

prostatectomy is a common treatment in prostate cancer, in this operation the prostate gland is removed and some of the tissue that surrounds the prostate, including seminal vesicles is also removed.

Another widely used method is Radiation; radiation therapy uses high-energy rays to destroy cancer cells. Radiation can be used as a first treatment options for tumuors that has not spread outside of the prostate or has grown only to the nearby tissue. Surgery and radiation are used in low-grade (Gleason score of 6 or less) and localised prostate cancers

Hormone therapy, also called androgen deprivation therapy (ADT) or androgen suppression therapy is another available treatment. Hormone therapy is directed towards the reduction of androgens in a man's body. Reducing androgens can slow the growth of the cancer and even shrink the tumor. Hormone therapy can be used in combination with radiation or surgery when there is a risk of cancer returning. ADT is also used when the cancer has spread to other tissues (metastatic disease) and cannot be removed by surgery. These therapies include:

- LHRH agonists. These drugs stop the body from making testosterone by blocking the LHRH receptor in the pituitary gland. Some examples are; goserelin (Zoladex®), histrelin (Vantas®), leuprolide (Lupron®), and triptorelin (Trelstar®).

- LHRH antagonists. These drugs prevent the secretion of LH by blocking the LHRH from binding to its recceptor. One example is degarelix (Firmagon®).

- Androgen inhibitors. These drugs target and block enzymes that the body use to produce testosterone. Some examples are enzalutamide (Xtandi), ketoconazole, and abiraterone (Zytiga®), which is given along with prednisone.

- Antiandrogens. These medicines normally are used along with LHRH agonists. Antiandrogens help block the body's supply of testosterone. Antiandrogens can be divided into steroidal antiandrogens and "pure" antiandrogens. One steroidal antiandrogens is megestrol (Megace®). An example of "pure" antiandrogens is bicalutamide (Casodex®), flutamide, and nilutamide (Nilandron®).

- Orchiectomy. This surgery is considered to be hormone therapy because the testicles are responsible for the production of the 90% of testosterone, and remove them decreases the amount of testosterone in the body.

Hormone therapy is based on the fact that prostate cancer cannot grow or survive without androgens, but the real fact is that one-third of the cases treated will experience disease recurrence and will progress into castration-resistant PCa (CRPC), and there is a lack of treatment options for this cases because they are no longer responsive to anti androgens therapy (20).

Although there are no standard treatments for CRPC, considerable progress has been made in understanding this disease. One mechanism of resistance to antiandrogen treatment is due to the increased expression of AR variants lacking the ligand binding-domain, the best characterized of which is AR-V7. AR-V7 is a constitutively

active variant lacking the LBD encoded by exon 4-8, owing to utilization of a cryptic exon7. AR-V7 expression is significantly higher in CRPC specimens than in specimens from hormone-naïve PCa and is associated with worse clinical outcome, suggesting an important role of AR-V7 in development of CRPC (21).

1.5. PIP5Ks

In addition to androgen signalling, current research is focussed on discovering other signalling pathways that may be important in prostate cancer, and in particular CRPC. Recent studies have identified a convergence of phosphoinositide signalling and AR function (22)(23). Semenas *et al*. 2014 showed for the first time that high levels of Phosphatidylinositol-4 phosphate 5-kinase type 1 Alpha were associated with worse prognosis and overexpression of AR in prostate cancer. The phosphoinositide family of lipids consists of seven derivatives of phosphatidylinositol (PtdIns) (Figure 4A) that are formed through the phosphorylation of 3-, 4-, 5-position of the inositol ring (24). All seven combinations of phosphorylated PtdIns have been found in animals, all except PtdIns (3,4,5) P3 has been found in plants. The importance of this family as lipid signalling molecules was first discovered by Lowell and Hokin in 1953 (25). PtdIns constitute less than 10% of all lipids in the eukaryotic cell membrane and phosphorylated derivatives (PIs) are important regulators of a large variety of cellular processes. All PIs have a glycerol backbone esterified to two fatty acids chains and a phosphate group, and like others

phospholipids they are attached to a polar head group that extends into the cytoplasm.

Among PI, phosphatidylinositol (4,5) biphosphate (PIP2) is the most studied. PIP2 is not only the substrate for phosphatidylinositol 3-kinase (PI3K), which produces phosphatidylinositol (3,4,5) triphosphate, it is also a second messenger itself. As a second messenger PIP2 affects cell migration, cell-cell adhesion, cytokinesis, stress response, apoptosis and nuclear processes such as cell-cycle progression and splicing.

The enzymes that are responsible for the synthesis of PIP2 are two distinct but related phosphoinositide kinases (PIPKs). Phosphatidylinositol 5-phosphate 4-kinase (PIP4K) that phosphorylates phosphatidylinositol 5-phosphate at the 4-position, and phosphatidylinositol 4-phosphate 5-kinase (PIP5K) which phosphorylates phosphatidylinositol 4-phosphate at the position 5. Because the amount of phosphatidylinositol 4-phosphate in the cell is approximately ten times higher than phosphatidylinositol 5-phosphate, the major pathway of production of PIP2 is through PIP5k.

PIPKs can be divided into three main types (type I, II and III), the three types share significant sequence homology but differ in the substrate specificities, subcellular localisations and functions (20).

- Type I: are PIPKs that phosphorylate phosphatidylinositol 4-Phosphate to phosphatidylinositol $(4,5)P_2$ and are called phosphatidylinositol 4-phosphate 5-kinases (PIP5Ks).

- Type II: PIPKs that phosphorylates phosphatidylinositol 5-phosphate at the 4-position and are called Phosphatidylinositol 5-phosphate 4-kinases (PIP4Ks).
- Type III: Phosphatidylinositol 3-Phosphate 5-kinases that phosphorylates phosphatidylinositol 5-phosphate at position 3 to phosphatidylinositol (3,5)-P_2.

PIPKs type I or PIP5Ks can phosphorylate phosphatidylinositol 4-Phosphate as mentioned above can phosphorylate PtdIns, phosphatidylinositol 3-phosphate and phosphatidylinositol (3,4) di-phosphate at position 5 (figure 4B) (26). Mammals encode three PIP5K genes PIP5K1A (PIP5Kα) (chromosome1), PIP5K1B (PIP5Kβ) (chromosome 9) and PIP5K1C (PIP5Kγ) (chromosome 19) (27)(28). The PIP5K1A and PIP5K1B genes give rise to additional splice variants (3 and 4, respectively), whereas a single variant is known for PIP5K1C. The PIP5K proteins show strong conservation within the central regions encoding the catalytic domain. The so-called activation loop is responsible for the substrate specificity of types I and II. The activation loop is a fragment of 22 to 27 aa, which is adjacent to the C-terminal end of the catalytic domain in all PIPKs. When the activation loops of a type I and a type II PIPK were swapped, the chimera with type I backbone showed specificity for type II substrate, and vice versa for the other chimera (29). PIP5K1A proteins have been localised to the plasma membrane, the Golgi complex and also in the nucleus. PIP5K1B proteins are also localised in the plasma membrane as well in

vesicles in the perinuclear region of the cell. Less is known about PIP5K1C.

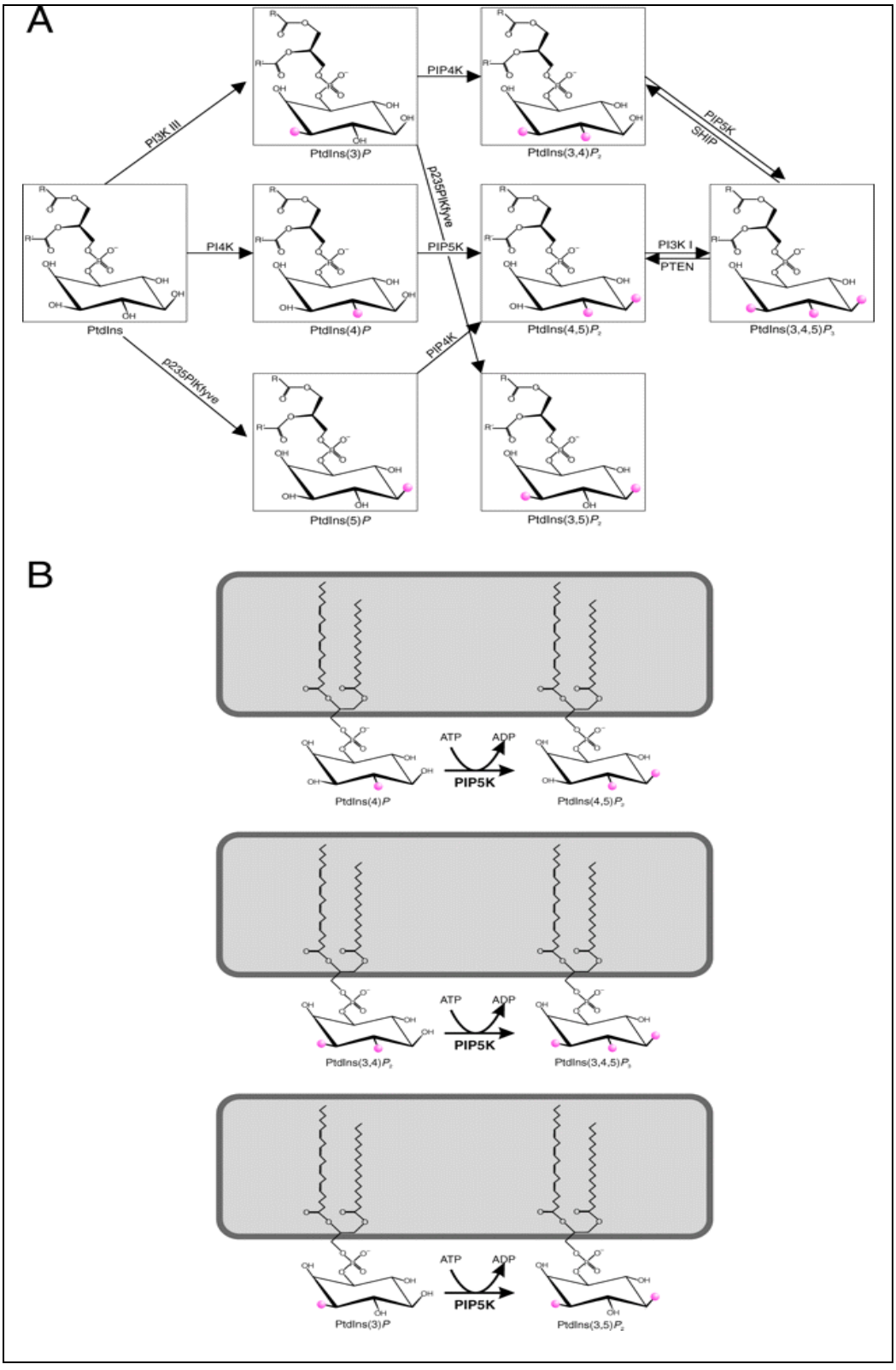

Figure 4

The generation of phosphoinositides by PIP5K. A) Diagram of the seven members of the phosphoinositide family. B) PIP5K can generate three different species of phosphoinositide using different substrates. Reproduced from Iman van den Bout *et al*. 2009.

1.5.1.PIP5K1A

PIP5K1A gene is located in chromosome 1 (1q21.3) (30). There are 5 isoforms, the isoform 1 encodes for the largest protein (562 aa) and has 16 exons. Isoform 2 encodes for a protein of 549 aa, this isoform lacks 3 nucleotides at a alternative splice site and one in-frame exon (exon 3). Isoform 3 encodes for a protein of 522 aa and lacks two in-frame coding exons (exon 3 and 10). The isoform 4 encodes for the smallest protein (500 aa) and lacks 3 nucleotides at an alternative splice variant and two in-frame exons (exon 3 and 12). Finally the isoform 5 encodes for a protein of 550 aa, and lacks an alternative in-frame exon (NCBI).

PIP5K1A is the major lipid kinase that is responsible for the synthesis of PtdIns-4,5-P2 (PIP2) by phosphorylating ptdIns-4-P (PI4P). PIP2 is a substrate used by PI3K to produce PtdIns-3,4,5-P3 (PIP3) (31), which activates the AKT family of serine/theonine kinases by recruiting them to the cell membrane (32)(33). This intracellular signalling pathway regulates the cell cycle. The aberrant activation of this pathway is one of the most common alterations in human cancer cells. In prostate cancer, it has been shown that the PI3K/AKT pathway can activate androgen receptor-mediated signalling, which promotes the progression of castration-resistant prostate cancer (34)(35).

1.6. CRISPR-Cas9

Recent advances in genome editing technology have derived from natural defence mechanisms present in bacterial cells such as CRISPR-CAS systems. Microbes have evolved mechanisms to deal with foreign nucleic acids that they are exposed to due to transduction, conjugation and transformation and viral predation. The clustered regulatory interspaced short palindromic repeats (CRISPR) system provides immunity against viruses and plasmids by targeting their DNA in a sequence-specific manner.

The CRISPR locus was discovered in *E. coli* in the 1980s, and is known to be present in around 84% of archaea and 45% of bacteria (CRISPRdb). It wasn't until 2007 that the function was confirmed by Barrangou and colleagues, who demonstrated that *S.thermophilus* can acquire resistance against a bacteriophage by integrating a DNA fragment of a virus into its CRISPR locus (36). CRISPR loci consititute an array of short repeated sequences separated by spacers (37). These spacers are often from foreign genomes and new spacers are added at one side of the sequence, making it possible to chronologically track the different viruses a cell and its ancestors have encountered. CRISPR can be found in both chromosomal and plasmid DNA, however the CRISPR activity requires additional proteins called Cas (CRISPR associated) proteins.

There are threed ifferent CRISPR-Cas systems (types I, II and III) and can be subdivided into 11 subtypes (I-A to IF, II-A to II-C, and III-A to III-B) (36)(38). They share a common function in providing

adaptive immunity, however the system exhibits extraordinary mechanism diversity. While type I and III systems occurs in bacteria and archea type II has only be found in bacteria.

The CRISPR-Cas mediated defence process can be divided into three stages (Figure 5). The first step, adaptation, it is the insertion of new spacers into the CRISPR locus. In the second step, expression, the Cas genes and the CRISPR locus are transcribed into a long-RNA precursor (pre-crRNA). This RNA is processed into the mature form of crRNA by Cas proteins and accessory factors. And finally in the third step, interference, the foreign DNA is recognized and degraded by CRISPR RNA and Cas proteins.

There are many different Cas proteins involved in cleaving the target DNA, processing the crRNA and enhance spacer acquisition in some cases. The Cas proteins are mainly responsible for the cleavage, producing a double strand break (DSB) in the DNA. This DSB can then be repaired by the cell through different pathways. The most common is the non-homologous end joining (NHEJ) (39). The NHEJ results in random insertions and deletions causing the disruption of the gene because of a frameshift in the sequence.

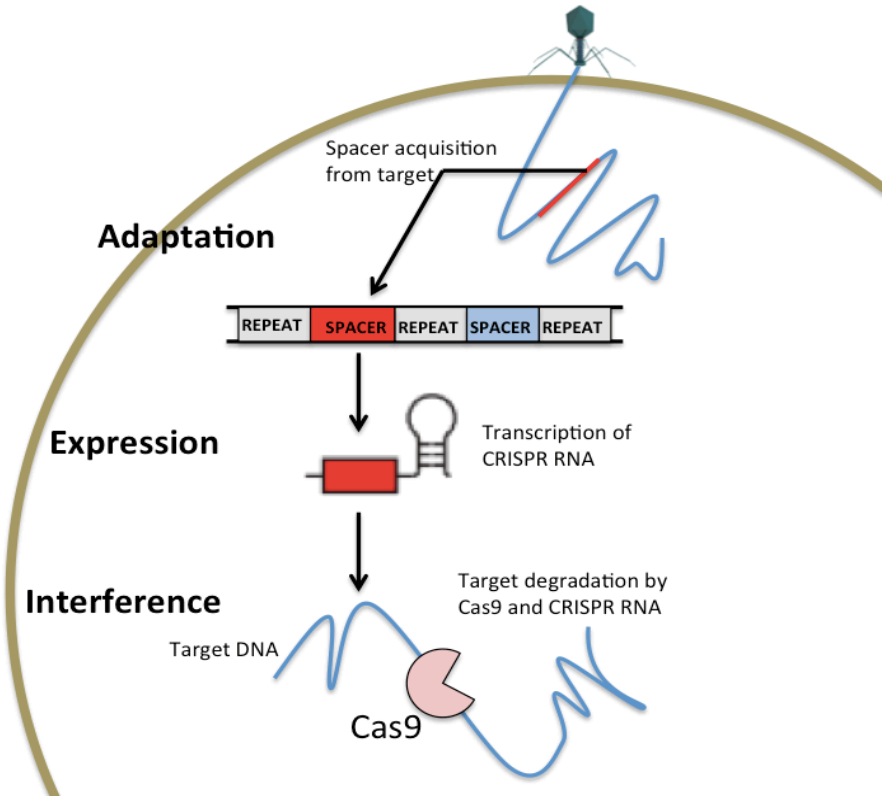

Figure 5

The CRISPR-Cas immune system process. Adapted from Rath *et al.* 2015.

1.6.1. CRISPR-Cas9 Genome Editing Technology

Doudna, Charpentier and colleagues (2012), showed that a modified streptococcal Cas9 (CRISPR-Cas system type II) can be programmed with a simple chimeric RNA (40). Cas9 is a DNA endonuclease with two different domains, RuvC and HNH, and each one of them cleaves the DNA in one strand (41). To create the CRISPR targeting RNA (crRNA), the CRISPR array is transcribed and the RNA is processed. In the type II system this process starts with the precursor of the CRISPR targeting RNA (pre-crRNA) synthesis, this transcript is then processed in presence of the trans-activating crRNA (tracrRNA) that has sequence

complementary to the palindromic repeat, and requires the presence of RNA-specific ribonuclease (RNase III). The tracrRNA hybridizes to the short palindromic repeat, and the crRNA and the tracrRNA can now bind to the Cas9, which becomes activated. The Cas9 protein will bind to a DNA sequence complementary to the crRNA and will produce the cleavage (40)(42).

There are some limitations in the system. First, any sequence potentially targeted by the system must have a specific sequence on its 3' end, the protospacer adjacent motif (PAM). The PAM is present on the DNA target but not on the crRNA that will target it. In the CRISPR Cas9 system the PAM sequence is NGG. Additionally, certain target sequences are believed to be problematic due to the RNA secondary structure they form. Another problematic issue is the off-target activity of the Cas9, the sgRNA allow some mismatches, and can direct the Cas9 to a non-desired sequence, producing indel mutation. To overcome this problem, the double nicking strategy uses a mutant Cas9, called Cas9n, this Cas9 has one of the catalytic domains mutated, been able to produce a single strand break. In this strategy two Cas9n are used, each one targets adjacent regions of different strands producing a single strand break close enough to be recognized by the cell as a double strand break. And be repaired by the NHEJ repair system causing indels mutations and a possible disruption of the gene targeted.

1.7. Aim of the project

The objective of this study was to generate mutant cell lines disrupting Androgen Receptor and PIP5K1A genes in prostate cancer cells. For this purpose we used a CRISPR-Cas9 genome editing approach that targets AR and PIP5K1A sequences close to the AUG start codons. The hypothesis under investigation was to explore the functional relationship between AR and PIP5K1A as suggested by studies using small molecule inhibitors of PIP5K1A. The cell lines would therefore be of potential value as cellular models to explore mechanisms of CRPC development.

CHAPTER 2

MATERIALS & METHODS

2.1 Sources of materials & their composition

2.1.1 General reagents

All general laboratory reagents were purchased as analytical grade quality from either Fisher Scientific or Sigma-Aldrich. Phosphate Buffered Saline (PBS) tablets from OXOID Ltd were used to prepare PBS solutions. All solutions were made in double deionised water; deionised water was purified by passage through a Pure Neptune water purification system. The pH of solutions was maintained by checking the pH with a pH meter (Jenway 3510), calibrated against standards of known pH. Sterilisation was achieved either by filtration through 0.2 micron filters, or via autoclaving at 121 °C for 15 minutes.

2.1.2 Bacterial reagents

Tryptone, and bacteriological growth medium ingredients were purchased from OXOID Ltd. The *Escherichia coli* strain DH5α (Hanahan 1983) was purchased from Stratagene.

Luria-Bertani (LB) medium = 1% (w/v) tryptone, 0.5% (w/v) yeast extract and 1% (w/v) NaCl in diH2O, made to pH 7.0 with NaOH. For solid medium, 2% (w/v) bacteriological agar was added. Autoclaved and stored at room temperature until required.

Solid LB medium = LB agar medium was melted and the desired antibiotic was added to make a final concentration of 1x. Approximately 20 ml was poured into each petri dish to cover the surface and allowed to solidify. Plates were stored at 4 ºC for up to 1 month.

Ampicillin = 1000x stock solution made as 100 mg/ml in 50% ethanol, 50% sterile diH2O. Stored at -20 °C.

2.1.3 Molecular biology reagents

Restriction digestion, kinase, phosphatase and ligase enzymes and their relevant buffers were purchased from either Roche or New England Biolabs (NEB).

DNA polymerases, along with their reaction buffers, supplements and dNTPs were purchased from NEB, as were DNA ladders.

Agarose gel purification kits were purchased from York Biosciences, Nucleospin® plasmid kits (miniprep kits) from Macherey-Nagel and plasmid maxi kits (maxiprep kits) from QIAGEN Ltd.

The GenEluteTM mammalian genomic DNA miniprep kits were purchased from Sigma-Aldrich. Oligonucleotides were purchased from Sigma-Aldrich as lyophilised and desalted pellets. They were resuspended in sterile water to a 100 µM concentration and stored at -20 °C.

Agarose gel loading dye (10x) = 3.9 ml glycerol, 500 µl 10% (w/v) SDS, 200 µl 0.5 M EDTA, 0.025 g bromophenol blue & made up to a total volume of 10 ml with diH2O. Used at a final concentration of 1x.

Tris-borate-EDTA (TBE) = 40 mM Tris-base, 40 mM boric acid, 0.5 mM EDTA (pH 8.0) and made up to 1 L with diH2O.

2.1.4 Biochemical reagents

The SDS-PAGE molecular weight protein marker and nitrocellulose membrane were purchased from Sigma-Aldrich and Bio-Rad, respectively.

The 30% (w/v) acrylamide mix was also purchased from Sigma-Aldrich, as was the N,N,N,N-tetramethylethylenediamine (TEMED).

The primary antibody mouse anti-VP16 was from Santa Cruz Biotechnology (sc- 7546), whereas the rabbit anti-LexA antibody was from EMD Millipore (06-719).

Horseradish peroxidise (HRP) conjugated secondary antibodies chicken anti-mouse and goat anti-rabbit were all purchased from Santa Cruz Biotechnology (sc-2954 & sc-2004, respectively).

SDS-PAGE loading buffer (4x) = 2.0 ml 1 M Tris-HCl pH 6.8, 0.8 g SDS, 4.0 ml 100% glycerol, 0.4 ml 14.7 M β-mercaptoethanol, 1.0 ml 0.5 M EDTA and 8 mg bromophenol blue, made up to 10 ml with diH2O.

Ponceau S solution = 0.1% (w/v) Ponceau S in 5% (v/v) acetic acid. Made up to a total volume of 1 L with diH2O.

Developing solution = 1 ml Tris-HCl pH 8.5, 6 µl H2O2, 22 µl solution A (90 mM p-Coumari acid in DMSO), 50 µl solution B (250 mM luminol in DMSO) to a final volume of 10 ml with diH2O.

5% milk blocking solution = 5% (w/v) Marvel 0% fat milk powder in 1x PBS. Z buffer = 60 mM Na2HPO4·7H2O, 40 mM NaH2PO4·H2O, 10 mM KCl, 1 mM MgSO4·7H2O and 50 mM β-mercaptoethanol, made up

to a total volume of 1 L with diH2O and pH adjusted to 7.0. Stored at 4 °C.

Tris-glycine-SDS (TGS) buffer (10x) = 250 mM Tris base, 2 M glycine, 35 mM SDS, pH 8.3.

Transfer buffer = 48 mM Tris base, 39 mM glycine, 0.037% (w/v) SDS, 20% (v/v) methanol.

2x HBS solution = 50 mM HEPES, 280 mM NaCl, 1.5 mM Na2HPO4, made up to a total volume of 100 ml with diH2O and pH adjusted to 7.07 with 1 M HCl. Filter sterilised and stored at -20 °C until use.

10% Ammonium persulphate (APS) solution = 10% (w/v) APS in diH2O. Stored at -20 °C until use.

10% Sodium dodecyl sulphate (SDS) solution = 10% (w/v) SDS in diH2O.

RIPA buffer use in cell lysis: 50 mM Tris-HCl, 1% Igepal CA-630, 150 mM NaCl and 0,1% SDS.

2.1.5 Tissue culture reagents

All tissue culture reagents were purchased from Lonza Biowhittaker and all plasticware was obtained from Nunc or Techno Plastic Products.

2.2 CRISPR-Cas9 synthetic guide RNA design

The CRISPR design tool from MIT (**http://crispr.mit.edu/**) was used for selecting the SgRNA. The sequences used for designing the guides were downloaded from NCBI Genbank database. The guides with highest score

were chosen and the oligo were purchased as normal oligonucleotides from Sigma-Aldrich.

2.3 Bacterial preparation and culture

2.3.1. Preparation of heat-competent *Escherichia coli*

Escherichia coli DH5α cells were streaked on an antibiotic-free LB agar plate and incubated at 37 °C overnight. A single colony was used to inoculate a 10 ml volume of LB medium (antibiotic-free) and cultured overnight at 37 °C, with shaking. This was then used to set up a 300 ml total volume culture in the same medium and was incubated at 37 °C, with shaking until the optical density (measured at 600 nm (OD600)) reached 0.5. From this point onwards, all work was performed at 4 °C. The culture was centrifuged for 10 minutes, 4,000 rpm at 4 °C in separate 50 ml Falcon tubes and the supernatants removed. Each pellet was washed in 8.3 ml of 0.1 M MgCl2, before the centrifugation was repeated. The supernatants were discarded and each pellet was resuspended in 8.3 ml of 0.1 M CaCl2. These suspensions were incubated on ice for a minimum of 20 minutes, before repeating the centrifugation. The supernatants were removed and all of the pellets were resuspended together in a total of 8.6 ml of 0.1 M CaCl2. 1.4 ml 100% glycerol was added to the suspension, inverted 5 times and 200 µl volumes were aliquot into microcentrifuge tubes. These were snap frozen immediately in liquid nitrogen and stored at -80 °C.

2.3.2. Bacterial transformations

Escherichia coli DH5α cells were utilised in all bacterial transformations. An aliquot of these heat-competent cells was allowed to thaw on ice for a maximum of 10 minutes after removal from storage at -80 °C. 50 ng of intact vector/plasmid, or 10 µl of ligation products was added to 50 µl of cells. The mixture was allowed to sit on ice for 30 minutes, then heat shocked at 42 °C for 2 minutes, before being placed on ice for a further 5 minutes. 950 µl of Luria-Bertani (LB) medium was added and each sample was cultured at 37 °C for 1 hour with shaking, before centrifugation for 2 minutes at 7,000 rpm. 920 µl of the supernatant was removed, the pellet was resuspended in the remaining liquid and the entire volume was plated on LB agar supplemented with ampicillin. Each agar plate was incubated overnight at 37 °C. For each set of transformations, a 50 µl volume of DH5α cells was subjected to the same procedure, although without the addition of any vector/plasmid, to act as a negative transformation control. A positive control of the original undigested vector was also used. In the case of transforming ligation products, a negative ligation agar plate was set up in addition, which contained the transformed products of the negative ligation sample.

2.4 Molecular biology techniques

2.4.1. Annealing of oligonucleotides

2 µL of each conplementary oligonucleotides of the CRISPR-Cas9 sgRNA are placed separately at 37 ºC for 1 hour mixed with 10x T4

DNA ligase buffer (2 µL), 1µL of polynucleotide kinase and 15µL of diH_2O, and incubated at 65 ºC for 20 minutes to inactivate the enzyme. Once complete, 5 µl of each complementary top and bottom oligonucleotide were mixed together and made up to a total volume of 50 µl with diH_2O. These mixtures were then placed in a thermocycler, which was used to heat the mixture to 95 °C for 3 minutes, then reduce by 1 °C each minute until a temperature of 26 °C was reached, followed by 5 minutes at 25 °C. The products were stored at -20 °C until required.

2.4.2. Restriction digestions

DNA was digested with the desired enzyme using the manufacturer's recommended buffer at the recommended temperature for 2 hours. DNA from plasmid were digested with restriction endonucleases for posterior cloning of SgRNA. Expression vector pspCas9n(BB)-2ª-GFP was digested with BbsI restriction enzyme for 2h. Later, phosphatase treatment was required so the plasmid was incubated at 37 ºC with 1 µL of TSAP (thermosensitive alkaline phosphatase) in order to dephosphorilate the 5' end and avoid religation of the plasmid. After, incubation for 20 minutes at 74 ºC was necessary to make TSAP inactive.

2.4.3. Ligation of DNA fragments

Previously annealed oligonucleotides needed to be ligated into the expression plasmid to construct a vector containing the SgRNA. The

enzyme used was T4 DNA ligase. Reactions of ligation were done in a molar ratio of 1:1 (vector: insert). 50 ng of DNA from the vector and the same amount of DNA of the insert were mixed with 2 µL of 10x T4DNA ligase buffer, 1µL of T4 DNA ligase enzyme to make a total volume of 20µL with diH_2O. After mixture, it was allowed to ligate overnight at 4 ºC. The same mixture without DNA insert was used as negative control to determine the level of undigested vector.

2.4.4. Agarose gel electrophoresis

Analysis and visualization of PCR products and restriction enzyme digestions was performed via agarose gel electrophoresis. Agarose gels were made by mixing the agarose (w/v) needed for the desired percentage agarose gel with 1 x TBE buffer, after mixing the solution was heat until a clear liquid was formed. 2 µl of ethidium bromide (10 mg/ml stock concentration) was added per 50 ml of liquid agarose, before pouring into the desired plate and allowing to solidify at 21 ºC (room temperature). All DNA samples were mixed with 10x DNA loading buffer to make a final concentration of 1x DNA loading buffer before loading the gel. 10 µl of the selected DNA ladder was also loaded. All agarose gels were ran at 90 V in 1x TBE buffer. The gels were ran for a different length of time depending on desired DNA band separation. DNA bands were visualized and imaged using a BIO-RAD Universal Hood II UV transilluminator.

2.4.5. Purification of DNA from an agarose gel

Products from PCR experiments were purified using the YORBIO Gel/PCR DNA purification kit. The bands present in the agarose gel could be visualized using an UV box, precisely excised and placed separately microcentrifuge tubes. The amount of agarose excised was weighted and purification was performed according to the protocol. DNA concentration was measured using a spectrophotometer and sent for sequencing or stored at -20 ºC.

2.4.6. Polymerase chain reaction (PCR)

Polymerase chain reaction was utilised to produce fragments of DNA necessary for ligation into vectors or plasmids, or for gene editing analysis of transfected cells. Taq Polymerase® (New England Biolabs) and Pfu DNA polymerase (thermo Fisher) were used for this purpose following the manufacter's recommendations of quantities and temperatures. The experiments were optimized adjusting the number of cycles and adding different concentrations of DMSO when it was required. All the reactions were performed in the Applied Biosystem 2720 thermal cycler.

2.4.7. Small-scale purification of vectors & plasmids from bacteria (miniprep)

A mini-prep kit (Nucleospin plasmid kits supplied by Macherey-Nagel) was used to purify plasmids at small-scale. A single bacterial colony is selected from the transformation plate and inoculated into 10 ml of ampicillin-

supplemented LB medium. Incubation at 37 ºC overnight with shaking is required and after this period the culture can be centrifuged at 4 ºC for 10 minutes at 4000 rpm. Once the pellet is obtained the purification of plasmid DNA was achieved by following manufacturer's protocol. Final DNA concentration was measured using a spectrophotometer.

2.4.8. Large-scale purification of vectors & plasmids from bacteria (maxiprep)

Purification of plasmids was performed also at large-scale. For this, a maxi-prep kit (Plasmid Maxi Kit® from QIAGEN Ltd.) was utilised. As in small-scale purification, the first step consists in inoculating a previously selected bacterial colony from a transformation plate in ampicillin-supplemented LB (10 ml) and posterior addition of 300 ml of the same supplemented medium. Then, cultures were incubated overnight with shaking at 37 ºC. After this period, centrifugation at 4 ºC for 50 minutes at 6500 rpm was necessary. Once the pellet is obtained, the manufacturer's protocol from the mentioned kit was followed. Final DNA product was diluted in 200 μL of diH_2O and DNA concentration was measured using a spectrophotometer.

2.4.9. Purification of genomic DNA from mammalian cells

For genomic DNA purification The GenElute mammalian genomic DNA miniprep kits (from Sigma-Aldrich) was used. For extracting genomic DNA from the cells, these were first harvested, placed in a 2 ml tube and centrifuged for 5 minutes at 1500 rpm to obtain a cell pellet. Then,

the pellet was resuspended with 1 ml of 1X PBS for a washing step and centrifugation repeated. After washing, remaining PBS was aspirated and genomic DNA purification was performed following manufacturer's protocol. DNA concentration was measured using a spectrophotometer and could be stored for further use at -20 ºC.

2.4.10. DNA concentration quantification

The quality of DNA, as well as its quantification was determined using a NanoDrop® ND-1000 UV-Vis Spectrophotometer. The purity of DNA was assessed using the ratio of sample absorbance at 260 and 280 nm, 1.8 was accepted as pure. The ratio of sample absorbance at 260 and 230 nm was also used as measurement of DNA purity, and a range of 1.8 to 2.2 was considered acceptable. Values of the ratios lower than 1.8 or not in the range 1.8-2.2 was considered as indication of cantamination in the sample.

2.4.11. Sequencing of DNA

For sequencing DNA samples, it was required the sequencing service from Source Bioscience Ltd, that uses an Applied BioSystems 3730 DNA Analyzer® and a PeakTrace® software for the analysis. Depending on the nature of the sample sent for sequencing different volumes and concentrations were needed. Purified PCR products were sent at 1 ng/μl per 100 bp of product, and DNA samples from mini and maxi-prep

purifications were sent at a concentration of 100 ng/μL, along with the correspondent amount of primers in each case.

2.5. Cell culture

2.5.1. Maintenance

The human embryonic kidney (HEK293) cell line was typically maintained in Dulbecco's Modified Eagle Medium® (DMEM), supplemented with 10% (v/v) heat-inactivated fetal calf serum (FCS), 1% (v/v) 0.2 M L-glutamine and 1% Penicillin/Streptomycin antibiotics. LNCaP C4-2 were maintained in Roswell Park Memorial Institute 1640® (RPMI) medium supplemented with 10% heat inactive foetal bovine serum (FBS), 2 mM L-glutamine, 1% (v/v) Penicillin/Streptomycin and Sodium Pyruvate (Gibco). The cells passage was realized when the cells were 80% confluent, before the passage the cells were washed with 1x PBS, and trypsin-EDTA was used.

The cells were grown in 10 cm^2 dishes or 7.5 cm^2 flasks in an incubator at 37 ºC in a 5% CO_2 atmosphere.

2.5.2. Calcium phosphate-mediated transfection of adherent cells

At least 24h before transfection the adherent cells were seeded in a 12 or 24 well plate at a density of 1 x 10^5 cell per well. On the day of experiment the medium was replaced 2-4 hours bedore transfection. Aliquots of 12.5 µl 2M calcium chloride solution was mixed with 500 ng

of plasmid DNA , diH_20 was added to the mixture to make a final volume of 100 µl. For each sample 100 µl 2x HEPES Buffered Saline® (HBS) solution was prepared. The amounts described correspond to a 12-well plate experiments. Then the calcium chloride and the plasmid mixture was added drop by drop to the HBS solution tube while vortexing. This mixture was incubated for 20 minutes at room temperature, and then added to the cells in a dropwise manner. The day after the experiments the cells were washed with PBS 1x and the medium was replaced. The cells were incubated at 37 ºC in humidified 5% $C0_2$ incubator.

2.5.3. Electroporation transfection

Transfection of LNCaP C4-2 cells was performed using Neon® Transfection System (Invitrogen). 24 h before tranfection the growth medium was replaced. The day of the experiment the cells were washed with an special PBS (without Mg or Ca) and then the cells were dislodge using Trypsin/EDTA for 3 minutes at 37 ºC. Then the Trypsine was neutralized using growth medium. The cells density was calculated and the cells were centrifuged at 1500 rpm for 5 minutes using PBS without Mg or Ca twice. The preparation for transfection was perfomed following the manufacter protocol for adherent cells. The electroporation was performed using the manufacters recommendation of plasmid DNA concentration, cell number and volume per well. 10 µL Neon tips were used as explain in the protocol, the LNCaP pulse conditions suggested

by the manufacters were used in our experiment. Transfection efficiency was assessed using confocal mycroscope.

2.6. Biochemical techniques

2.6.1. Preparations of whole-cell extracts

For harvesting the adherent cells, the cells were washed twice with cold 1x PBS and then dislodged from the dishes using a scraper. The PBS containing the cells was centrifuged for 5 minutes at 1500 rpm. The supernatant was removed and the pellet was resuspended in 150 µL of lysis buffer (RIPA buffer supplemented with 1 mM PMSF and complete protease inhibitors), then was incubated on ice for 20 minutes. Then the cells were sonicated (Diagenode water bath sonicator), and the supernatant was then transferred into a new microcentrifuge tube properly labelled to determine protein concentration.

2.6.2. Protein concentration determination

The Bradford assay method was used for measuring the protein concentration on the cell extracts. The concentration was calculated using a spectrophotometer and a linear regression previously calculated. The linear regression was performed with Bovine Serum Albumina (BSA). In a normal assay, 2 µL of protein extract were mixed by inverting 5 times with 800 µL of diH_20 and 200 µL of Sigma- Aldrich Bradford dye. The samples were then incubated at room temperature

for 15 minutes. The absorbance was read at 595 nm and concentration calculated.

2.6.3. Sodium-dodecyl-sulphate polyacrylamide gel electrophoresis (SDS-PAGE)

The protein samples were loaded and analysed by one dimensional polyacrylamide electrophoresis. SDS-PAGE gels were prepared following the manufacter's protocool.

The gel was loaded with 150 µg of protein extract mixed with SDS-PAGE loading buffer, no more than 30 µl per sample, and ran with 1x TGS buffer at 170V.

2.6.4. Western blotting

Proteins were transferred from SDS-PAGE gels onto a nitrocellulose membrane by electrophoresis at 30 V and overnight at 4 ºC,using a glycine transfer buffer. After that, the 1% Ponceau S solution was applied to the membrane for no more than 1 minute, in order to confirm the protein transfer form the gel to the membrane. Then the membrane was blocked for 1 hour in 5% milk blocking solution. Then the membrane was incubated with the selected primary antibody at 4 ºC overnight. Finally the membrane was washed three times before incubated with the secondary antibody for 90 minutes at room temperature. Then the membrane was washed with PBS mixed with tween 20. ECL developing solution was used to detect the protein. The chemiluminiscent signal was detect by Luminescent image analyser

(Fujifilm LAS-4000® system). Exposure times were typically 30 seconds,5 minutes or 15 minutes. The protein marker was also imaged as a reference of molecular weight.

CHAPTER 3

RESULTS

3.1 CRISPR-Cas Synthetic RNA Guides Design for CRISPR-Cas Genome Editing Tool

For this project the CRISPR-Cas system type II from *S. pyogenes* was chosen, and in order to minimize the off target activity of the Cas9 the double nicking strategy was performed. In this strategy, instead of a normal Cas9, which produces a double strand break as explained in the introduction, a mutated Cas9 is used. This enzyme has one of the catalytic domains mutated, and it is known as Cas9n. Then, two single strands breaks are produced, by two different Cas9n targeting adjacent regions in the DNA at the same time. If both single strand breaks are achieved, this double nicking will produce the same outcome as a double strand break, being repaired by the cell through the non-homologous end joining repair pathway, and producing indel mutations in the targeted region causing a gene disruption. These RNA guides must be sequence of 20 nucleotides long, with a consensous PAM sequence in their 5`end –NGG. This PAM sequence is not part of the sgRNA, but part of the target DNA. To design the sgRNA the CRISPR Design Tool platform provided by the Massachussets Institute of Technology (MIT) was used.

The CRISPR Design tool used provided different pairs ranked according to the quality and the off-target score. In the figures 6 and 7 the sgRNA design is shown, and a double nicking strategy model. The target region selected was the ATG, in order to produce indel mutations causing the gene disruption.

The RNA guides were chosen according to quality scores as well as being close to the desired target region. In some cases no RNA guides were generated after the ATG with an acceptable score. In consequence, some target regions selected were in the 5'UTR region.

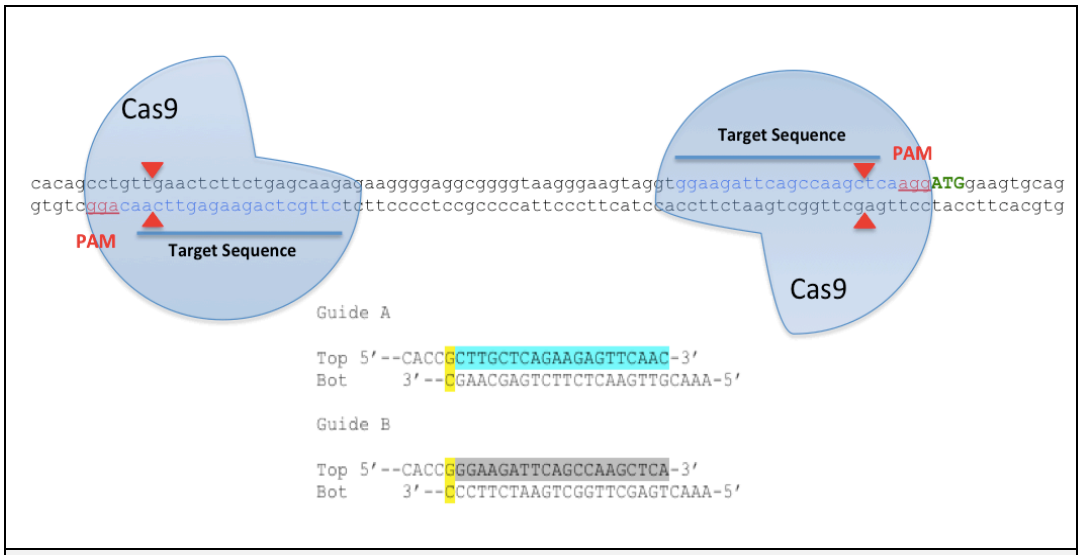

Figure 6

AR sgRNA and double nicking model of the Cas9n.

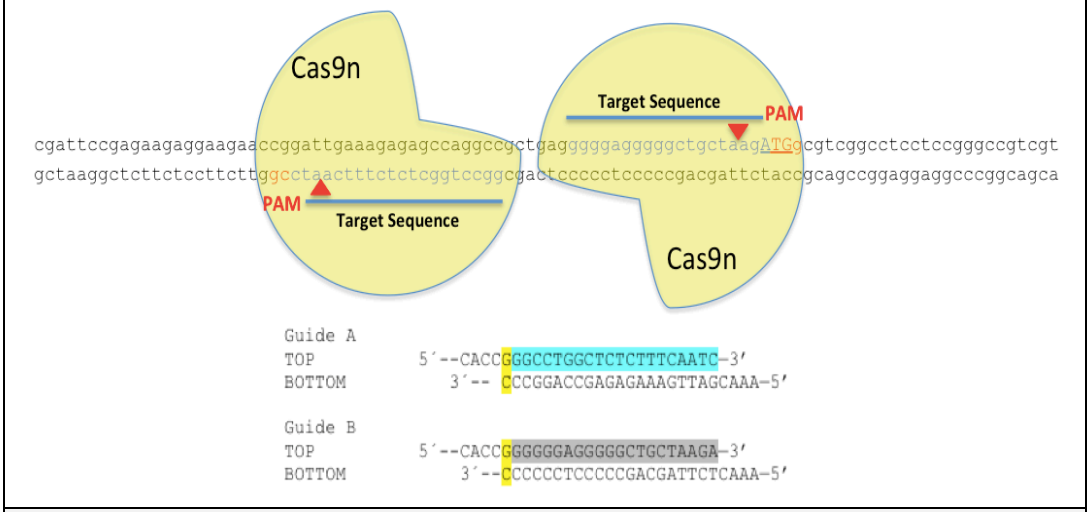

Figure 7

PIP5K1A sgRNA and double nicking model of the Cas9n.

3.2 Cloning the sgRNA into the vector and bacterial transformation

After the guide design, the next step was the construction of the expression vector that was going to be used. For this experiments the selected vector was pSpCas9n(BB)-2A-GFP (Px61 from Addgene repository)(Figure 8). This vector encodes for Cas9n and a conserved sgRNA scaffold; the vector also has U6 RNA polymerase III promoter, BbsI restriction site and 2A-GFP marker gene which it is used as positive transfection marker; the vector also encodes for ampicillin resistance gene, which it is used as a positive transformation marker in *E.coli.*

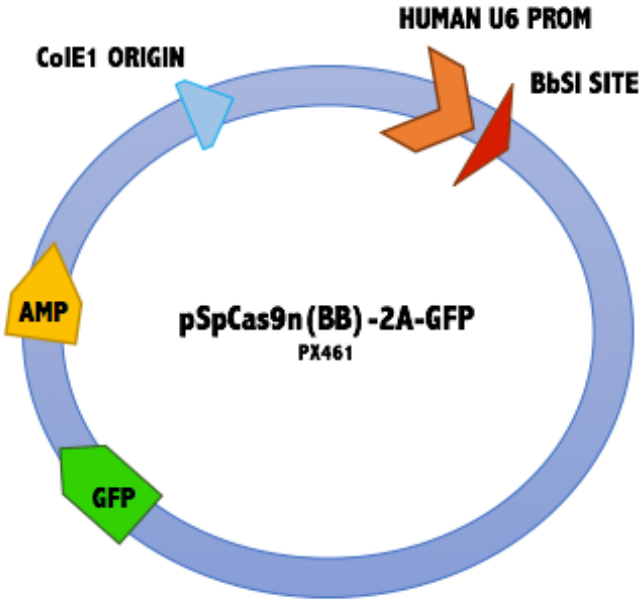

Figure 8

Px461 vector showing AMP, GFP, Human U6 promoter and BbSI site.

The plasmid was digested with BbsI restriction enzyme (2h at 37 ºC). The digested vector was separated in an agarose gel, and then purified. The sgRNAs were inserted into the vector using T4 DNA ligase enzyme. After the ligation the expression vector was then transformed into the competent cells (*E.coli* DH5α) using the $CaCl_2$ technique. In this technique the plasmid DNA and the competent cells are incubated together in a calcium rich buffer, and after a sudden increase of the temperatures which produces pores in the plasma membrane, the plasmid can pass through the pores into the bacteria cell. The transformants were selected on LB agar plates containing ampicillin. Only the cells that were expressing the plasmid could grow, a negative control plate was performed as well. After 24h single colonies were selected from the plate and cultured overnight in LB media supplemented with ampicillin. The DNA was purified using a Mini-Prep kit, and then using a primer complementary to the U6 promoter the samples were sent for DNA sequencing to check that the correct sgRNA sequence was inserted. The Figure 9 shows the sequencing results and the digested vector.

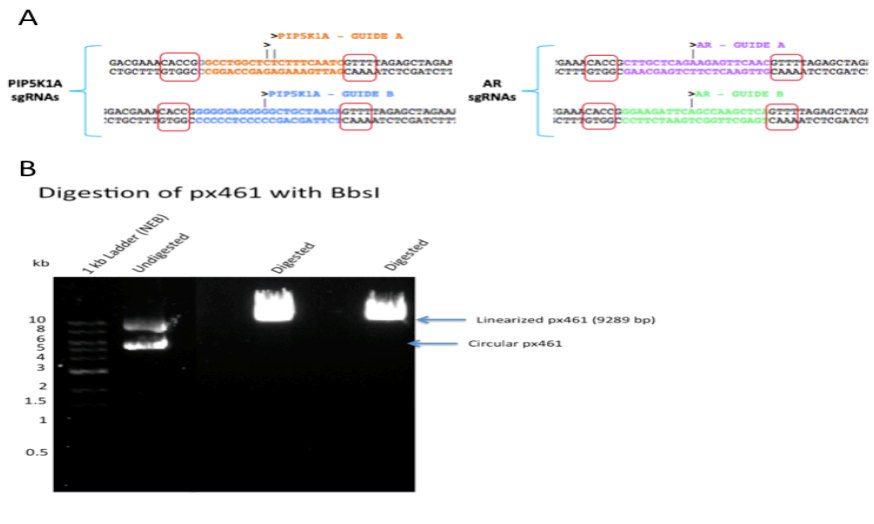

Figure 9

A) Sequencing results showing the correct insertion of the guides into the vector and the BbsI site. B) Image of BbsI digested vector.

3.3 Checking the transfectability of the plasmids

In order to verify that the constructs were sufficiently pure for cell transfection, a preliminary transfection was performed in Human Embryonic Kidney 293® (HEK293) cells. This cell line was selected because they are easily transfected and also easy to maintain. The method used for the transfection was the calcium phosphate-mediated transfection. This method for introducing DNA into mammalian cells is based on producing calcium phosphate-DNA precipitates, the calcium enhance the binding of the DNA to the cell surface, and then the DNA enters the cell by endocytosis. This method is widely used because of its high cost-efficiency rate. The experiment is described in Figure 10.

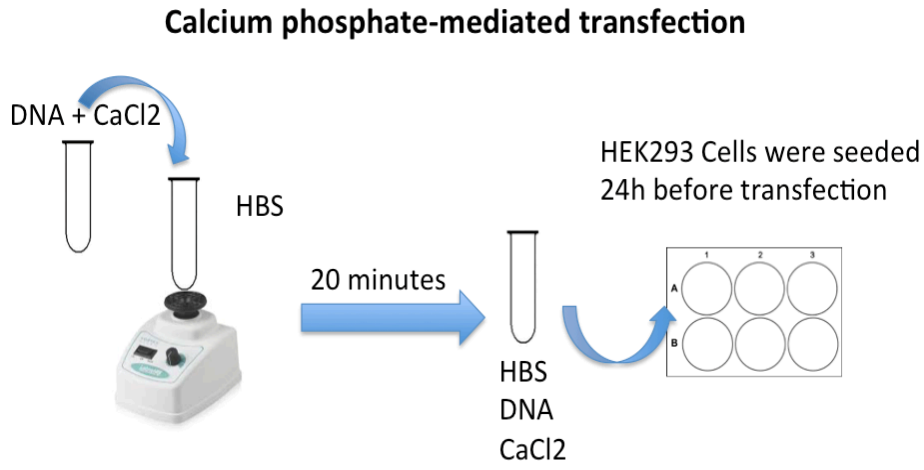

Figure 10

Calcium Phosphate-mediated transfection diagram.

250 ng of DNA were mixed with Cacl 2M, then was added to the HEPES Buffered Saline® solution while vortexing. After 20 minutes the mixed solution was added to the HEK293 cells that were seeded 24h prior to transfection ($1x10^5$ cells). The growth media was replaced 2h before transfection. After 16 h the cells were washed with 1x PBS and growth media was replaced to remove excess of precipitate. Then the transfection efficiency was assessed by confocal microscopy. As explained before the vector encodes for the EGFP protein, the green fluorescent protein can be observed using confocal microscopy and it is a marker of positive transfected cells. The results showed high transfection efficiency for all single and double sgRNA vectors used. The images from this experiment are shown in the figure 11.

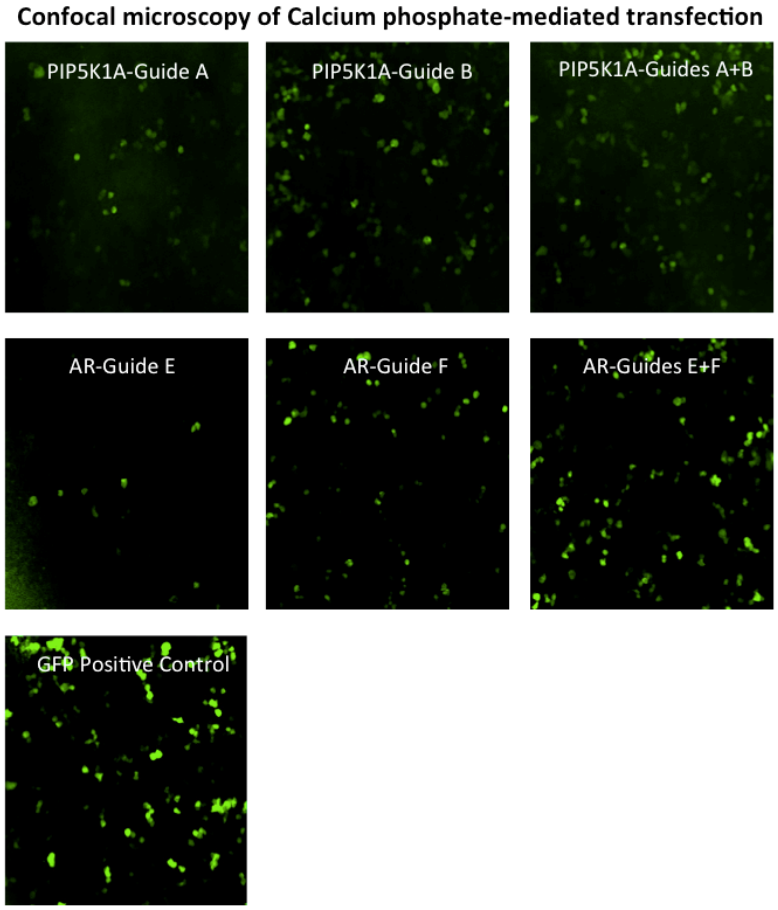

Figure 11

Confocal microscopy images from Calcium phosphate-mediated transfection of HEK-293 cells.

3.4 Protein expression analysis of targeted genes

The expression of androgen receptor and PIP5K1A was analysed by Western Blot in different cell lines in order to verify that the proteins were expressed. The Western Blot technique is an analytical technique used to detect specific proteins. The protein is extracted from the samples, separated in a SDS polyacrylamide gel electrophoresis and transfered to a nitrocellulose membrane. Then the membrane is incubated with specific antibody to detect the expression of the protein

of interest. For this experiments the protein was extracted when the cells were between 70 and 90% confluent. After harvesting the cells were incubated with lysis buffer for 20 minutes and cell disruption was achieved by sonication. The protein concentration was assessed by Bradford colorimetric assay method. 150 µg samples of protein were resolved on a 12% polyacrylamide gel by SDS-PAGE technique. The samples were then transferred to a Nitrocellulose membrane, and the blot was incubated with the desired primary and secondary antibody.
AR expression was tested in six different cell lines, MDA-MB-231 and MDA-MB-468 breast adenocarcinoma cell lines, human enbryonic kidney HEK-293, and three prostate cancer cell lines (PC3, 22rv1 and LNCaP C4-2).

Figure 12 shows protein expression of AR in the selected cell lines, as expected MDA-231, MDA-468 and PC3 showed no expression of AR. Although PC3 is a prostate cancer cell line, it does not express AR according to the literature (43). HEK-293 showed a low level of expression; LNCaP C4-2 and 22rv1 showed the highest expression. Note that 22rv1 also expresses AR-V7, this AR isoform without ligand-binding domain is characteristic from the CRPC, so detection of both isoforms is consistent with the literature (22).

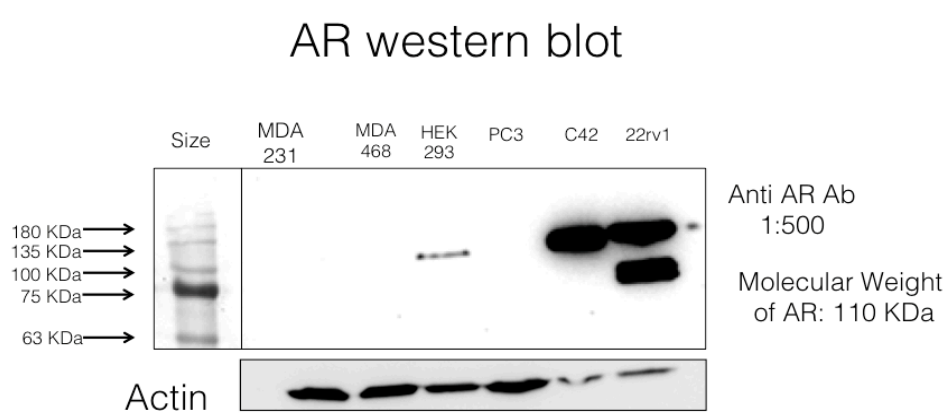

Figure 12

Androgen receptor Western Blot.

In Figure 13 the PIP5K1A expression was analysed only in prostate cancer cell lines, as expected LNCap C4-2 showed the highest expression, consistent with the literature (22). Note that two bands were visible in each cell line consistent with other studies (23).

The analysis of the expression of our target genes in different cell lines was accomplished in order to confirm the suitability of these cell lines for gene targeting. And also to assess if the expression was consistent with the literature. LNCaP C4-2 cells were chosen as a suitable cell line for genome editing because shows robust expression of both AR and PIP5K1A and also these cells are used in xenograft models.

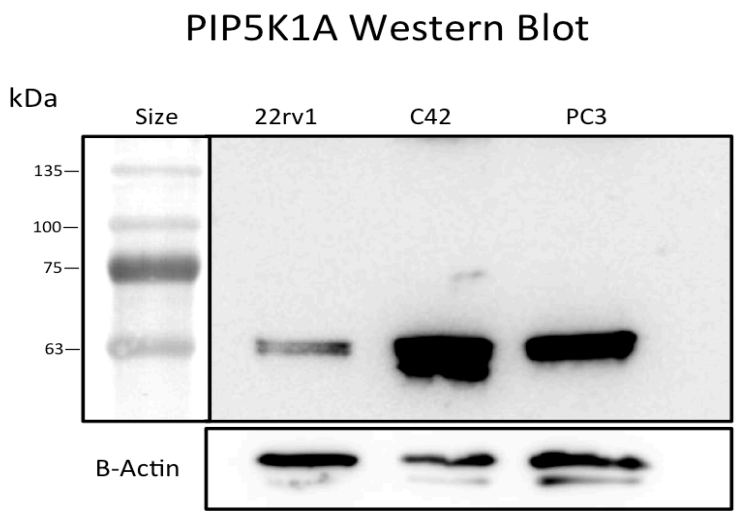

Figure 13

PIP5K1A Western Blot.

3.5 Transfection of LnCaP C4-2 by electroporation

Having established the successful transfection of HEK293 cells with the AR and PIP5K1 targeting vectors we next focussed on transfection of LNCaP C4-2. This cell line is a variant of LNCaP, which is a human prostate adenocarcinoma cell line derived from the left supraclavicular lymph node metastasis from a 50-year-old caucasian male in 1977. The C4-2 variant is the most similar to LNCaP among all variants, but shows features of progressed disease such as metastatic capability and hormone Independence (44). In addition this cell line presents an accelerated growth rate compared to LNCaP. In order to achieve the highest efficiency in the transfection different methods were tested. In addition to calcium phophate transfection we also tested the TransIT-X2 Dynamic delivery system (Mirus) and electroporation using Neon transfection system from invitrogen. Due the its highest efficiency

(GFP+) and the survival rate of the cells Neon Transfection method was used.

Neon transfection system (Invitrogen) is an electroporation based system, in which an electrical field is applied to the cells producing an increase of membrane permeability allowing the DNA to enter in the cell. Following the manufacturer's instructions, Maxi Prep DNA was prepared with a protocol to eliminate endotoxins, in order to obtained the highest cell viability after transfection. 24h after the transfection the efficiency was analysed by scoring for GFP+ cells using confocal microscopy and subsequently after a further 24hr culture, GFP+ cells were sorted by flow cytometry. The parameters used in the experiment are shown in Table 1.

Cell Density (cells/mL)	Pulse Voltage (v)	Pulse width (ms)	Pulse number
7×10^6	1,200	20	2

Table 1

Parameters for Neon Transfection System obtained from the protocols and cell line data.

3.6 Fluorescence Activated cell Sorting (FACS) and clonal expansion

Fluorescence activated cell sorting is a special type of flow cytometer in which the cells are sorted by the specific light scattering and fluorescent characteristics of each cell. This method is useful for separating a

heterogenuos mixture of cells, the flow cytometer analyzes the fluorescent characteristics of one cell at a time, and then the cells can be separated in different containers. Also the cells that have a different size are separated, in order to avoid dead cells or aggregates in the final container. For the purpose of our experiment the cells were sorted on the basis of GFP expression. First the cells were divided into positive GFP and negative, being an indicator of transfection efficiency. The cells that were found GFP positive were again sorted, in order to remove the cells that show a low GFP expression. Then these cells were isolated in a 96 well plate, one cell in each well. To analyse the CRISPR-Cas event that had happened in the cell, the next step was clonal expansion.

During the next 4 weeks the media was replaced several times in the 96-well plate, after one week with the same media the wells that presented brown colour were transferred to a T25 flask with fresh new RPMI complete medium. Once the amount of cells was considered appropiate the cells were harvested and the genomic DNA was extracted using Genelute Mammalian genomic DNA miniprep kit. From PIP5K1A-targeted cells, nine clones were expanded, named as: PI-1C, PI-2B, PI-2F, PI-3B, PI-3C, PI-6, PI-8G, PI-10H and PI-12H. For AR-targeted cells, four clones, named as: AR-B5, AR-D5, AR-D6 and AR-F5. As a positive control from the cells transfected with one sgRNA guide one clone was expanded, named as: SG-B8.

3.7 Molecular genotyping

3.7.1. Genomic extraction and PCR

Genomic DNA from the clones was extracted and the quality of the DNA was assessed by nanodrop. In order to detect CRISPR-Cas mediated events in the target locus PCR primers were designed to amplify the relevant genomic regions. For the AR gene the PCR was optimized using different amounts of DMSO in order to minimise the impact of DNA secondary structure in the target region and the polymerase that was used was the proof-reading PFU polymerase. The primers were designed using Primer3 and then the chosen primers (Figure 14) were tested in USCS genome browser in-silico PCR to check the primer specificity.

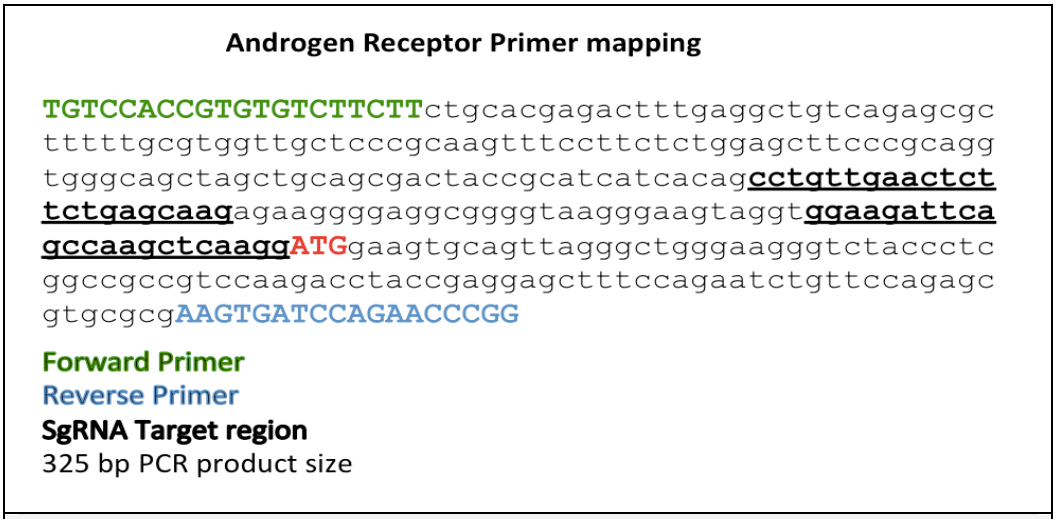

Figure 14

AR sequence indicating primer location and sgRNA target region.

The AR gene is present on the X chromosome and the male derived cell lines such as prostate cancer lines would generally be expected to

encode a single allele. This cell line has a variable number of chromosomes, the range is 61-90 chromosomes, the modal number is 83 chromosomes (45). For AR the PCR was performed using PFU polymerase, the results are shown in Figure 15. The analysis of the control cells (C 4-2 transfected with one guide) produced a fragment of 317 bp as expected. Clones AR-B5 appeared larger than wild type suggesting a insertion, the clones AR-D5 and AR-D6 appeared below the expected size, indicating deletions in the target region. Clone AR-F5 appeared to produce a freagment of the expected size.

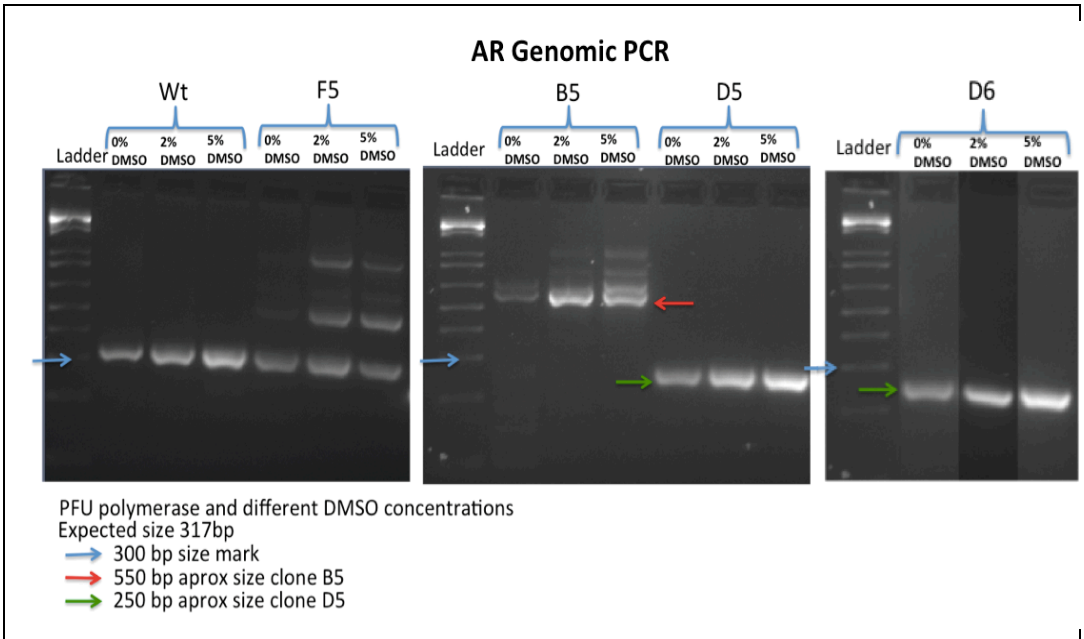

Figure 15

Agarose gel products from AR expanded clones.

After the Gel DNA purification the samples AR-B5 (2% DMSO), AR-D5 (5% DMSO), AR-D6 (5% DMSO) and wild type sample (5% DMSO)

were sent for sequencing. The clone AR-F5 was not sent for sequencing due to poor quality of DNA.

For PIP5K1A the Taq polymerase was chosen and no optimization with DMSO was needed. The primers were designed using Primer3 and checked against the UCSC genome browser In-silico PCR. During this process we found a pseudogene called PIPSL, with high sequence similarity with PIP5K1A. However the primers were chosen in a region that is not conserved and would therefore we not amplify PIPSL in the PCR reaction. The location of the primers and the PCR Product size is shown in the Figure 16.

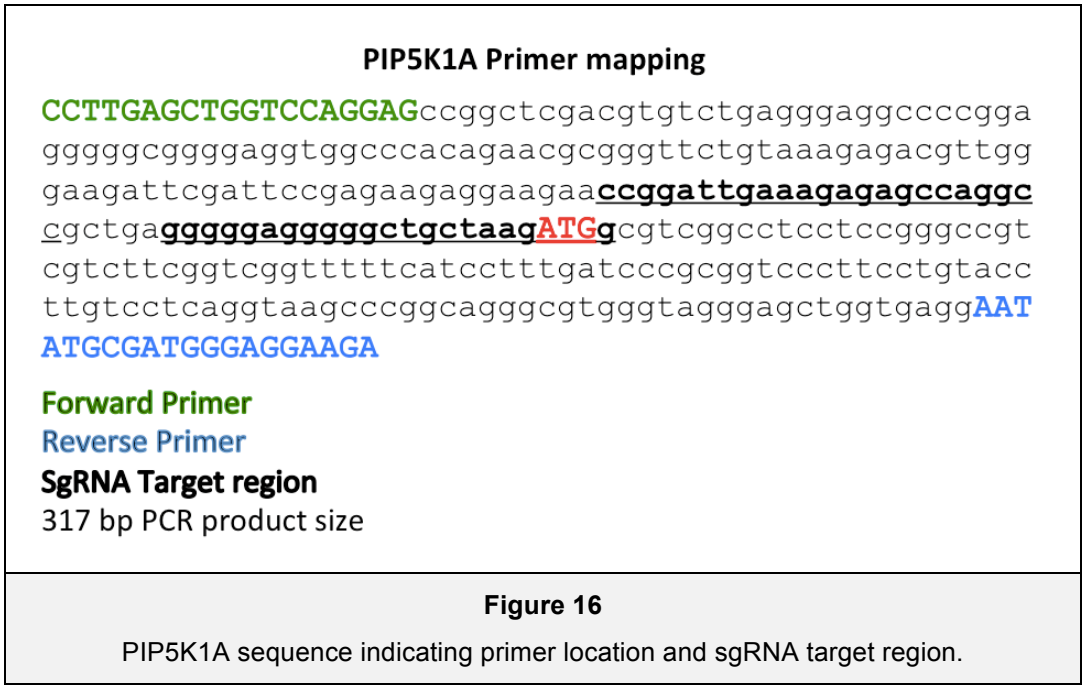

Figure 16

PIP5K1A sequence indicating primer location and sgRNA target region.

As shown in Figure 17 PCR analysis of genomic DNA from the control cells (C42 transfected with single guide) produced a fragment of the expected size (325bp). Several clones (1C, 3C, 8G, 10H) appeared

to produce single PCR fragments below the expected size, indicating the presence of deletions. Clones, 2B, 3B and 6A produced 2 bands of different sizes indicating indel events in one or more alleles. Clone 12H had a single product which appeared to be larger than the WT fragment. Clone 2F failed to produce any PCR fragment with the primers used (despite several attempts) suggesting the loss of WT alleles due to indels. These data provide strong evidence that the CRISPR Cas9 targeting of PIP5K1A and AR was successful and that all of the expanded GFP+ clones harbour indels of the target locus, in contrast to the single guide control, which appeared wild type. These PCR products were purified by Gel DNA purification and were sent for sequencing, all except 2F (due no amplification) and 3C (poor quality of the DNA).

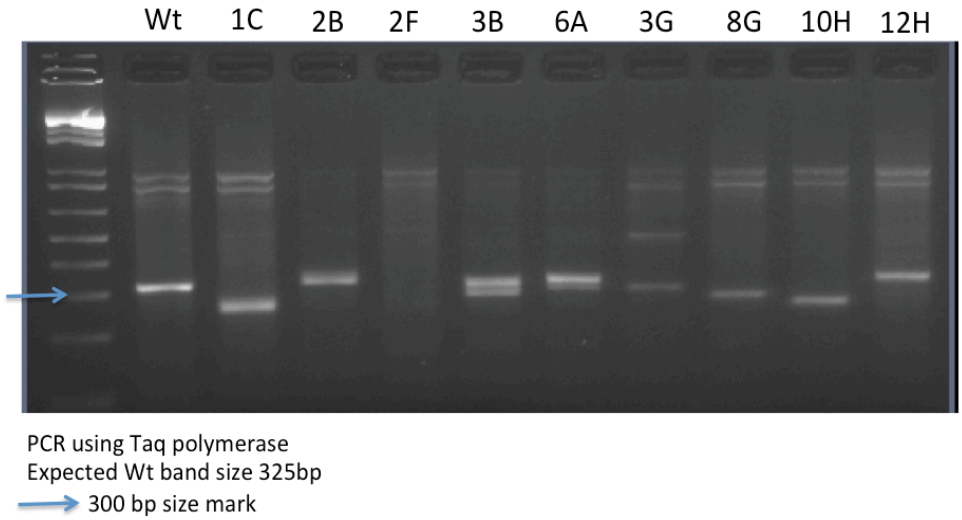

Figure 17

Agarose gel products from PIP5K1A expanded clones.

3.7.2. Sequencing analysis of genomic PCR bands

The PCR product were sequenced using the same primers used in the PCR and all samples analysed were sequenced with the forward primer and the reverse.

AR sequencing analysis

Genotyping was performed for a single guide control (Clone SG-B8) and three AR targeted clones; B5, D5, D6. The clone B5 appeared to produce a fragment of 600 bp indicating 280 bp insertion, clones D5 and D6 appeared to produce fragment below the expected size, indicating deletion of 75 bp approx. The wild type sequence was compared with the NCBI-AR sequence obtained from the NCBI web page, the blast showed 100% similarity. The sequence of the samples was compared against the wild type using the APE software for DNA sequence analysis.

Sequence analysis of clone D5 revealed a 73 bp deletion in the 5'-UTR region of the AR gene (Figure 18), consistent with the short PCR product detected by gDNA genotyping. This deletion occurs within the 5'UTR region of the AR transcript (1031-1100) at a position 15bp upstream of the AR ATG and may therefore affect AR expression via transcription, transcript stability or translation. Recurrent germline mutations in the 5'UTR of AR have been identified in Androgen Insensitive patients leading to reduced AR expression (46). It has also been reported that the 5'UTR region is regulated by transcription factors such as NFkB, SP1 and E2F (47)(48)(49). Similarly, Clone D6 was also

found to contain a 70bp deletion (1037-1106) within 9 bp upstream of the ATG (Figure 19). We also sequenced the B5, which showed an increased size compared to wild type in the genotyping. The sequence confirmed an insertion of 276 bp between the sgRNA target regions within the AR transcript (Figure 20). The sequence 1052-1073 (21bp) is deleted and replaced by the insertion. However the kozak ATG is still intact and may therefore be able to produce a protein (Figure 21).

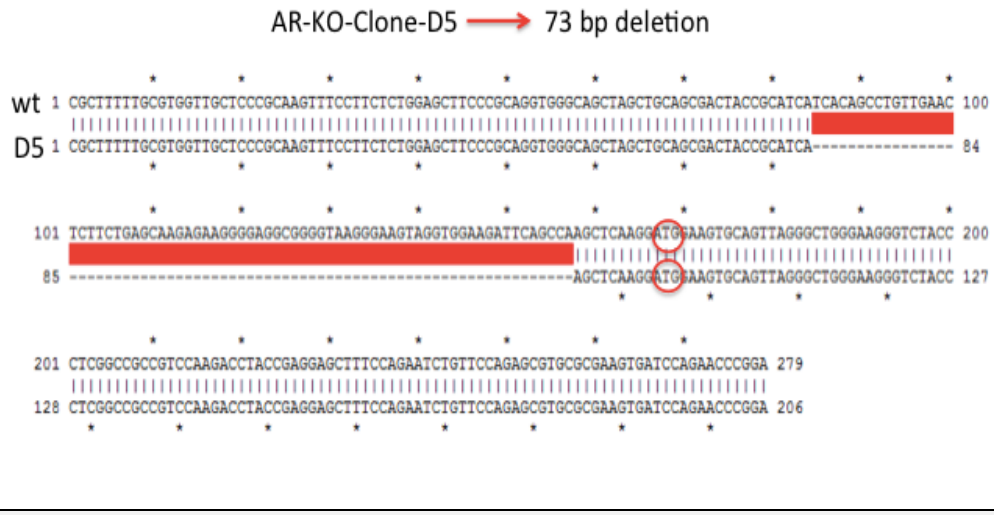

Figure 18

Clone D5 sequence and wild-type sequence alignment. In red, 73 bp deletion in the 5'-UTR region of AR.

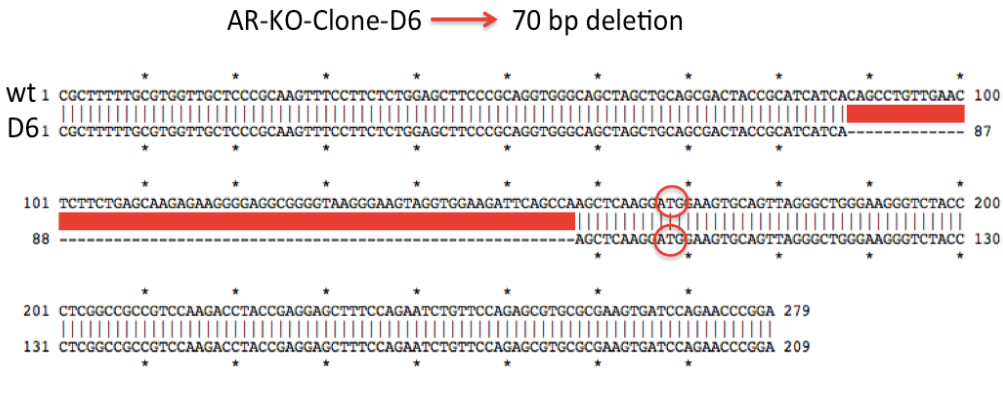

Figure 19

Clone D6 sequence and wild type sequence alignment. In red, 70 bp deletion in the 5'UTR region of AR.

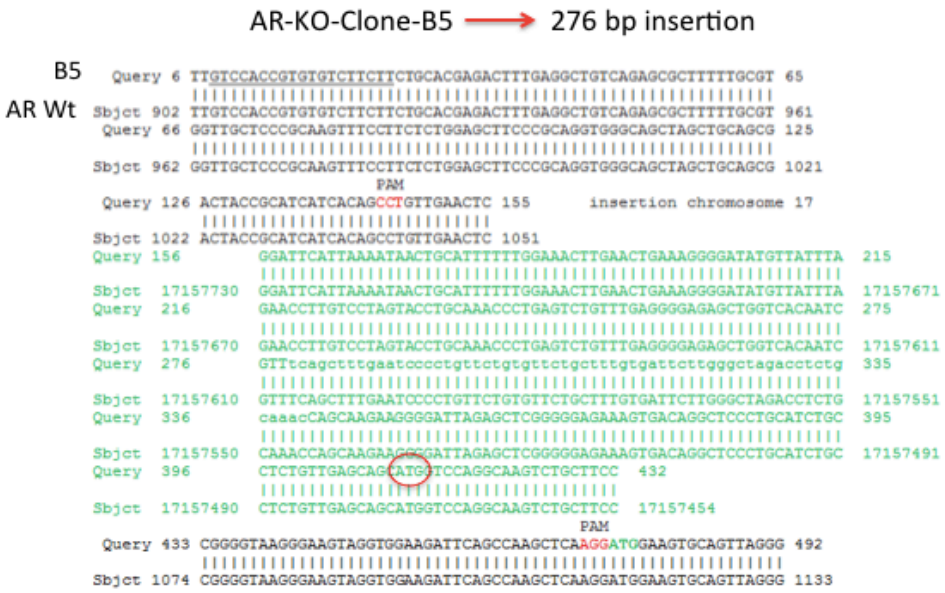

Figure 20

Clone B5 sequence and wild-type sequence alignment. In green, 276 bp insertion. Inserted sequence showed 100% identitity to a sequence from non-coding region of chromosome 17.

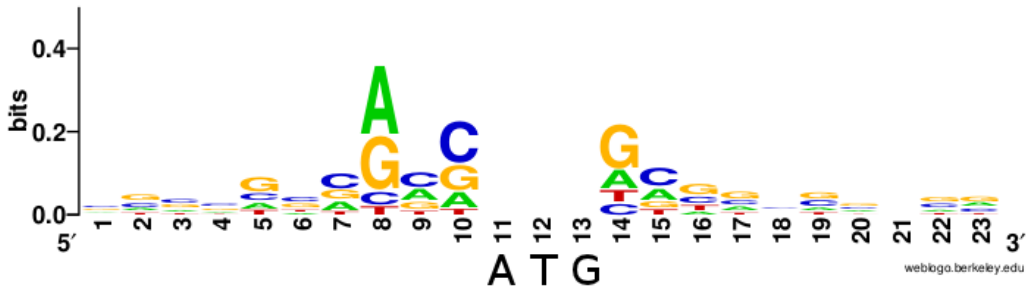

Figure 21

Kozak ATG sequence from the clone B5.

BLAST searches against the human genome revealed that the 276 bp fragment has 100% identity to a non-coding region on human chromosome 17. Thus it appears that in clone B5, sequence from chr 17 was used as a repair template by the NHEJ complex in response to the DSB generated by CRISPR targeting. It is unknown how sequences are selected to repair the damage sites. We checked for possible off-target hits of the single sgRNAs in this region of the chromosome 17. The analysis revealed no potential off-target sites in this region. The AR transcripts in these clones are shown schematically in Figure 22.

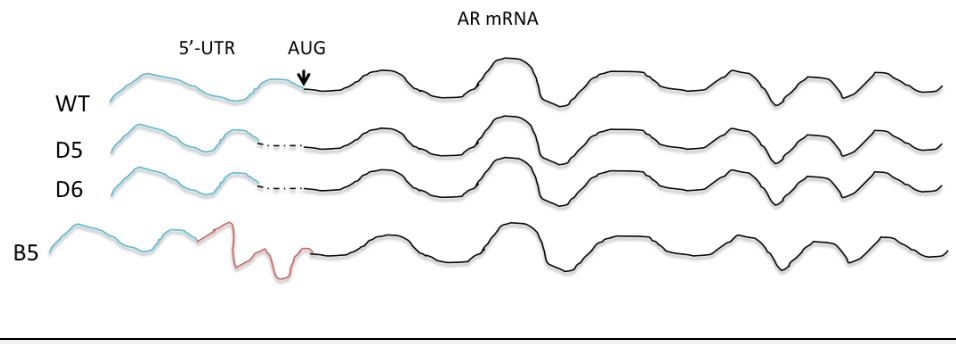

Figure 22

Different mRNA transcripts of AR clones.

PIP5K1A sequence analysis

The sequence analysis of seven PIP5K1A clones (1C, 2B, 3B, 6A, 8G, 10H and 12H) was performed and analysed using the analysis program APE. The samples were sequenced with the forward and reverse primers, as used in the genomic PCR. Most of the sequences analysis of the PIP5K1A-KO clones have mixed sequences, indicating multiple alleles, different events can had occurred in each allele and also there is a probability that some allele can be wild type based on the size of the PCR products. But analysing the sequence data we can predict

where are the approximate boundaries of the changed sequences. As a positive control a single sgRNA targeted clone, and confirmed to have wild type sequence as shown in figure 23.

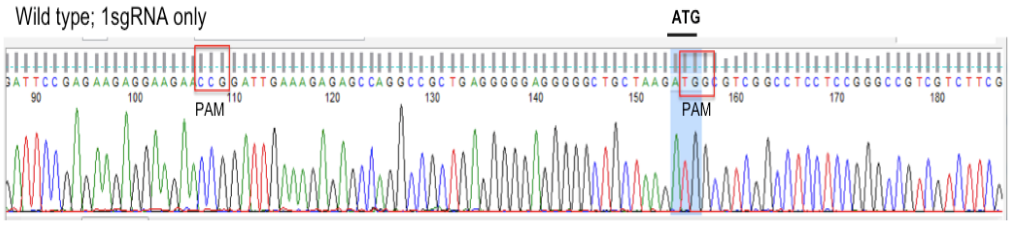

Figure 23

One sgRNA samples showing wild type sequence of the Human PIP5K1A gene. In the figure the red square indicates the PAM motif of the sgRNAs and the ATG.

As evident in the electropherograms of the PIP5K1A targeted clones, most show wild type sequence up to the targeted region. For several sequence reactions, it is seen that the peak height is reduced within the targeted region indicating multiple alleles. Moreover mixed sequence signals can be observed between the PAMs (Figure 24), meaning that different CRISPR-Cas-mediated events have occurred, leading to several alleles. We predict that some of the alleles have been disrupted, due to deletions of the ATG.

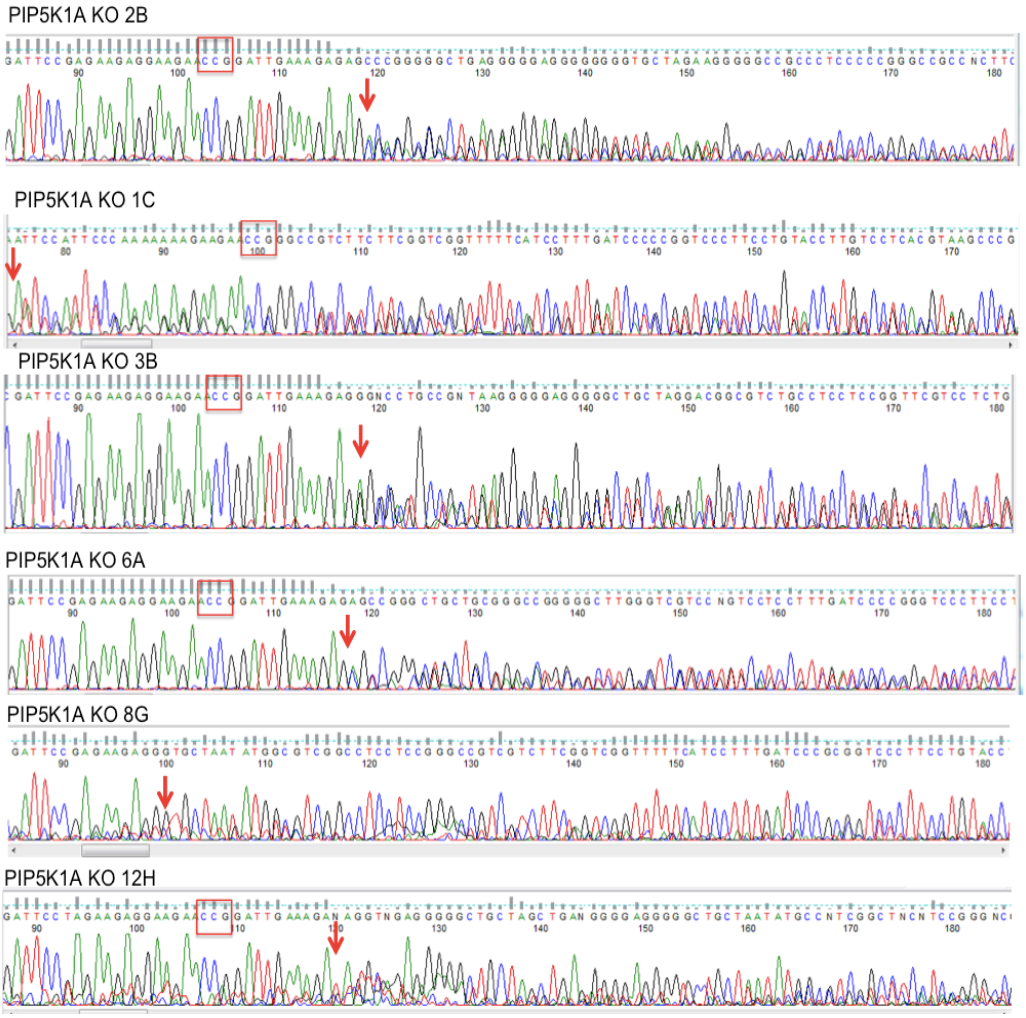

Figure 24

Electropherogram of the PIP5K1A clones that showed mixed sequences. Red squares indicates PAM motif and red arrow indicates approximate boundaries.

Only one of the PIP5K1A targeted clones (10H) gave a clear sequence (Figure 25), suggesting that only a single allele was being analysed. There are several possible explanations for this. It is possible, but unlikely, that this clone has only a single PIP5K1A allele is present, perhaps due to chromosomal loss or a change in the karyotype during the selection. Another explanation could be that all alleles have exactly the same genome editing event. This is also unlikely due to the random nature of NHEJ repair of DSBs, which would be more likely to generate

diffent indels in different alleles. Note that C42 cells have a complex karyotype. The most likely explanation is that the indels generated in other alleles affect the regions corresponding to one or both primers, hence the PCR failed to amplify these alleles.

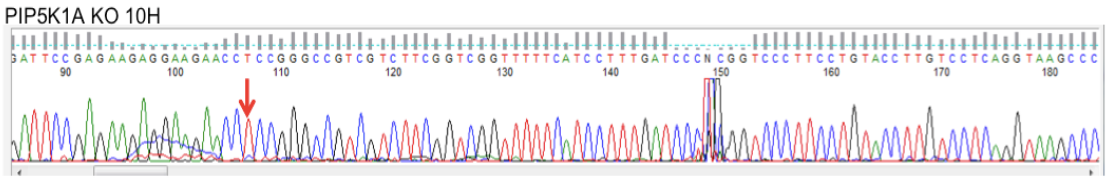

Figure 25

Clone 10H electropherogram. Indicating one sequencing signal.

```
                                                                M  A  S   A  S  S   G  P  S   S  S  V
CGAGAAGAGG AAGAACCGGA TTGAAAGAGA GCCAGGCCGC TGAGGGGGAG GGGGCTGCTA AGATGGCGTC GGCCTCCTCC GGGCCGTCGT CTTCGGTCGG WT
CGAGAAGAGG AAGAACC--- ---------- ---------- ---------- ---------- ---------- -------TCC GGGCCGTCGT CTTCGGTCGG 10H
              PAM                                                   PAM
```

Figure 26

Clone 10H sequence compared to the wild type. Indicating the PAM, the ATG and the expected protein sequence.

The sequencing of PIP5K1A Clone 10H identified a single allele with a 60 bp deletion that has deleted the ATG, and therefore should produce a disruption of the gene.

3.8 Detection of AR and PIP5K1A proteins in targeted clones

After the sequence analysis the next step was the analysis of clones at protein level, to compare the expression of the targeted genes against wild type cells. Western Blots were performed following the same protocol described previously. The antibody used to detect AR and PIP5K1A were N-20 (sc-816) from Santa Cruz Biotechnology and #9693 PIP5K1A antibody from Cell Signaling Technology, respectively. We observed that the AR targeted clones showed a reduced growth rate compared to control C42 cells or the single targeted clone SG-B8. This meant that not all clones reach sufficient cell numbers for analysis at the same time. Thus we performed some westerns on clones as they became available for analysis. Figure 27 shows western analysis of clone AR-B5, which contained an insertion of 276 bp in the 5'UTR region. This result showed a clear reduction in detectable AR protein in comparison to the single guide (WT) clone. Equal loading of the samples was confirmed by detection of beta actin (42 kda).

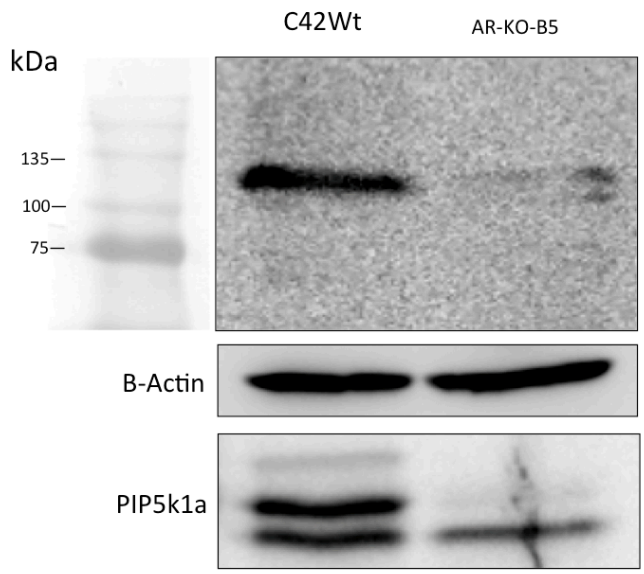

Figure 27

Comparison of the expresion of AR, B-actin and PIP5K1A in the clones AR-KO-B5.

A second Western Blot was performed to repeat this result and confirmed the strong reduction of AR expression in the AR-B5 clone. As the AR gene is on the X chromosome and thus there should be a single allele in male-derived cells, we conclude that the insertion in the 5'UTR disrupts the expression of AR by an unknown mechanism.

We also assessed AR expression in the AR-D5 and AR-D6 clones, however, although somewhat lower than control, the effect on AR levels was not as great as in AR-B5. Thus deletion of a 70bp sequence proximal to the ATG may not be critical for AR expression, although the insertion of the 276 bp sequence appears to have a more deleterious effect.

To assess whether the depletion of AR expression impacts on PIP5K1A, a Western Blot was performed on the same blot. As shown in Figure 27 lower panel, we detected two PIP5K1A bands in the wild type clone of 65 KDa. In contrast, the major PIP5K1A band was absent the AR-B5 clone. The β-actin levels are equal in both samples indicating that the same amount of protein was loaded. This suggests a direct link between AR and PIP5K1A expression. Interestingly, a previous microarray study to detect AR regulated genes in LNCAP cells identified PIP5K1A (50). In that study suppression of AR expression using antisense oligos caused an upregulation of PIP5K1A transcripts, whereas, treatment of the LNCAP cells with bicalutamide, and AR antagonist ligand, had no effect.

PIP5K1A clones analysis

While sequence analysis showed that most of the PIP5K1A targeted clones have more than one allele, several contained alleles in which the ATG was deleted, suggested protein expression would be disrupted. We therefore performed Western Blots on cell free extracts of these clones to assess effects on PIP5K1A and prioritise clones for further analysis.

As a preliminary analysis we chose the clones PI-1c, PI-2C and PI-2f, along with the PI-WT clone. Protein extracts from these clones were equally loaded as confirmed by detection of beta actin. While the wild type clone PI-WT exhibited two bands around 55 and 63 KDa as expected, all three targeted clones showed a reduction in the levels of

PIP5K1A detected (Figure 28). This result is quite consistent with the genotyping results (Figure 17). Clone PI-1C displayed a single band of weaker intensity that migrated between the two bands present in wild type clones. The sequence analysis suggests alterations that may delete the N-terminal domain of one allele. Thus a single truncated allele may be expressed in this clone.

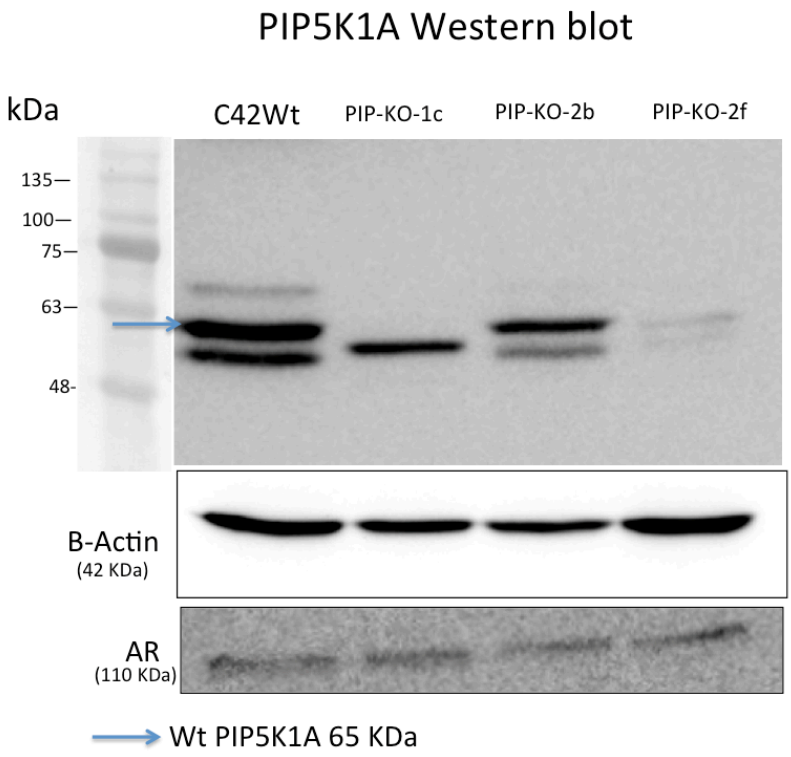

Figure 28

Analysis of the expression of PIP5K1A, AR and B-Actin in PIP5K1A-KO clones.

Clone 2C shows a similar protein expression pattern similar to wild type but shows a reduced overall expression level. Finally PI-2F shows the greatest depletion of PIP5K1A levels. This clone failed to produce PCR products in the genotyping suggesting indels that disrupt the regions

complementary to the primers. A second western was performed which confirmed the results for PI-1C, PI-2F and PI-2B and also detected reduced expression in clone PI-10H, which also carries deletions extending into the ATG sequence. These results confirm the successful editing of the PIP5K1A gene in C42 cells. We also detected PIP5K1A levels in clones PI-3B, PI-3C, PI-6A, PI-8G, PI-10H and PI-12H and found them to have variable expression levels of PIP5K1A. However because of an air bubble on the blot, the signal for the control cells was maksed therefore, we could not compare to levels in the untargeted control. Further work will be required to complete the analysis.

To assess if loss of PIP5K1A expression had any impact on AR expression, Western Blots were performed. We were unable to make a conclusive assessment due to variations in the protein levels. However, our preliminary data suggest that full-length AR levels are not substantially affected by depletion of PIP5K1A levels.

CHAPTER 4

DISCUSSION

4. Discussion

Prostate cancer is the most common non-cutaneous malignancy in men in most developed countries, and a major cause of cancer deaths (1). Different treatment options are available for PCa, including radical prostatectomy, radiation therapy, androgen deprivation and 'watchful waiting' (prostatecanceruk.org). However, approximately one-third of treated PCa patients will experience disease recurrence and will progress into CRPC, which no longer responds to hormone therapy (20)(51).

Hormone therapies can be used to achieve a reduction of androgen levels in the body (which exert their biological effects through the AR) or can block AR function by acting as antagonists that associate with the ligand binding domain (LBD). Both the N-terminal domain and the LBD can function in transcriptional activation and these functions of AR are regulated by androgens. The LBD undergoes a conformational change in response to androgen binding that enables it to recruit coactivators such as SRC1 and GRIP1. In advanced disease e.g. CRPC mutations in the AR gene render it resistant to androgen therapy. Mutations in the LBD permit the AR to be activated by drugs that antagonise the normal receptor. In addition, constitutively active AR variants lacking the ligand binding domain (e.g. AR-V7) are expressed in CRPC. These altered AR proteins continue to drive growth of the tumour thus new therapies are needed.

A recent study indicates that PIP5K1A might play an important role in prostate tumour growth, inducing the activation of AR via an alternative pathway involving PI3K/AKT (22). PIP5K1A is a lipid kinase responsible for the synthesis of PIP2, substrate of PI3K which activates the AKT family

of serine/threonine, inducing cell growth and proliferation. PIP5K1A is overexpressed in prostate tumours and PCa cell lines, and it was shown that a small molecule inhibitor of PIP5K1A (ISA-2011B) can inhibit growth of PCa cells in culture and reduce tumour volume in xenograft studies (22). However, the mechanism behind the interactions between AR and PI3K/AKT remains unknown. In a follow up study it was shown that ISA-2011B selectively reduces expression of the ARv7 variant on 22Rv1 cells, a CRPC cell line that readily forms tumours in xenograft models (23). This suggests a regulatory feedback loop in the expression of AR and PIP5K1A In order to study the underlying mechanisms of AR and PIP5K1A interactions, and to validate the effects of ISA-2011B on PCA cells, we set out to disrupt these genes in LNCAP-C42 cells using CRIPR-Cas9 genome editing technology.

The nickase strategy was used to generate indels close to the ATG in both genes, producing knock-outs in LNCaP C4-2 prostate cancer cells. With this approach we could successfully established mutant PIP5K1A and AR cell lines. CRISPR-mediated events were studied in each mutant, in order to study the genomic alterations in the target region. In AR-targeted cells we established two mutant cell lines with deletions in the 5'UTR region which showed no alteration in AR expression according to our preliminary Western Blot data. This suggests that deletion of the 70bp region directly preceding the ATG has little effect on transcription or translation of AR. However, the AR-targeted mutant cell line (AR- B5) contained a 276 bp insertion in 5'UTR region upstream of the ATG, and showed a strong reduction in AR protein, in several independent Western

Blot experiments. The mechanism is not known but may interfere with translation of the AR transcript.

The 5'UTR region regulates translation by different ways; the uORFs (upstream open reading frames) and uAUGs are the main regulatory elements in the 5'UTR, by an unknown mechanism the number and position of these sequences can disrupt or enhance the translation process (52). Secondary structures are also regulatory tools in 5′ UTR, the hairpin loops formed by the sequences with a high percentage of GC modulates the translation, a change in these structures can be enough to block the access of the preinitiation complex to the mRNA (53). In addition, RNA binding proteins regulates translation by binding to the 5'UTR region, these proteins binds to certain motifs and can increase or inhibit translation (54).

Preliminary analysis indicates expression of PIP5K1A in the AR-B5 clone was reduced, suggesting that PIP5K1A is regulated by AR. This needs to be confirmed by RTqPCR to ascertain if PIP5K1A transcript levels are reduced. For PIP5K1A-targeted C4-2 cells we successfully established nine different mutant cell lines, with different events in each allele, wild type alleles might be still present and further analysis must be required in order to confirm the knock-out. At protein level several clones (AR-B5, PI-1C, PI-2F, PI-10H) showed alteration in expression indicating successfully disruption of the gene. CRISPR-Cas9 genome editing tool has been proved to be able to edit LNCaP C4-2 cells, also the suitability of this cell line for clonal expansion, been able to grow in isolation, has been demonstrated. In order to confirm the interaction between the target genes

at protein level, a preliminary mRNA study by RT-q-PCR in the clones has been done, the data suggest a possible feedback regulatory interaction but further study is required to confirm this.

In terms of future work, further PCR analysis of genomic DNA is required, perhaps expanding the sets of primers used to more precisely identify different alleles. Some of the clones failed and did not produce any product suggesting mutations or deletions in the primer target area caused by the CRISPR-Cas9 mediated event. New primers need to be designed further the targeted region. We can also perform mRNA extraction and cDNA synthesis to confirm the alterations at transcript level. This would involve subcloning the PCR products into pGEMT easy for sequencing of individual alleles. Once the genotype alterations and Western Blot validations are confirmed, RNA-seq analysis to compare wild type transcription profiles with the cells deficient in PIP5K1A and AR expression can be performed. This would discover differentially expressed genes in PIP5K1A and AR knock-out clones. These mutants can provide the information for the understanding of the PIP5K1A/AR regulation, and by the understanding of this pathway new therapeutics target can be discovered to provide treatment options for CRPC patients.

References

1. Attard G, Parker C, Eeles RA, Schroder F, Tomlins SA, Tannock I, *et al*. Prostate cancer. Lancet. 2016 Jan 2;387(10013):70-82. PubMed PMID: 26074382.

2. Pentyala S, Whyard T, Pentyala S, Muller J, Pfail J, Parmar S, *et al*. Prostate cancer markers: An update. Biomedical reports. 2016 Mar;4(3):263-8. PubMed PMID: 26998261. Pubmed Central PMCID: 4774403.

3. Vlaeminck-Guillem V, Ruffion A, Andre J. [Value of urinary PCA3 test for prostate cancer diagnosis]. Progres en urologie : journal de l'Association francaise d'urologie et de la Societe francaise d'urologie. 2008 May;18(5):259-65. PubMed PMID: 18538269. Place du test urinaire PCA3 pour le diagnostic du cancer de la prostate.

4. Gohiji KaK, S. Molecular Mechanism of Prostate cancer invasion and metastasis. Prostate cancer: Science and Clinical Practice. 2003:11-27.

5. Greenberg R. Does prostate cancer represent more than one cancer? Prostate cancer: Science and Clinical Practice. 2003:29-34.

6. Roy AK, Lavrovsky Y, Song CS, Chen S, Jung MH, Velu NK, et al. Regulation of androgen action. Vitamins and hormones. 1999;55:309-52. PubMed PMID: 9949684.

7. Heinlein CA, Chang C. Androgen receptor in prostate cancer. Endocrine reviews. 2004 Apr;25(2):276-308. PubMed PMID: 15082523.

8. Russell DW, Wilson JD. Steroid 5 alpha-reductase: two genes/two enzymes. Annual review of biochemistry. 1994;63:25-61. PubMed PMID: 7979239.

9. Nuclear Receptors Nomenclature C. A unified nomenclature system for the nuclear receptor superfamily. Cell. 1999 Apr 16;97(2):161-3. PubMed PMID: 10219237.

10. Tsai MJ, O'Malley BW. Molecular mechanisms of action of steroid/thyroid receptor superfamily members. Annual review of biochemistry. 1994;63:451-86. PubMed PMID: 7979245.

11. Tan MHE, Li J, Xu HE, Melcher K, Yong EL. Androgen receptor: structure, role in prostate cancer and drug discovery. Acta Pharmacol Sin. 2015 Jan;36(1):3-23. PubMed PMID: WOS:000347389000002. English.

12. Aranda A, Pascual A. Nuclear hormone receptors and gene expression. Physiological reviews. 2001 Jul;81(3):1269-304. PubMed PMID: 11427696.

13. Huang P, Chandra V, Rastinejad F. Structural overview of the nuclear receptor superfamily: insights into physiology and therapeutics. Annual review of physiology. 2010;72:247-72. PubMed PMID: 20148675. Pubmed Central PMCID: 3677810.

14. Gao W, Bohl CE, Dalton JT. Chemistry and structural biology of androgen receptor. Chemical reviews. 2005 Sep;105(9):3352-70. PubMed PMID: 16159155. Pubmed Central PMCID: 2096617.

15. Marhefka CA, Moore BM, 2nd, Bishop TC, Kirkovsky L, Mukherjee A, Dalton JT, *et al*. Homology modeling using multiple molecular dynamics simulations and docking studies of the human androgen receptor ligand

binding domain bound to testosterone and nonsteroidal ligands. Journal of medicinal chemistry. 2001 May 24;44(11):1729-40. PubMed PMID: 11356108.

16. Tyagi RK, Lavrovsky Y, Ahn SC, Song CS, Chatterjee B, Roy AK. Dynamics of intracellular movement and nucleocytoplasmic recycling of the ligand-activated androgen receptor in living cells. Molecular endocrinology. 2000 Aug;14(8):1162-74. PubMed PMID: 10935541.

17. Buchanan G, Irvine RA, Coetzee GA, Tilley WD. Contribution of the androgen receptor to prostate cancer predisposition and progression. Cancer metastasis reviews. 2001;20(3-4):207-23. PubMed PMID: 12085963.

18. Cunha GR, Donjacour AA, Cooke PS, Mee S, Bigsby RM, Higgins SJ, et al. The endocrinology and developmental biology of the prostate. Endocrine reviews. 1987 Aug;8(3):338-62. PubMed PMID: 3308446.

19. Denis LJ, Griffiths K. Endocrine treatment in prostate cancer. Seminars in surgical oncology. 2000 Jan-Feb;18(1):52-74. PubMed PMID: 10617897.

20. Montgomery RB, Mostaghel EA, Vessella R, Hess DL, Kalhorn TF, Higano CS, et al. Maintenance of intratumoral androgens in metastatic prostate cancer: a mechanism for castration-resistant tumor growth. Cancer research. 2008 Jun 1;68(11):4447-54. PubMed PMID: 18519708. Pubmed Central PMCID: 2536685.

21. Qu Y, Dai B, Ye D, Kong Y, Chang K, Jia Z, et al. Constitutively active AR-V7 plays an essential role in the development and progression

of castration-resistant prostate cancer. Scientific reports. 2015;5:7654. PubMed PMID: 25563505. Pubmed Central PMCID: 4288210.

22. Semenas J, Hedblom A, Miftakhova RR, Sarwar M, Larsson R, Shcherbina L, et al. The role of PI3K/AKT-related PIP5K1alpha and the discovery of its selective inhibitor for treatment of advanced prostate cancer. Proceedings of the National Academy of Sciences of the United States of America. 2014 Sep 2;111(35):E3689-98. PubMed PMID: 25071204. Pubmed Central PMCID: 4156761.

23. Sarwar M, Semenas J, Miftakhova R, Simoulis A, Robinson B, Wingren AG, et al. Targeted suppression of AR-V7 using PIP5K1alpha inhibitor overcomes enzalutamide resistance in prostate cancer cells. Oncotarget. 2016 Aug 31. PubMed PMID: 27588408.

24. van den Bout I, Divecha N. PIP5K-driven PtdIns(4,5)P2 synthesis: regulation and cellular functions. Journal of cell science. 2009 Nov 1;122(Pt 21):3837-50. PubMed PMID: 19889969.

25. Hokin MR, Hokin LE. Enzyme secretion and the incorporation of P32 into phospholipides of pancreas slices. The Journal of biological chemistry. 1953 Aug;203(2):967-77. PubMed PMID: 13084667.

26. Halstead JR, Roefs M, Ellson CD, D'Andrea S, Chen C, D'Santos CS, et al. A novel pathway of cellular phosphatidylinositol(3,4,5)-trisphosphate synthesis is regulated by oxidative stress. Current biology : CB. 2001 Mar 20;11(6):386-95. PubMed PMID: 11301249.

27. Loijens JC, Anderson RA. Type I phosphatidylinositol-4-phosphate 5-kinases are distinct members of this novel lipid kinase family. The

Journal of biological chemistry. 1996 Dec 20;271(51):32937-43. PubMed PMID: 8955136.

28. Weernink PA, Meletiadis K, Hommeltenberg S, Hinz M, Ishihara H, Schmidt M, et al. Activation of type I phosphatidylinositol 4-phosphate 5-kinase isoforms by the Rho GTPases, RhoA, Rac1, and Cdc42. The Journal of biological chemistry. 2004 Feb 27;279(9):7840-9. PubMed PMID: 14681219.

29. Kunz J, Wilson MP, Kisseleva M, Hurley JH, Majerus PW, Anderson RA. The activation loop of phosphatidylinositol phosphate kinases determines signaling specificity. Molecular cell. 2000 Jan;5(1):1-11. PubMed PMID: 10678164.

30. Xie Y, Zhu L, Zhao G. Assignment of type I phosphatidylinositol-4-phosphate 5-kinase (PIP5K1A) to human chromosome bands 1q22--> q24 by in situ hybridization. Cytogenetics and cell genetics. 2000;88(3-4):197-9. PubMed PMID: 10828584.

31. Shaw RJ, Cantley LC. Ras, PI(3)K and mTOR signalling controls tumour cell growth. Nature. 2006 May 25;441(7092):424-30. PubMed PMID: 16724053.

32. Hennessy BT, Smith DL, Ram PT, Lu Y, Mills GB. Exploiting the PI3K/AKT pathway for cancer drug discovery. Nature reviews Drug discovery. 2005 Dec;4(12):988-1004. PubMed PMID: 16341064.

33. Saito K, Tolias KF, Saci A, Koon HB, Humphries LA, Scharenberg A, et al. BTK regulates PtdIns-4,5-P2 synthesis: importance for calcium signaling and PI3K activity. Immunity. 2003 Nov;19(5):669-78. PubMed PMID: 14614854.

34. Semenas J, Dizeyi N, Persson JL. Enzalutamide as a second generation antiandrogen for treatment of advanced prostate cancer. Drug design, development and therapy. 2013;7:875-81. PubMed PMID: 24009414. Pubmed Central PMCID: 3762760.

35. Wang Y, Kreisberg JI, Ghosh PM. Cross-talk between the androgen receptor and the phosphatidylinositol 3-kinase/Akt pathway in prostate cancer. Current cancer drug targets. 2007 Sep;7(6):591-604. PubMed PMID: 17896924.

36. Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, et al. CRISPR provides acquired resistance against viruses in prokaryotes. Science. 2007 Mar 23;315(5819):1709-12. PubMed PMID: 17379808.

37. Rath D, Amlinger L, Rath A, Lundgren M. The CRISPR-Cas immune system: biology, mechanisms and applications. Biochimie. 2015 Oct;117:119-28. PubMed PMID: 25868999.

38. Sinkunas T, Gasiunas G, Fremaux C, Barrangou R, Horvath P, Siksnys V. Cas3 is a single-stranded DNA nuclease and ATP-dependent helicase in the CRISPR/Cas immune system. The EMBO journal. 2011 Apr 6;30(7):1335-42. PubMed PMID: 21343909. Pubmed Central PMCID: 3094125.

39. Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. Genome engineering using the CRISPR-Cas9 system. Nature protocols. 2013 Nov;8(11):2281-308. PubMed PMID: 24157548. Pubmed Central PMCID: 3969860.

40. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science. 2012 Aug 17;337(6096):816-21. PubMed PMID: 22745249.

41. Hsu PD, Lander ES, Zhang F. Development and applications of CRISPR-Cas9 for genome engineering. Cell. 2014 Jun 5;157(6):1262-78. PubMed PMID: 24906146. Pubmed Central PMCID: 4343198.

42. Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, et al. RNA-guided human genome engineering via Cas9. Science. 2013 Feb 15;339(6121):823-6. PubMed PMID: 23287722. Pubmed Central PMCID: 3712628.

43. Tai S, Sun Y, Squires JM, Zhang H, Oh WK, Liang CZ, et al. PC3 is a cell line characteristic of prostatic small cell carcinoma. The Prostate. 2011 Nov;71(15):1668-79. PubMed PMID: 21432867. Pubmed Central PMCID: 3426349.

44. Liu AY, Brubaker KD, Goo YA, Quinn JE, Kral S, Sorensen CM, et al. Lineage relationship between LNCaP and LNCaP-derived prostate cancer cell lines. The Prostate. 2004 Jul 1;60(2):98-108. PubMed PMID: 15162376.

45. Wu HC, Hsieh JT, Gleave ME, Brown NM, Pathak S, Chung LW. Derivation of androgen-independent human LNCaP prostatic cancer cell sublines: role of bone stromal cells. International journal of cancer. 1994 May 1;57(3):406-12. PubMed PMID: 8169003.

46. Hornig NC, de Beaufort C, Denzer F, Cools M, Wabitsch M, Ukat M, et al. A Recurrent Germline Mutation in the 5'UTR of the Androgen

Receptor Causes Complete Androgen Insensitivity by Activating Aberrant uORF Translation. PloS one. 2016;11(4):e0154158. PubMed PMID: 27110943. Pubmed Central PMCID: 4844194.

47. Chang JH, Wang LF, Lu SL, Huang CF, Lu HK. NF-kappaB Enhances Androgen Receptor Expression through 5'-UTR Binding in Gingival Cells. Journal of dental research. 2015 Oct;94(10):1439-45. PubMed PMID: 26152187.

48. Valdez CD, Davis JN, Odeh HM, Layfield TL, Cousineau CS, Berton TR, et al. Repression of androgen receptor transcription through the E2F1/DNMT1 axis. PloS one. 2011;6(9):e25187. PubMed PMID: 21966451. Pubmed Central PMCID: 3180375.

49. Hay CW, Hunter I, MacKenzie A, McEwan IJ. An Sp1 Modulated Regulatory Region Unique to Higher Primates Regulates Human Androgen Receptor Promoter Activity in Prostate Cancer Cells. PloS one. 2015;10(10):e0139990. PubMed PMID: 26448047. Pubmed Central PMCID: 4598089.

50. Eder IE, Haag P, Basik M, Mousses S, Bektic J, Bartsch G, et al. Gene expression changes following androgen receptor elimination in LNCaP prostate cancer cells. Molecular carcinogenesis. 2003 Aug;37(4):181-91. PubMed PMID: 12891627.

51. Zong Y, Goldstein AS. Adaptation or selection--mechanisms of castration-resistant prostate cancer. Nature reviews Urology. 2013 Feb;10(2):90-8. PubMed PMID: 23247694.

52. Araujo PR, Yoon K, Ko D, Smith AD, Qiao M, Suresh U, et al. Before It Gets Started: Regulating Translation at the 5' UTR. Comparative

and functional genomics. 2012;2012:475731. PubMed PMID: 22693426. Pubmed Central PMCID: 3368165.

53. Davuluri RV, Suzuki Y, Sugano S, Zhang MQ. CART classification of human 5' UTR sequences. Genome research. 2000 Nov;10(11):1807-16. PubMed PMID: 11076865. Pubmed Central PMCID: 310970.

54. Abaza I, Gebauer F. Trading translation with RNA-binding proteins. Rna. 2008 Mar;14(3):404-9. PubMed PMID: 18212021. Pubmed Central PMCID: 2248257.